

Levelized Product Cost: Concept and Decision Relevance

Stefan Reichelstein*

Graduate School of Business

Stanford University

Anna Rohlfig-Bastian

WHU - Otto Beisheim School of Management

November 2013

*Email addresses: reichelstein@stanford.edu; anna.rohlfing@whu.edu

We are grateful to seminar participants at Carnegie Mellon, CUNY, Tübingen, the “Ausschuss Unternehmensrechnung” of the “Verein für Socialpolitik” and the University of Alberta Accounting Research Conference for helpful comments, with particular thanks to Gunther Friedl, Hans-Ulrich Küpper, Alexander Nezlobin and Ansu Sahoo.

Levelized Product Cost: Concept and Decision Relevance

Abstract:

This paper examines a life-cycle cost concept that applies to both manufacturing and service industries in which upfront capacity investments are essential. Borrowing from the energy literature, we refer to this cost measure as the levelized product cost (LC). Per unit of output, the levelized cost aggregates a share of the initial capacity expenditure with periodic fixed- and variable operating costs. The resulting cost figure exceeds the full cost of a product, as commonly calculated in managerial accounting. Our analysis shows that the LC can be interpreted as the long-run marginal product cost. In particular, this cost measure is shown to be the relevant unit cost that firms should impute for investments in productive capacity.

1 Introduction

Numerous surveys and studies have suggested that managers frequently rely on a product's full cost for a range of planning decisions, including product pricing.¹ Some economists have decried this practice as “irrational,” arguing that fixed costs are typically sunk and therefore irrelevant for decision making. In particular, many industrial organization models point to the issue of double marginalization when managers internalize a cost higher than marginal cost (Sutton 1991, p. 28). The continued use of full cost for pricing and other short-run decisions has prompted some economists to advocate behavioral models as alternatives to the paradigm of fully rational decision-making in order to explain managerial practices (Al-Najjar, Baliga, and Besanko 2008).

At the same time, the identification and measurement of marginal cost remains controversial in economics (Pittman 2009; Carlton and Perloff 2005). This measurement issue has been particularly prominent in antitrust investigations of excess profitability and monopoly pricing. Surveying different strands of the industrial organization literature, Pittman (2009) notes that many economic studies have used variable production cost as an effective proxy for marginal cost. Pittman succinctly summarizes the resulting tension as follows: “*It is difficult to understand how a firm that sets prices at true marginal cost is able to survive as a going concern unless that true marginal cost includes the marginal cost of capital.*”² This paper confirms the notion that from a long-run planning perspective the “true marginal cost” must include the cost of capital (capacity) as well as periodic fixed costs to be incurred in later periods. We demonstrate that this long-run marginal cost exceeds the measure of full cost as commonly calculated in managerial accounting.

Accounting textbooks, such as Horngren, Datar, and Rajan (2012) or Zimmerman (2010), portray the importance of full cost as a benchmark for breaking even in terms of accounting profit. The corresponding measure of full cost typically includes charges related to depreciation and fixed overhead costs. For short-run decisions, like product pricing, these textbooks

¹According to the survey by Govindarajan and Anthony (1983) more than 80% of respondents use a product's full cost when setting list prices. Similar findings were obtained in later studies by Shim and Sudit (1995), Drury, Osborne, and Tayles (1993), and Bouwens and Steens (2008). The extensive survey evidence on intracompany pricing also points to full cost as the prevalent basis for valuing transfers of goods and services in vertically integrated firms; see, for instance, Eccles (1985).

²Borenstein (2008) suggests that economies of scale can defuse the tension between marginal cost pricing and non-negative economic profits.

advocate the use of *incremental costs* which typically exclude fixed costs but include variable cost and applicable opportunity costs. For long-run decisions, such as those in plant, property and equipment, accounting textbooks generally do not advocate an accrual-based measure as the relevant cost. Instead they follow the standard corporate finance prescription of evaluating the stream of discounted cash flows associated with a particular decision. While this approach is uncontroversial, it does not point to a measure of relevant cost for capacity investment decisions.

The central concept of this paper is the so-called Levelized Cost (LC). Borrowing from the energy literature, which coined the term Levelized Cost of Electricity, this concept is a formalization of the following verbal definition: “*the levelized cost of electricity is the constant dollar electricity price that would be required over the life of the plant to cover all operating expenses, payment of debt and accrued interest on initial project expenses, and the payment of an acceptable return to investors*” (MIT 2007). We note that LC is defined entirely in terms of cash flows and calibrated as the minimum price that investors would have to receive on average in order to break-even in terms of discounted cash flows.³

The first part of our analysis demonstrates that the LC can be calculated as a *fully loaded cost* in a way that is similar to how accounting textbooks typically determine the full cost of a product.⁴ This alignment, however, requires depreciation to be calculated so that book values reflect the replacement cost value of assets (Rogerson 2008). In addition, the measure of full cost has to include capital charges on the remaining book value of the capacity assets and the overall capacity charge must be marked up by a tax factor that reflects the delayed amortization of the investment for income tax purposes. The inclusion of imputed interest charges and the tax factor cause our measure of fully loaded cost to exceed the usual concept of full cost in management accounting.

³The LC concept is naturally related to the notion of life cycle costing in the cost accounting literature; see, for instance, Horngren, Datar, and Rajan (2012), Atkinson et al. (2011), Coenenberg, Fischer, and Günther (2012), and Ewert and Wagenhofer (2008). All of these approaches seek to aggregate product related costs over different stages of the product’s life cycle. In contrast, our LC concept is concerned with the cost of producing a good or service at a particular facility.

⁴In a similar vein, Küpper (1985; 2009) advocates for cost accounting to provide cost measures that can be used for investment decisions. To that end, Küpper also develops an accrual-based cost metric that is consistent with present value considerations. Küpper (2009) demonstrates the use of this concept in the context of multiple long-term decision problems including production planning and the identification of price floors for individual products.

We then examine the economic relevance of the LC concept for alternative market settings including those where firms can set prices and those where competition forces price taking behavior. Given the break-even conceptualization of the levelized cost, one might expect that in a competitive market with price taking firms the equilibrium price will, “on average,” be equal to the levelized product cost. We confirm that in a competitive equilibrium, the present value of future expected market prices is equal to the annuity value of the levelized product cost, where the annuity is taken over the life-cycle of the productive facility. Furthermore, in a stationary environment, where the distribution of market prices and production costs do not change over time, the expected equilibrium market price will be equal to the LC in each period. This finding confirms our interpretation of the levelized cost as the long-run marginal product cost. It should be noted, though, that unlike the usual microeconomic textbook description, long-run and short-run marginal cost will not coincide in equilibrium due to the presence of capacity constraints. In fact, the former will always be below the latter for any output level below the capacity limit.

The presence of capacity constraints will prevent prices from being bid down to the variable cost of production, even in a competitive market. The aggregate capacity level in equilibrium is shown to depend on the degree of price volatility in the product market. Our notion of *limited price volatility* is that the maximal percentage deviation from the average market price (holding quantity fixed) does not exceed the ratio of the unit variable cost to the levelized product cost. This condition is more likely to be satisfied in capital intensive industries. If this condition is met, firms in the industry will in equilibrium always deploy the entire available capacity, even for unfavorable shocks to market demand, and the aggregate capacity level will correspond to the expected demand at the market price corresponding to the LC.

With *significant price volatility*, we find that the aggregate capacity in equilibrium will be *larger* than that obtained in a setting with limited volatility. While this finding may seem counter-intuitive at first glance, the argument is that firms retain the option of idling parts of that capacity in subsequent periods under unfavorable market conditions. Only that portion of the LC that corresponds to the capacity cost is a sunk cost. Since the market price will not fall below the short-run marginal cost, that is, the unit variable cost of production, the payoff structure associated with a capacity investment for firms effectively looks like a call option and this option becomes more valuable with significant price volatility.

For a firm with monopoly power in a stationary product market, we demonstrate that the optimal capacity level satisfies the following condition: the *expected* marginal revenue of output chosen optimally in each subsequent period, subject to the constraint imposed by the initial capacity choice, is equal to the levelized cost, LC. At the same time, the marginal revenue corresponding to full capacity utilization will generally be below the LC, unless the condition of limited price volatility is met. Accordingly, the capacity level at which the expected marginal revenue is equal to the LC constitutes a lower bound for the optimal capacity level. One obtains a corresponding upper bound by imputing the *levelized fixed cost*, defined as the LC less the unit variable cost of production. This adjustment reflects that the variable cost portion of the LC is not a sunk cost in subsequent periods.

The pattern of results we obtain for competitive industries and monopolies extends to oligopolistic competition. In particular, we consider a setting in which two firms choose their output levels in a standard Cournot fashion in each period, given their variable production costs and the constraints imposed by the initial capacity choices. For the first stage capacity decisions it is then a (subgame perfect) Nash equilibrium outcome for each firm to choose a capacity level at which the expected marginal revenue of output in future periods is equal to the LC.

Our characterization of the levelized cost as the long-run marginal cost is consistent with the findings in a number of recent studies (Rogerson 2008; Rajan and Reichelstein 2009; Nezlobin 2012; Nezlobin, Rajan, and Reichelstein 2012). In contrast to our model framework, these studies rely on overlapping capacity investments in an infinite horizon setting. Capacity can be added continuously on an “as needed basis.” Furthermore, market demand is assumed to expand over time and there are no periodic price shocks. As a consequence, firms never find themselves in a position of excess capacity. Our setting is motivated by the observation that in many settings of interest, initial capacity investments will turn out to be excessive for unfavorable realizations of market demand.

The main theme in this paper is also directly related to a branch of the managerial accounting literature that has sought to provide a rationale for full cost pricing.⁵ This literature concludes that the sufficiency of full cost for product pricing depends on several conditions, including the timing of the pricing decision relative to the point in time when the firm commits to capacity resources. Other conditions include whether capacity constraints

⁵Papers in this literature include Banker and Hughes (1994) Balakrishnan and Sivaramakrishnan (2001; 2002), Göx (2002) and Banker, Hwang, and Mishra (2002).

are “soft” and whether firms learn additional information about the product market after deciding on capacity levels. While our findings are consistent with this earlier literature, the focus of our analysis is not on product pricing. Our primary goal is to identify a unit cost measure that can effectively serve as the marginal cost for initial capacity investments. The initial capacity choice, of course, anticipates the subsequent volume and pricing of the product in response to subsequent market conditions, including the overall capacity levels available in the industry. The resulting product prices are shown to be equal to the levelized product cost plus some mark-up that varies with the extent of competition in the industry. In particular, the mark-up on levelized cost is shown to be zero under atomistic competition, and decreasing in the number of firms in the industry.

The remainder of the paper is organized as follows. Section 2 formalizes the Levelized Product Cost (LC) concept and establishes how it relates to the customary measure of full cost. Section 3 analyzes the equilibrium price and aggregate capacity level in a competitive market setting with price-taking firms. We consider a market structure with price-setting firms, in particular monopoly and duopoly, in Section 4. Section 5 provides conclusions and proofs are relegated to the Appendix.

2 Levelized Product Cost versus Full Cost

2.1 The LC Concept

The levelized cost of a product or service seeks to identify a per unit break-even value that a producer would need to obtain as sales revenue in order to justify an investment in a particular production facility. In the context of electricity generation, the Levelized Cost of Electricity concept is widely used by academic and business analysts to compare the cost effectiveness of alternative energy sources which differ substantially in terms of upfront investment cost and periodic operating costs. As noted in Reichelstein and Yorston (2013), however, the formulaic implementation of this concept has been lacking in uniformity.⁶

We formalize the levelized product cost concept in the context of a generic capacity investment in a new production facility. The levelized cost of one unit of output at this

⁶In particular, some authors have conceptualized the Levelized Cost of Electricity as the ratio of “total lifetime cost” to “total lifetime electricity produced” (EPIA September 2011; Campbell 2008, 2011; Werner 2012). This turns out to be generally incompatible with breaking-even over the life-cycle of the project as articulated in the verbal definition of the MIT (2007) study quoted in the Introduction.

facility aggregates the upfront capacity investment, the sequence of output levels generated by the facility over its useful life, the periodic operating costs required to deliver the output in each period, and any tax related cash flows that apply to this type of facility. We treat the applicable cost of capital, r , as exogenous and denote the corresponding discount factor by $\gamma \equiv \frac{1}{1+r}$.⁷

Investment in the production facility may entail economies of scale. In particular,

- $v(k)$: the cost of installing k units capacity. We normalize units so that one unit of capacity can produce one unit of output in the initial year of operation.
- T : the useful life of the output generating facility (in years) .
- x_t : the capacity decline factor: the percentage of initial capacity that is functional in year t . Production in year t is limited to $q_t \leq x_t \cdot k$.

The capacity decline factor reflects that in certain contexts, the available output yield changes over time. For instance, with photovoltaic solar cells it has been observed that their efficiency diminishes over time. The corresponding decay is usually represented as a constant percentage factor ($x_t = x^{t-1}$ with $x \leq 1$) which varies with the particular technology (Reichelstein and Yorston 2013). On the other hand, production processes requiring chemical balancing frequently exhibit yield improvements over time due to learning-by-doing effects, e.g., semiconductors and biochemical production processes. The analysis in this paper will pay particular attention to the “one-hoss shay” asset productivity scenario, in which the facility has undiminished capacity throughout its useful life, that is $x_t = 1$ for all $1 \leq t \leq T$ and thereafter the facility is obsolete.⁸

The unit cost of installed capacity, $v(k)$, represents a joint cost of acquiring one unit of capacity for T years. In order to obtain *the cost of capacity for one unit of output*, the joint cost $v(k)$ will be divided by the present value term $\sum_{t=1}^T x_t \cdot \gamma^t$ and the units of capacity installed, k :

⁷If the firm’s leverage ratio is held constant, it is well known from standard corporate finance that, in reference to the above quote in the MIT study, equity holders will receive an “*acceptable return*” and debt holders will receive “*accrued interest on initial project expenses*” provided the project achieves a zero Net Present Value (NPV) when evaluated at the Weighted Average Cost of Capital (WACC); see, for instance, Ross, Westerfield, and Jaffe 2005.

⁸The one-hoss shay scenario is commonly considered in the regulation literature, see, for instance, Laffont and Tirole (2000) and Rogerson (2011).

$$c(k) = \frac{v(k)}{k \cdot \sum_{t=1}^T x_t \cdot \gamma^t}. \quad (1)$$

We shall refer to $c(k)$ as the *unit cost of capacity*. Absent any other operating costs or taxes, $c(k)$ would yield the break-even price identified in the verbal definition above. To illustrate, suppose the firm makes an initial capacity investment of $\$v(k)$ and therefore has the capacity to deliver $q_t = x_t \cdot k$ units of product in year t . If the revenue per unit is c , then revenue in year t would be $c \cdot x_t \cdot k$ and the firm would exactly break even on its initial investment over the T -year horizon.⁹

In addition to the initial investment expenditure $v(k)$, the firm may incur periodic fixed operating costs. The notation, $F_t(k)$, indicates that the magnitude of these costs may vary with the scale of the initial capacity investment. Applicable examples here include insurance, maintenance expenditures, and property taxes. Unless otherwise indicated, we assume that the firm will incur the fixed operating cost $F_t(k)$ regardless of the output level q_t it produces in period t .¹⁰ The initial investment in capacity triggers a stream of future fixed costs and a corresponding stream of future (expected) output levels. By taking the ratio of these, we obtain the following time-averaged fixed operating costs per unit of output:

$$f(k) \equiv \frac{\sum_{t=1}^T F_t(k) \cdot \gamma^t}{k \cdot \sum_{t=1}^T x_t \cdot \gamma^t} \quad (2)$$

With regard to variable production costs, we assume a constant returns to scale technology in the short run, so that the variable costs per unit of production up to the capacity limit are constant in each period, though they may vary over time. We again define a time-averaged unit variable cost by:

$$w \equiv \frac{\sum_{t=1}^T w_t \cdot x_t \cdot k \cdot \gamma^t}{k \cdot \sum_{t=1}^T x_t \cdot \gamma^t}. \quad (3)$$

⁹Throughout this section, it will be assumed that the available capacity is fully exhausted in each period, that is $q_t = x_t \cdot k$. This specification will no longer apply in Sections 3 and 4, where uncertainty and demand shocks are introduced.

¹⁰We will also consider an alternative scenario wherein the cost $F_t(k)$ is incurred only if $q_t > 0$. Thus, the firm can *avoid* the fixed operating cost $F_t(k)$ in period t if it idles the production facility in that period.

Corporate income taxes affect the levelized cost measure through depreciation tax shields and debt tax shields, as both interest payments on debt and depreciation charges reduce the firm's taxable income. While the debt related tax shield is already incorporated into the calculation of the firm's discount rate, the depreciation tax shield is determined jointly by the effective corporate income tax rate and the allowable depreciation schedule for the facility. These variables are represented as:

- α : the effective corporate income tax rate (in %),
- \hat{T} : the facility's useful life for tax purposes (in years), which is usually shorter than the projected economic life, i.e., $\hat{T} < T$, and
- \hat{d}_t : the allowable tax depreciation charge in year t , as a % of the initial asset value $v(k)$.

For the purposes of calculating the levelized product cost, the effect of income taxes can be summarized by a *tax factor* which amounts to a “mark-up” on the unit cost of capacity, $c(k)$.

$$\Delta = \frac{1 - \alpha \cdot \sum_{t=1}^{\hat{T}} \hat{d}_t \cdot \gamma^t}{1 - \alpha}. \quad (4)$$

Since the assumed useful life for tax purposes generally satisfies $\hat{T} < T$, we will from hereon simply refer to T as the useful life with the understanding that $\hat{d}_t = 0$ for $\hat{T} \leq t \leq T$. The tax factor Δ exceeds 1 but is bounded above by $\frac{1}{1-\alpha}$ if $t = 0$.¹¹ It is readily verified that Δ is increasing and convex in the tax rate α . Holding α constant, a more accelerated tax depreciation schedule tends to lower Δ closer to 1. In particular, Δ would be equal to 1 if the tax code were to allow for full expensing of the investment immediately, that is, $\hat{d}_0 = 1$ and $\hat{d}_t = 0$ for $t > 0$.¹²

Combining the preceding components, we are now in a position to state the following characterization of the levelized product cost:

¹¹To illustrate, for a corporate income tax rate of 35%, and a tax depreciation schedule corresponding to the double declining balance rule over 20 years, the tax factor would amount to roughly $\Delta = 1.3$.

¹²For investments in solar power, the U.S. federal tax code allows for a 30% Investment Tax Credit and a five-year accelerated depreciation schedule. The effect of these tax subsidies is to lower the tax factor from about 1.3 to about .7 with a major effect on the corresponding LC since for solar installations virtually all costs are capacity related.

Proposition 0 *The Levelized Cost (LC) is given by*

$$LC(k) = w + f(k) + c(k) \cdot \Delta \quad (5)$$

with $c(k)$, w , $f(k)$ and Δ as given in (1) - (4).

To see that the expression in (5) does indeed satisfy the verbal break-even definition provided above, let p denote the unit sales price. Figure 1 illustrates the sequence of annual pre-tax cash flows and annual operating incomes subject to taxation.

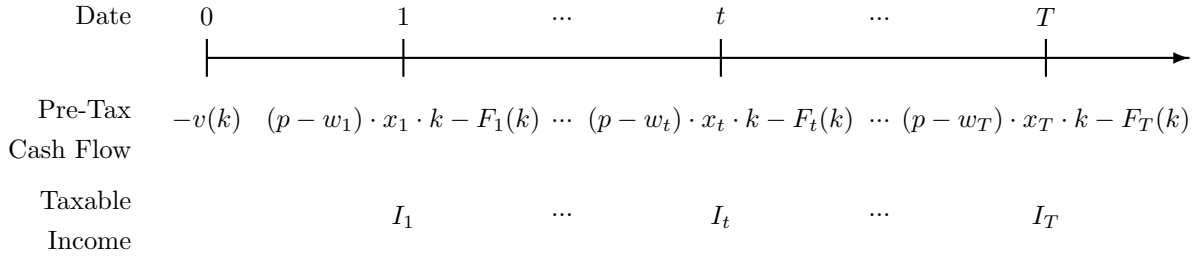


Figure 1: *Cash Flows and Taxable Income*

Taxable income in period t is given by the contribution margin in the respective period minus fixed operating costs minus the depreciation expense allowable for tax purposes:

$$I_t = (p - w_t) \cdot x_t \cdot k - F_t(k) - d_t^o \cdot v(k).$$

As the firm pays an α share of its taxable income as corporate income tax, the annual *after-tax* cash flows, CF_t , are:

$$CF_0 = -v(k)$$

and

$$CF_t = (p - w_t) \cdot x_t \cdot k - F_t(k) - \alpha \cdot I_t.$$

In order for the firm to break even on this investment, the product price p must be such that the present value of all after-tax cash flows is zero. Solving the corresponding linear equation yields $p = LC$. In conclusion, the levelized product cost is a break-even price which includes three principal components: the (time-averaged) fixed operating cost per

unit of output produced, $f(k)$, the (time-averaged) unit variable cost, w , and the unit cost of capacity, $c(k)$, marked-up by the tax factor Δ . It should be noted that the LC identified in Proposition 0 is entirely a cash flow concept. Depreciation enters only through the cash flow corresponding to the depreciation tax shield.

2.2 Relation to Full Cost

In the cost accounting literature, the concept of full cost is usually articulated as a unit cost measure that comprises variable production cost plus (allocated) overhead costs. Overhead costs, in turn, usually include both fixed and variable components. These components may contain accruals that arise due to cash expenditures being allocated cross-sectionally across products or inter-temporally across time periods, e.g., depreciation charges. This section explores to what extent a properly constructed measure of full cost can align with the levelized product cost identified in Proposition 0. In the context of a single-product firm, the *traditional* measure of full cost in period t would be:

$$\hat{FC}_t(k) = \frac{w_t \cdot q_t + F_t(k) + d_t \cdot v(k)}{q_t}, \quad (6)$$

where q_t is the quantity produced in period t , d_t is the percentage depreciation charge in period t , which may of course differ from the charge that is applicable for tax purposes. The total depreciation charge in period t is given by $d_t \cdot v(k) = d_t \cdot v \cdot k$, with $\sum d_t = 1$. It is intuitively obvious that a traditional full cost measure cannot capture the LC since the latter is a discounted cash flow concept, yet (6) does not take the time value of money into consideration. This consideration leads to the following measure of *fully loaded cost*:

$$FC_t(k) = \frac{w_t \cdot q_t + F_t(k) + [d_t + r \cdot (1 - \sum_{i=1}^{t-1} d_i)] \cdot v(k) \cdot \Delta}{q_t}. \quad (7)$$

By construction, $FC_t > \hat{FC}_t$ in each period because the former concept includes an imputed capital charge on the remaining book value and a mark-up corresponding to the tax factor.

Proposition 1 *Suppose the asset's productivity profile conforms to the one-hoss shay scenario ($x_t = 1$), $w_t = w$, and $F_t(k) = f \cdot k$. With full capacity utilization, that is, $q_t = x_t \cdot k$, fully loaded cost is equal to the levelized product cost in each period, that is,*

$$FC_t(k) = LC(k),$$

provided depreciation is calculated according to the annuity method, that is, the depreciation schedule $\{d_t\}$ satisfies $d_{t+1} = d_t \cdot (1 + r)$.

The claim in Proposition 1 relies on the well-known observation that with annuity depreciation:

$$d_t + r \cdot (1 - \sum_{i=1}^{t-1} d_i) = \frac{1}{\sum_{i=1}^T \gamma^i}.$$

As a consequence, the last term in the numerator of (7) is equal to $c(k) \cdot \Delta$, establishing the claim.

The conditions for Proposition 1 appear rather restrictive. A more general result can be obtained provided full cost is calculated on the basis of additional accrual accounting concepts. First, suppose that either the unit variable costs w_t or the unit fixed operating costs $\frac{F_t(k)}{k}$ change over time. In order for the fully loaded cost FC_t to align with LC , variable and fixed operating costs can no longer be recognized on a cash basis but instead their overall present value must be prorated across time periods by means of accruals.¹³

Secondly, the one hoss-shay assumption in Proposition 1 can be relaxed, provided depreciation is calculated according to the so-called *Replacement Cost Rule*. In particular, there exists a unique depreciation rule, (d_1, \dots, d_T) , such that for any productivity profile (x_1, \dots, x_T) :

$$x_t \cdot c(k) = [d_t + r \cdot (1 - \sum_{i=1}^{t-1} d_i)] \cdot v(k).$$

The label replacement cost accounting indicates that for this depreciation rule the remaining book value at date t ,

$$BV_t = v(k) \cdot [1 - \sum_{i=1}^t d_i]$$

¹³Similarly, in their chapter on life cycle costing Ewert and Wagenhofer (2008) advocate the use of accruals to assign an appropriate share of cash expenditures to the product costs reported at different stages.

would exactly be equal to the fair market value of a used asset if there was a competitive rental market for capacity services (Rogerson 2008).

Given the conditions of Proposition 1, we note that if the depreciation schedule $\{d_t\}$ is accelerated relative to the benchmark of the annuity method, our expected measure of full cost in (7) will still be equal to the LC *on average* in the sense that the present value of the two cost measures will be identical. In particular, the use of straight-line depreciation implies that the capital charges $d_t + r \cdot (1 - \sum_{i=1}^{t-1} d_i)$ will increase over time. Thus, $FC_t > LC$ up to some date $1 \leq \hat{t} \leq T$, but $FC_t < LC$ thereafter. The common reliance on straight-line depreciation also motivates the following observation:

Corollary 1 *Suppose $x_t = 1$, $w_t = w$, and $F_t(k) = f \cdot k$. With full capacity utilization, the levelized product cost exceeds full cost in each period:*

$$LC(k) > \hat{F}C_t(k)$$

for $1 \leq t \leq T$, provided depreciation is calculated according to the straight-line rule.

The traditional accounting measure of full cost based on full straight-line depreciation entails a capacity charge of $\frac{v(k)}{T}$ per unit of capacity. In contrast, the LC requires this charge to be $\frac{v(k)}{\sum \gamma^t}$. The corollary is based on the simple observation that for $\gamma < 1$:

$$T > \sum_{t=1}^T \gamma^t.$$

Most of the economics literature has abstracted from irreversible capacity investments by assuming instead that firms can obtain production capacity in a rental market. Rental capacity then effectively becomes a consumable input, like labor and raw materials. With a rental market for “capital,” there is effectively no distinction between long-run and short-run costs, as the firm can freely adjust all production inputs in any given period. Such a setting is nonetheless of interest in order to compare economists’ conceptualization of marginal cost with the above LC concept. In our notation, Carlton and Perloff (2005, p. 254) conceptualize marginal cost as:

$$MC = w + (r + \delta) \cdot v, \tag{8}$$

where δ denotes “economic depreciation.” Carlton and Perloff (2005) posit that with a competitive market for capacity services, one unit of capacity rented for one period of time

should trade for $(r + \delta) \cdot v$, if v is the constant unit price per unit of capacity (that is, $v(k) = v \cdot k$), r is the required rate of return and δ reflects the physical decay rate of capacity. This specification is indeed compatible with the LC formulation in Proposition 0 under the additional assumptions that assets are infinitely lived ($T = \infty$) and the decline in capacity follows a geometric pattern, that is, $x_t = x^{t-1}$. The denominator in the expression for the unit cost of capacity $c(k)$ in (1) then amounts to:

$$\sum_{t=1}^{\infty} x^{t-1} \cdot \gamma^t = 1 - x + r.$$

Therefore the unit cost of capacity $c(k)$ in (1) coincides with the marginal cost of capital in (8), provided the economic depreciation rate δ is equated with $1 - x$, the capacity 'survival' factor. This equivalence should not come as a surprise to the extent that both the LC in Proposition 0 and the notion of a competitive rental market for capacity are pegged to an economic break-even condition. We note that Carlton and Perloff (2005) do not include the tax factor Δ in their measure of marginal cost. Perhaps more importantly, they also do not include fixed operating costs. We demonstrate below that the long-run marginal product cost must include the fixed operating costs $F_t(k)$, if one seeks a unit cost measure that is to be imputed for capacity investment decisions. Fixed operating costs may be incurred in future periods once the initial capacity is acquired, but nonetheless they are "in play," and therefore incremental, at the initial investment stage.

In concluding this section, we contrast our framework with that in recent studies by Rogerson (2008), Rajan and Reichelstein (2009), Dutta and Reichelstein (2010) and Nezlubin (2012). Our analysis seeks to identify the relevant cost of capacity acquisition in a setting where such acquisitions constitute an irreversible investment in order to produce a stream of future outputs. Once the investment decision has been made, the corresponding expenditure, $v(k)$, and the subsequent operating fixed costs, $F_t(k)$, are sunk. In response to subsequent demand fluctuations for its product, the firm is then left to optimize within its capacity constraint, with the unit variable cost, w_t , left as the only relevant cost. In contrast, the recent studies mentioned above assume that the firm makes a sequence of overlapping capacity investments. With an infinite planning horizon, an expanding product market and the absence of periodic shocks to demand, the firm never finds itself in a position with excess capacity. In effect, the sunk cost nature of past capacity investments never manifests itself

in such a setting.¹⁴

3 Price Taking Firms

This section examines the role of leveled product costs in a market setting with a large number of identical firms. In particular, firms are assumed to be price-takers, they have identical cost structures and there are no barriers to entry. For expositional simplicity, we suppose that the market for the product in question opens at date 0 (there are no market incumbents at date 0) and effectively closes at date T , possibly because the current product or production technology will be replaced by a superior one at that point in time.

The standard textbook description of equilibrium in a competitive industry posits that the market price will be equal to both marginal- and average cost. If a firm is to cover its periodic operating fixed costs so as to obtain zero economic profits, marginal cost must then be below the market price for some range of output levels to the left of the equilibrium output level.¹⁵ In the context of our model, suppliers make irreversible capacity investments at date 0 on the terms described in the previous section. In each subsequent period, firms adjust prices and output to current demand conditions subject to their current capacity constraints. The expected aggregate market demand in period t is given by $Q_t = D_t^o(p_t)$. The functions $D_t^o(\cdot)$ are assumed to be decreasing and we denote by $P_t^o(\cdot)$ the inverse of $D_t^o(\cdot)$. The actual price in period t is a function of the aggregate supply Q_t and the realization of a random shock $\tilde{\epsilon}_t$:

$$P_t(Q_t, \tilde{\epsilon}_t) = \tilde{\epsilon}_t \cdot P_t^o(Q_t). \quad (9)$$

The specification of multiplicatively separable shocks will be convenient in order to quantify a threshold value for the magnitude of the periodic uncertainty.¹⁶ The random variables

¹⁴Dutta and Reichelstein (2010) allow for temporary shocks to demand, but since they also assume zero variable costs, the firm will always produce at capacity.

¹⁵Borenstein (2000) articulates this point as follows: “*It is important to understand that a price-taking firm does not sell its output at a price equal to the marginal cost of each unit of output it produces. It sells all of its output at the market price, which is set by the interaction of demand and all supply in the market. The price-taking firm is willing to sell at the market price any output that it can produce at a marginal cost less than that market price.*”

¹⁶In contrast, the research reviewed by Balakrishnan and Sivaramakrishnan (2002) on capacity choice and full cost pricing exclusively considers an additive error term for the aggregate demand function.

$\tilde{\epsilon}_t$ are assumed to be serially uncorrelated and to have the common density $h(\cdot)$ whose support is contained in the interval $[\underline{\epsilon}, \bar{\epsilon}]$ with $\bar{\epsilon} > 1 > \underline{\epsilon} > 0$. In order for $P_t^o(\cdot)$ to be interpreted as the *expected* inverse demand curve, we also normalize the periodic random fluctuations such that:

$$E[\tilde{\epsilon}_t] \equiv \int_{\underline{\epsilon}}^{\bar{\epsilon}} \epsilon_t \cdot h(\epsilon_t) d\epsilon_t = 1.$$

Firms in the industry are assumed to be risk neutral and to have the same information regarding future demand.¹⁷ In particular, they anticipate that $\tilde{\epsilon}_t$ will be realized at the beginning of period t , prior to each supplier deciding its current level of output, less than or equal to its capacity level. Provided each firm is a price-taker, the supply curve in period t is illustrated by the sequence of vertical and horizontal red lines in Figure 2. In particular, a firm exhausts its full capacity whenever the market price covers at least the short-run marginal cost w_t .

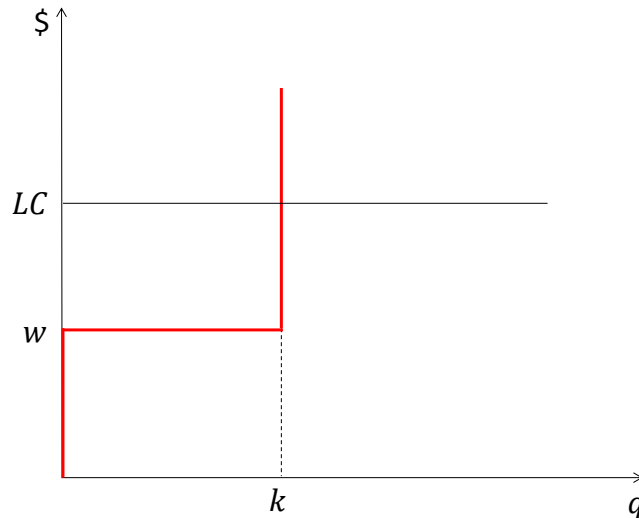


Figure 2: *Competitive Supply Curve with Capacity Constraint*

To develop the results in this section, it will be convenient to begin with a long-run constant returns to scale technology. Thus $v(k) = v \cdot k$ (and therefore $c(k) = c \cdot k$) and

¹⁷We do not consider inventory build-ups in our model. In some industries, holding inventory is not a viable option because of high inventory holding costs (e.g., electricity). The general effect of low inventory holding costs will be to smooth out demand fluctuations. As a consequence, we would then expect the equilibrium prices and capacity levels to approach those obtained under conditions of demand certainty.

$F_t(k) = f_t \cdot k$. As a consequence, the levelized product cost $LC(k)$ is independent of k and will be denoted simply by LC .¹⁸ Initially, we shall also focus on a time-invariant cost- and capacity structure such that $x_t = 1$, $w_t = w$, $h_t(\cdot) = h(\cdot)$, and $F_t(k) = f \cdot k$ for $1 \leq t \leq T$. Furthermore, we consider first a setting where the expected aggregate demand is unchanged over the T -period horizon, that is $P_t^o(\cdot) = P^o(\cdot)$. We refer to the combination of these assumptions as a *stationary environment*.

Since the levelized product cost is the threshold price at which firms break-even on their capacity investments, one would expect that, with frictionless entry into the industry, the competitive equilibrium price will on average be equal to the long-run unit cost. The following result confirms this intuition and characterizes the aggregate level of capacity in equilibrium. We denote by $P_t(w, \epsilon_t, K)$ the equilibrium price in period t , contingent on the unit variable cost w , the aggregate industry capacity level K , and the realization of the periodic shock ϵ_t :

$$P(w, \epsilon_t, K) = \begin{cases} \epsilon_t \cdot P^o(K) & \text{if } \epsilon_t \geq \epsilon(K, w) \\ w & \text{if } \epsilon_t < \epsilon(K, w). \end{cases}$$

Here, $\epsilon(K, w)$ denotes the cut-off level for the periodic shock ϵ_t below which the available capacity will no longer be fully exhausted. Thus, $\epsilon(K, w) \cdot P^o(K) = w$ for values of ϵ in the range $[\underline{\epsilon}, \bar{\epsilon}]$.

It will also be useful to define the *Levelized Fixed Cost (LFC)* as the levelized product cost minus the unit variable cost. Thus, $LFC = LC - w = f + c \cdot \Delta$. Let:

$$LC^- \equiv \max \left\{ \frac{LC}{\bar{\epsilon}}, LFC \right\}.$$

Finally, we identify a condition that relates the volatility in market prices to the variable cost of production as a percentage of the overall levelized cost.

Definition 1 *Market demand is said to exhibit limited price volatility if*

$$\text{Prob} \left[\bar{\epsilon} \geq \frac{w}{LC} \right] = 1. \tag{10}$$

We note that the limited volatility condition is more likely to be satisfied in capital

¹⁸The assumption of fixed operating costs that are unavoidable (even if $q_t = 0$) is of obvious importance here.

intensive industries characterized by high capacity investment costs.¹⁹ If condition (10) is not met, we shall refer to the setting as one of *significant price volatility*.

Proposition 2 *Given a stationary environment, a competitive equilibrium entails a unique aggregate capacity level K^* such that the expected product price satisfies:*

$$E[P(w, \tilde{\epsilon}_t, K^*)] = LC.$$

The equilibrium capacity level K^ is bounded by:*

$$D^o(LC^-) \geq K^* \geq D^o(LC), \quad (11)$$

with $K^ = D^o(LC)$ if and only if price volatility is limited.*

Since in a competitive equilibrium, prices are equal to the long-run marginal cost, Proposition 2 justifies the interpretation of the levelized product cost as the long-run marginal product cost. Suppose first that there are no shocks to price, that is, $P(Q_t, \tilde{\epsilon}_t) = P^o(Q_t)$ for sure. Firms will then produce at full capacity in each period provided $P_t^o(K^*) \geq w$. The capacity constraint prevents the industry from bidding the market price down to w . At the investment stage, the condition of zero economic profits for all participants dictates that the aggregate capacity level must satisfy $P^o(K^*) = LC$. Furthermore, the stationarity of the environment implies that in equilibrium, all capacity investments will be made initially at date 0. In other words, in equilibrium, all firms move in lock-step with their investments at date zero and effectively foreclose the possibility of entry in the remaining $T - 1$ periods.

When market demand is subject to periodic shocks, the equilibrium condition for the optimal aggregate capacity level, K^* becomes:

$$E_\epsilon [\tilde{\epsilon}_t \cdot P^o(Q_t^*(\tilde{\epsilon}_t, K^*))] = LC, \quad (12)$$

where $Q_t^*(\epsilon_t, K^*)$ denotes the optimal aggregate output level, given the initial capacity level and the realization of the current shock ϵ_t . This quantity is equal to K^* if and only if $\epsilon_t \cdot P^o(K^*) \geq w$, that is the industry will produce at capacity given the short-run incremental cost w . An atomistic firm can effectively “commit” to exhausting its capacity since the market price will not drop below w as otherwise some firms would idle their capacity.²⁰

¹⁹Applicable examples include semiconductors, electricity generation and airlines.

²⁰An implicit assumption here is that $\underline{\epsilon} \cdot P^o(0) \geq w$.

Thus, the zero profit condition for an individual firm is given by (12); the expected product price (bounded below by w) must be equal to the long-run marginal cost LC.

If market demand is subject to limited price volatility in the sense of condition (10), the zero-profit condition in (12) is met at $K = K^o$ because:

$$\underline{\epsilon} \cdot P^o(K^o) \equiv \underline{\epsilon} \cdot LC \geq w,$$

and therefore $Q_t^*(\epsilon_t, K^o) = K^o$. Intuitively, one might expect that higher price volatility results in a lower aggregate capacity level than $K^o \equiv D^o(LC)$, because if condition (10) is not met, some of the industry's aggregate capacity will be idle with positive probability. However, as $P^o(\cdot)$ is decreasing, the zero economic profit condition in (12) can only be met for some $K^* > K^o$. In effect, the aggregate investment in capacity will be larger than under conditions of limited volatility because there is a need to commit the entire LC at the investment stage. Firms have a *call option* to idle parts of their capacity in subsequent periods which allows them to avoid the variable cost in case of unfavorable demand realizations. Figure 3 illustrates the equilibrium capacity levels identified in Proposition 2.

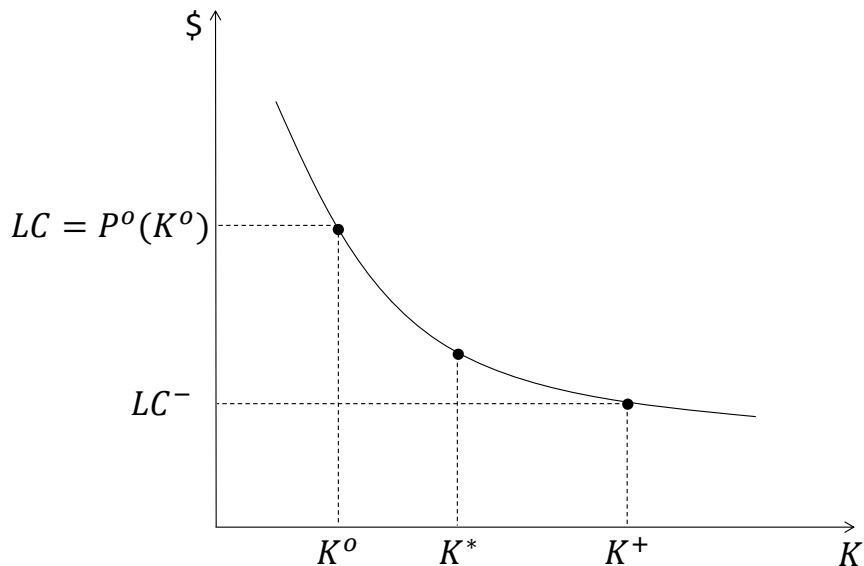


Figure 3: *Equilibrium Capacity Levels with Significant Price Volatility*

The upper bound on capacity, as given by the relation $D^o(LFC) \geq K^*$, follows from the observation that if the variable production costs were to be zero, that is, $w = 0$, firms would always fully use the available capacity. As a consequence, the expected market price

would then be equal to $LFC = LC - w$ and $D^o(LFC) = K^+$. With positive variable production costs, the aggregate capacity level must in equilibrium be correspondingly in order to maintain zero economic profits for all firms.

The distribution of equilibrium prices for two volatility scenarios is illustrated in Figure 4. With limited volatility, the support of $\tilde{\epsilon}$ is given by $[\underline{\epsilon}_1, \bar{\epsilon}_1]$ and the market price will vary proportionally with $\tilde{\epsilon}$ at the rate $P^o(K_1^*)$. With significant volatility, the support of $\tilde{\epsilon}$ broadens to $[\underline{\epsilon}_2, \bar{\epsilon}_2]$ and equilibrium prices will be a piecewise linear function of $\tilde{\epsilon}$. We note that in the low volatility scenario, prices are *more responsive* to a given magnitude of the shock ϵ , since $P^o(K_1^*) > P^o(K_2^*)$.

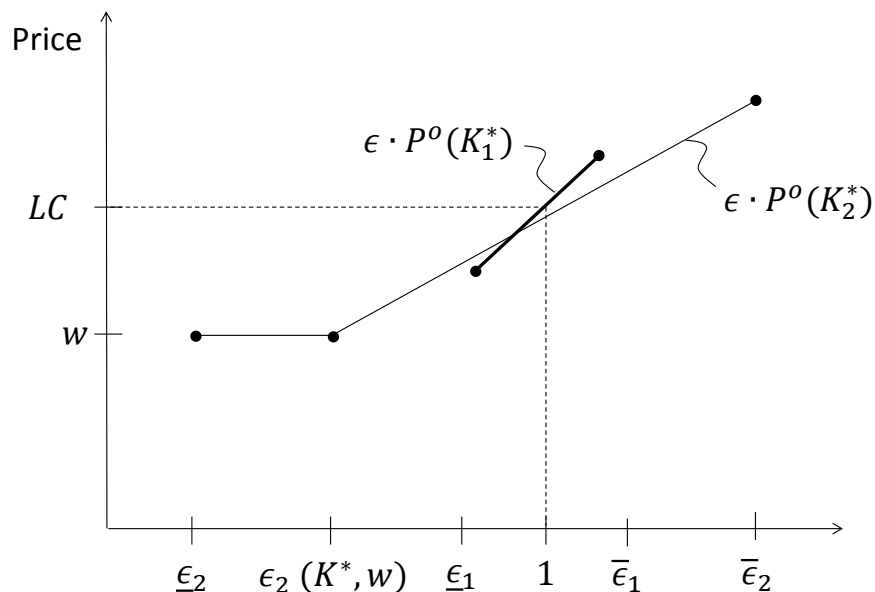


Figure 4: *Distribution of Equilibrium Prices*

The baseline result in Proposition 2 has assumed that the fixed operating costs, f , are *unavoidable*. Suppose now that the firm will not incur the unit cost f if it were to idle its entire capacity in period t . Proposition 2 then remains valid as stated, except that the aggregate level of capacity may increase. Once the fixed cost f are “in play”, the relevant price floor for capacity utilization in each period increases to $w + f$ because suppliers would be better off withholding their capacity if the price were to drop below $w + f$.

Corollary 2 *If the periodic operating fixed costs per unit of capacity, f , are avoidable, the expected equilibrium price remains equal to the levelized product cost. The equilibrium level of aggregate capacity increases relative to the scenario where f is unavoidable if and only if*

$$\text{Prob} \left[\bar{\epsilon} \geq \frac{w + f}{LC} \right] < 1.$$

The identification of the competitive market price in Proposition 2 yields several predictions for the expected profitability that firms report in equilibrium. In the context of our model, net income becomes operating income less taxes paid. If economic profitability is assessed in terms of residual income, that is, net income minus the imputed interest charge on book value, the Conservation Property of residual income (Preinreich 1938) implies that the present value of future residual incomes will be zero if the expected revenue per unit of output is indeed equal to the levelized cost. The inter-temporal distribution of the net income numbers depends on both the financial- and tax reporting rules for operating assets. Given the prevalence of straight-line depreciation in practice, the following result assumes that firms rely on this rule to calculate net income.

Corollary 3 *Suppose firms use straight-line depreciation for financial reporting purposes. In equilibrium, expected net income will be positive in each period provided*

$$\alpha \leq 1 - \frac{\sum_{i=1}^T \gamma^i}{T}. \quad (13)$$

Equation (13) is satisfied for many reasonable parameter configurations. *Ceteris paribus*, it will be met for a higher discount rate and investment projects with a longer useful life. The result in Corollary 2 holds regardless of the applicable tax depreciation schedule. In fact, the condition in (13) considers a lower bound on first-period profits that would emerge if the initial investment were directly expensed for tax purposes. A sharper prediction regarding net income can be obtained if one assumes that straight-line depreciation is applied for both financial reporting and tax purposes. We then obtain the following prediction for expected net income per unit of capacity acquired:

$$\pi = (1 - \alpha) \left[LC - w - f - \frac{v}{T} \right] = (1 - \alpha) \left[v \left(\frac{\Delta}{\sum_{i=1}^T \gamma^i} - \frac{1}{T} \right) \right] > 0. \quad (14)$$

Equation (14) immediately yields an expression for what is frequently termed a “fair profit” in government contracts. For sole-source contracting, governments frequently rely on a cost-plus formula in which the “plus” is calibrated as a profit allowance that gives the firm an adequate return on its invested capital. As argued in Section 2, the overall cost-plus price must then be the levelized cost. For simplicity, suppose the contractor is obligated to deliver one widget in each of the next T , following an initial investment of v . If cost is calculated as full cost based on straight-line depreciation, the cost-plus price per widget becomes:

$$LC = (w + f + \frac{v}{T}) + \eta \cdot v,$$

where, consistent with (14), the profit allowance, η , is calculated as a percentage of the initial investment with

$$\eta = \frac{\Delta}{\sum_{i=1}^T \gamma^i} - \frac{1}{T}.$$

To illustrate, if $r=10\%$ and $\alpha = .35$, the mark-up factor η stays in a relatively narrow range of 8-10% as the duration of the contract T varies between 15 and 30 years.

The remainder of this section extends the benchmark result in Proposition 2 by relaxing several of the assumptions invoked thus far. One consequence of the assumed constant returns to scale technology in our model has been that the efficient scale of operation for individual firms, that is the individual k , remains indeterminate. If the cost of acquiring capacity $v(k)$ and the periodic fixed operating costs $F_t(k)$ are non-linear, the condition of zero expected economic profits dictates that the efficient scale of operation must be chosen such that the LC per unit of output is minimized. Referring back to the generalized definition of $LC(k)$ in Section 2, we define the efficient scale of operation by:

$$k^* \in \operatorname{argmin}\{LC(k)\}.$$

The capacity level k^* effectively determines the efficient size of firms in the industry. The expected equilibrium price in Proposition 2 becomes $LC(k^*)$, while the aggregate capacity level in equilibrium, K^* , is determined by $P^o(K^*) = LC(k^*)$.²¹

²¹Non-linearities in $LC(k^*)$ may give rise to a traditional U-shaped average cost curve. We note, however, that unlike the usual depiction in microeconomic textbooks the short-run marginal cost (w) is always below the long run marginal cost ($LC(K^*)$) for any output level $Q < K^*$.

Suppose next that the aggregate demand is no longer stationary but instead contracts over time.²² In particular, suppose that $P_t^o(\cdot) = \lambda_t \cdot P^o(\cdot)$ for all $1 \leq t \leq T$ such that $\lambda_{t+1} < \lambda_t < 1$ for all $1 \leq t \leq T$. We refer to such a scenario as a *declining product market*. We introduce the notation $\boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_T)$ and define:

$$m(\boldsymbol{\lambda}) = \frac{\sum_{t=1}^T \gamma^t}{\sum_{t=1}^T \lambda_t \cdot \gamma^t}.$$

Clearly, $m(\boldsymbol{\lambda}) > 1$ in a declining product market. Intuitively, one would expect the equilibrium capacity in a declining product to be lower compared to a stationary market. The following result confirms this intuition and shows that on average, the expected equilibrium market price will also be higher as a consequence of market demand that diminishes over time. For the following result, we also allow the unit variable cost, w_t , to change over time. Consistent with the notion of a declining product market, the restriction imposed is that $w_{t+1} \geq w_t$. The importance of that restriction is that in a competitive equilibrium, all capacity investments are made at date 0. We recall from Section 2 that in the one-hoss shay scenario where all $x_t = 1$,

$$LC \equiv LFC + \frac{\sum_{t=1}^T w_t \cdot \gamma^t}{\sum_{t=1}^T \gamma^t}.$$

Let $P_t(w_t, \epsilon_t, \lambda_t, K)$ denote the equilibrium market price in period t when aggregate market demand is given by $\lambda_t \cdot P^o(\cdot)$ and the current demand shock is realized. By construction,

$$P_t(w_t, \epsilon_t, \lambda_t, K) = \begin{cases} \epsilon_t \cdot \lambda_t \cdot P^o(K) & \text{if } \epsilon_t \geq \epsilon(w_t, K, \lambda_t) \\ w_t & \text{if } \epsilon_t < \epsilon(w_t, K, \lambda_t), \end{cases}$$

with

$$\epsilon(w_t, K, \lambda_t) \cdot \lambda_t \cdot P^o(K) \equiv w_t.²³$$

²²In contrast, the recent work on capacity investments, such as Rogerson (2008, 2011), Rajan and Reichelstein (2009), Dutta and Reichelstein (2010), Nezlobin (2012) and Nezlobin, Rajan, and Reichelstein (2012), has consistently focused on expanding markets, with the consequence that firms are never left with excess capacity.

²³As before, $\epsilon(w, K, \lambda_t)$ is bounded by $\underline{\epsilon}$ and $\bar{\epsilon}$.

Proposition 3 *With a declining product market, a competitive equilibrium entails a unique aggregate capacity level K^* such that the expected equilibrium prices satisfy:*

$$\sum_{t=1}^T E[\lambda_t \cdot P_t(w_t, \tilde{\epsilon}_t, \lambda_t, K)] \cdot \gamma^t = LC \cdot \sum_{t=1}^T \gamma^t.$$

The capacity level K^ is bounded by:*

$$D^o(LC^- \cdot m(\boldsymbol{\lambda})) \geq K^* \geq D^o(LC \cdot m(\boldsymbol{\lambda})) \quad (15)$$

with $K^ = D^o(LC \cdot m(\boldsymbol{\lambda}))$ if and only if:*

$$\text{Prob} \left[\tilde{\epsilon} \cdot \lambda_T \geq \frac{w_t}{LC \cdot m(\boldsymbol{\lambda})} \right] = 1. \quad (16)$$

For a declining product market, the zero-profit condition applies over the entire planning horizon rather than on a period-by-period basis. In particular, the annuity value of the levelized cost must then be equal to the present value of future expected market prices. With a declining product market and weakly increasing variable costs, it becomes *ceteris paribus* more likely that market demand exhibits significant volatility, at least in later time periods. Formally, this can be seen from the fact that the inequality in (16) is more stringent than (10) because $m(\boldsymbol{\lambda}) \cdot \lambda_T < 1$.

Proposition 3 can be extended to a growing product market, where $\lambda_{t+1} > \lambda_t > 1$, provided the growth rates are “sufficiently small.” Once demand grows too quickly, it will no longer be an equilibrium to have all capacity investments undertaken at date 0. The resulting scenario of sequential and overlapping investments will be similar in nature to the recent models of Rogerson (2008, 2011), Rajan and Reichelstein (2009), Dutta and Reichelstein (2010), and Nezlobin (2012). In particular, the main finding of Proposition 2, namely that the expected market price in each period is equal to the levelized product cost, will continue to hold provided (i) firms can add capacity continuously, (ii) the product market expands monotonically over time, and (iii) there is an infinite planning horizon.

4 Price Setting Firms

4.1 Monopoly

We now turn to a monopolistic firm that faces the same stationary environment posited in connection with Proposition 2. In particular, suppose that expected aggregate market demand in period t is given by $q_t = D^o(p_t)$. The function $D^o(\cdot)$ is assumed to be decreasing and we denote by $P^o(\cdot)$ again the inverse of $D^o(\cdot)$. The actual price in period t is a function of the quantity supplied q_t and the realization of the random shock $\tilde{\epsilon}_t$:

$$P(q_t, \tilde{\epsilon}_t) = \tilde{\epsilon}_t \cdot P^o(q_t). \quad (17)$$

Similar to our assumptions in Section 3, the monopolist observes the realization of $\tilde{\epsilon}_t$ before deciding on the quantity supplied in that period. Let $MR^o(q)$ denote the marginal revenue, with

$$MR^o(q) \equiv \frac{d}{dq}[P^o(q) \cdot q].$$

Throughout this section, we assume that marginal revenue is decreasing in q . Assuming the short-run variable costs are again constant over time and equal to w and fixed operating costs in each period are unavoidable and equal to f per unit of capacity, we obtain the following result.

Proposition 4 *Given a stationary environment, the optimal capacity investment, k^* , for a monopolist satisfies $k^* \in [k^o, k^+]$, where k^o and k^+ are given by:*

$$MR^o(k^o) \equiv LC \quad \text{and} \quad MR^o(k^+) \equiv LC^-. \quad (18)$$

Furthermore, $k^ = k^o$ if and only if the condition of limited price volatility in (10) is satisfied.*

Proposition 4 reinforces our general claim that the LC is the relevant cost for capacity investment decisions. Specifically, the first-order condition for the optimal capacity level k^* is:

$$E_\epsilon [\tilde{\epsilon}_t \cdot MR^o(q_t^*(\tilde{\epsilon}_t, w, k^*))] = LC, \quad (19)$$

where $q_t^*(\epsilon_t, w, K^*)$ denotes the optimal monopoly output level in period t , given the initial capacity choice and the realization of the current shock ϵ_t . We note that $MR^o(q_t^*(\epsilon_t, w, k^*)) =$

w if the capacity constraint does not bind for a particular ϵ_t , while $MR^o(q_t^*(\epsilon_t, w, k^*)) = MR^o(k^*)$ in case $\epsilon_t \cdot MR^o(k^*) \geq w$. If the limited price volatility condition in (10) holds and therefore

$$\epsilon_t \cdot MR^o(k^o) = \epsilon_t \cdot LC \geq w$$

for all realizations of ϵ_t , the first-order condition in (19) implies $k^* = k^o$.

Once price volatility becomes *significant*, in the sense that condition (10) does not hold and the deviation from market price exceeds the ratio of variable cost to levelized costs, the monopolist will withhold capacity for unfavorable realizations of market demand. Like in the competitive setting, this has the effect of driving up the amount of initial capacity investment, because the variable cost is not yet sunk. An upper bound on the optimal capacity level is obtained by equating marginal revenue with the levelized fixed cost. *LFC* would be the relevant cost of capacity for a hypothetical firm that has no variable costs and therefore always produces at capacity. The claim in Proposition 4 then follows, because the marginal return on capacity investments for the hypothetical monopolist (that is, $w = 0$) is always at least as large as the marginal return on investment in the actual problem.

Given our finding in Proposition 2, we conclude that with limited price volatility, the monopolist's capacity investment is lower than the aggregate capacity level that obtains under competition. With limited price volatility, we also confirm the usual prediction that the monopoly price entails a mark-up over the marginal cost, where the mark-up is given by:

$$p^*(k^o, \epsilon_t) = \epsilon_t \cdot LC \cdot \frac{E(k^o)}{E(k^o) - 1},$$

and $E(k^o) > 1$ denotes the price elasticity of demand. In contrast to the customary textbook solution, though, the basis for the mark-up is now the levelized product cost rather than the short-run marginal cost.

Holding other parameters fixed, Proposition 4 suggests that the monopoly capacity investment will be higher for firms operating in environments with high price volatility. To formalize this comparative static result, we rely on the usual metric of calling one probability distribution as "riskier" than another distribution if the first one can be expressed as a compound lottery of the second. Formally, the distribution $h_2(\cdot)$ is obtained by a mean-preserving spread of $h_1(\cdot)$, if:

$$\tilde{\epsilon}_2 = \mu + \tilde{\epsilon}_1,$$

where μ is a random variable distributed on some interval $[\underline{\mu}, \bar{\mu}]$, such that the conditional probability distributions $g(\mu|\epsilon_1)$ satisfy:

$$\int_{\underline{\mu}}^{\bar{\mu}} \mu g(\mu|\epsilon_1) d\mu = 0,$$

for all ϵ_1 .²⁴ Thus, $\tilde{\epsilon}_2$ is obtained from $\tilde{\epsilon}_1$ by adding a second lottery which preserves the original mean of $\tilde{\epsilon}_1$.

Corollary 4 *Higher price volatility results ceteris paribus in a larger capacity investment by the monopolist.*

One caveat of the preceding comparative statics result is that the cost of capital is held constant. In other words, higher price volatility in the product market for the investment under consideration is not assumed to alter any risk premium embedded in the cost of capital. This specification is plausible if the investment decision under consideration is small relative to the firm's overall portfolio. It should also be noted that the prediction of a larger capacity investment due to more volatility does not necessarily translate into a prediction of lower monopoly prices on average. For a given distribution $h(\cdot)$, more capacity will, of course, ensure that the quantity delivered to the market in each state will be at least as large as if capacity were more constrained. Yet, a higher variance distribution $h(\cdot)$ not only results in more capacity, it also shifts the probability mass such that the impact on expected monopoly prices appears difficult to predict in general.

For the special case of a uniform distribution, it turns out that the expected monopoly price is indeed lower as volatility increases. Suppose $h(\epsilon) = \frac{1}{(2\cdot\theta)}$. Higher values of θ then correspond to higher volatility in the sense of a mean-preserving spread. Assuming in addition a constant elasticity of demand, we established through numerical simulation that the expected monopoly price is monotonically decreasing in θ . Thus, for this particular setting, we conclude that higher price volatility leads to a higher capacity investment and to lower average monopoly prices.

²⁴Our definition of mean-preserving spreads here follows the standard approach as in Mas-Colell, Whinston, and Green (1995), except that we posit the existence of probability densities. The only restriction imposed by this specification is that the distributions cannot have probability point-mass.

4.2 Duopoly

Our findings in the previous two sections have demonstrated that the levelized product cost is the relevant cost for capacity investments both in a monopoly and a competitive setting. These findings strongly suggest that this attribute of the LC also carries over to oligopolistic settings in which the incumbent firms make initial capacity investments and subsequently compete subject to their capacity limitations. To model the interactions of the oligopolists, we focus on quantity games in the sense of Cournot competition.²⁵

Suppose two firms, identical in their products and cost structure, first make simultaneous capacity decisions and then in each subsequent period choose production quantities simultaneously. These choices then determine the market price in that period. In the simplest setting, the two identical firms i and j face a *stationary environment* as described in Section 3. If the firms choose the production quantities $\mathbf{q}_t = (q_t^1, q_t^2)$ the product price in period t is given by:

$$P(q_t^1 + q_t^2, \tilde{\epsilon}_t) = \tilde{\epsilon}_t \cdot P^o(q_t^1 + q_t^2). \quad (20)$$

The initial capacity investments are denoted by $\mathbf{k} = (k^1, k^2)$ and thus $K(\mathbf{k}) = k^1 + k^2$ becomes the aggregate capacity level. Given the quantity and capacity choice of firm j , the marginal revenue for firm i is given by:

$$MR^o(q^i | q^j) \equiv \frac{d}{dq^i} [P^o(q^1 + q^2) \cdot q^i].$$

We assume that the marginal revenue of firm i is decreasing in both q^i and q^j . Clearly, that condition will be met for "standard" willingness to pay curves, including the special case of a linear function.

Proposition 5 *Given a stationary environment, suppose the limited price volatility condition in (10) holds. The following then constitutes a subgame perfect equilibrium outcome:*

i) At date 0, both firms choose identical capacity investments, $k^{1} = k^{2*} = k^*$ which satisfy the equation:*

$$MR^o(k^* | k^*) = LC. \quad (21)$$

²⁵For ease of notation, we rely on the duopoly setting for the derivation of our results. It appears that the results reported here carry over to settings with n competing firms without major modifications.

ii) In subsequent periods $1 \leq t \leq T$, each firm supplies the maximal production quantity $q_t^1 = q_t^2 = k^*$.

The proof of the Proposition exploits that, regardless of the initial capacity choices and regardless of the realization of the noise term $\tilde{\epsilon}_t$, there is a unique Nash-equilibrium in the output quantities chosen at stage t . The restriction imposed by subgame perfection therefore requires that the initial capacity choices constitute an equilibrium if followed by the subsequent output levels corresponding to the unique equilibrium in each stage game. In particular, these equilibrium output levels call for the firm with the lower capacity level to exhaust its capacity provided the marginal revenue at the capacity limit still covers the unit variable cost w . With limited price volatility, firm j will therefore anticipate full capacity utilization whenever $k^j \leq k^*$. Finally, it is shown in the proof that k^* is indeed a best response to the same capacity level chosen by the other firm.

There are several promising directions for extending the baseline result in Proposition 5. First, it would be desirable to characterize the entire set of equilibrium outcomes. Absent any price volatility, the first-order conditions for a Nash-equilibrium immediately show that the capacity levels identified in Proposition 5 constitute the unique equilibrium. Furthermore, our finding for the monopoly scenario (Proposition 4) suggests that regardless of the degree of price volatility, the equilibrium capacity levels will be bounded above by levels that correspond to a relevant cost equal to the levelized fixed cost, LFC.

Second, it would be natural to consider a scenario where one of the firms has a first-mover advantage with its capacity choice. Once both firms have entered the market, albeit sequentially, they decide their output levels simultaneously in each subsequent period. At the initial investment stages, the Stackelberg leader can then credibly preempt capacity investments by the follower to the extent that the price of the output to be produced still exceeds the unit variable cost. Anticipating the follower's reaction, the first mover then has an incentive to choose a capacity level larger than the equilibrium level, k^* , identified in Proposition 5.

Third, beginning with the work of Kreps and Scheinkman (1983) earlier industrial organization literature has established an equivalence between Cournot quantity competition and Bertrand price competition subject to capacity constraints. In future work, it would be useful to establish this equivalence for the basic set-up examined in this paper. For particular settings, a number of papers have highlighted that this equivalence depends on additional

specifications, including the specific “rationing rule” employed by Kreps and Scheinkman (1983). Furthermore, the degree of price volatility in the product market demand is likely to have an impact on any equivalence result between Cournot and Bertrand competition.²⁶

5 Concluding Remarks

A central feature of many manufacturing and service industries is that firms have to make irreversible upfront capacity investments in order to deliver products and services to customers. This paper has examined the economic relevance of the so-called levelized product cost (LC), a concept widely used in the electricity literature under the label Levelized Cost of Electricity. This life-cycle cost is based entirely on discounted cash flows and is calibrated as the minimum average price the firm would have to receive for the product in question in order to justify investment in a particular production facility. We demonstrate that the LC can be represented as a fully loaded cost, provided depreciation is calculated in a manner that reflects the decline in productive capacity over time. In addition, this fully loaded cost must include capital charges on the outstanding book value of the productive asset and the pre-tax charge for capacity must be marked up by a tax factor which reflects the depreciation schedule allowable for tax purposes. As a consequence, our notion of fully loaded cost exceeds the measure of full product cost as usually portrayed in accounting textbooks.

We demonstrate that the levelized cost can be interpreted as the long-run marginal product cost. In particular, this cost becomes the prediction for the expected equilibrium price in a competitive setting. For a range of alternative market settings, we show that the LC is the relevant unit cost in the sense that firms equate the levelized product cost with the expected marginal revenue that is obtained through the sequentially optimal output levels in subsequent periods, given the initial capacity choice. Our results thus confirm that there is no inherent tension between the cost accountant’s notion of fully loaded product cost of a product and the long-run marginal cost, as commonly conceptualized in the industrial organization literature.

The equilibrium level of capacity will generally exceed that corresponding to the expected demand (marginal revenue) at the levelized cost if market demand exhibits significant price volatility. Our classification of limited versus significant volatility hinges on the ratio of

²⁶See, in particular, Davidson and Deneckere (1986), Osborne and Pitchik (1986), Hviid (1991) and Grant and Quiggin (1996).

short-run variable costs to levelized product costs. This ratio tends to be small in capital intensive industries. In general, the optimal capacity level is such that the corresponding marginal revenue is above the LFC, because firms do not need to commit to use their entire capacity at the investment stage. Instead they retain a call option to idle parts of their installed capacity in future periods as a response to unfavorable market conditions. This call option becomes more valuable with higher volatility, resulting *ceteris paribus* in higher capacity investments.

Moving further afield, we note that the setting in this paper has confined attention to single product environments. In contrast, the existing work on product costing has focused on settings where capacity installations and their attendant fixed costs are shared among multiple products (Cooper and Kaplan 1988). In those settings, particular cost allocation rules, like activity-based costing systems, are usually justified by the need to identify the long-run cost of individual products, even though capacity investments are not treated as endogenous. Our analysis lends itself to extending the existing work on alternative product costing systems to environments in which multiple products share the same capacity installations and relevant costs are determined by means of both intertemporal and cross-sectional cost allocations.²⁷

The literature on intra-company pricing has struggled to find support for the common practice of valuing transfers across segments of a firm at full cost. This difficulty may have been in part a consequence of the assumption in those studies that upfront investments cannot be verified by the firm's accounting system; see, for instance, Baldenius, Reichelstein, and Sahay (1999), Johnson (2006) and Pfeiffer, Schiller, and Wagner (2011).²⁸ One promising avenue for future research is to revisit the desirability of full-cost transfer prices in the context of capacity investments. The work of Dutta and Reichelstein (2010) takes a step in that direction, though, as mentioned above, their model confines attention to settings in which the divisions never find themselves with excess capacity.

²⁷With uncertain market demand, there will be a natural diversification effect arising from multiple products using the same scarce capacity resources.

²⁸This literature takes the perspective that divisional investments are "soft" and relationship specific.

6 Appendix

Proof of Proposition 2.

Denote the aggregate capacity investment at date 0 by K . We first characterize the equilibrium level of K^* and the expected equilibrium prices, $E[P(w, \tilde{\epsilon}_t, K^*)]$, under the assumption that no firm makes additional capacity investments at any date t , $1 \leq t < T$. We subsequently confirm that in equilibrium there are indeed no capacity investments after date 0.

If the aggregate capacity level is K at date 0, the equilibrium price in period t is given by:

$$P(w, \epsilon_t, K) = \begin{cases} \epsilon_t \cdot P^o(K) & \text{if } \epsilon_t \geq \epsilon(K, w) \\ w & \text{if } \epsilon_t < \epsilon(K, w). \end{cases}$$

Here $\epsilon(K, w)$ denotes the unique cut-off level defined as follows:

$$\epsilon(K, w) = \begin{cases} \bar{\epsilon} & \text{if } \bar{\epsilon} \cdot P^o(K) \leq w \\ \frac{w}{P^o(K)} & \text{if } \bar{\epsilon} \cdot P^o(K) > w > \underline{\epsilon} \cdot P^o(K) \\ \underline{\epsilon} & \text{if } \underline{\epsilon} \cdot P^o(K) \geq w \end{cases}$$

Clearly, $\epsilon(K, w)$ is weakly increasing in K .

Consider now an individual firm that has invested a capacity level k . There is no loss of generality in assuming that this firm will indeed produce k units of output since the product market price never falls below the short-run marginal cost w . This follows from our assumption that

$$\underline{\epsilon} \cdot P^o(0) \geq w.$$

Obviously, individual firms are indifferent about idling part as all of their capacity in case $P(w, \epsilon_t, K) = w$.²⁹

The present value of expected future cash flows for the representative firm then becomes:

$$\Gamma(k|K) = \sum_{t=1}^T \left\{ E \left[[P(w, \tilde{\epsilon}_t, K) - w] \cdot k - f \cdot k - \alpha \cdot \tilde{I}_t \right] \right\} \gamma^t - v \cdot k.$$

Here \tilde{I}_t denotes taxable income given by:

$$\tilde{I}_t = [P(w, \tilde{\epsilon}_t, K) - w] \cdot k - f \cdot k - \hat{d}_t \cdot v \cdot k.$$

²⁹The following arguments would lead to the same conclusion if one were to assume that the firm chooses some $0 \leq q \leq k$ in case the market price equals w in period t .

In a competitive equilibrium, the aggregate capacity level must be chosen so that for each atomistic firm: $\Gamma(k|K^*) = 0$. Solving this equation for the expected market price under the assumption that $q = k$, the expected profit per unit of capacity acquired must be zero. Thus:

$$(1 - \alpha) \sum_{t=1}^T E [P(w, \tilde{\epsilon}_t, K^*)] \cdot \gamma^t = \quad (22)$$

$$v \cdot \left[1 - \alpha \cdot \sum_{t=1}^T \hat{d}_t \cdot \gamma^t \right] + (1 - \alpha) \cdot \sum_{t=1}^T \{f + w\} \cdot \gamma^t.$$

Dividing by $(1 - \alpha)$ in (22) and recalling the definition of the tax factor, Δ :

$$\Delta = \frac{1 - \alpha \sum_{t=1}^T \hat{d}_t \cdot \gamma^t}{1 - \alpha}$$

as well as the definitions of the time-averaged unit fixed cost, f , and the time-averaged unit variable cost, w , respectively, we conclude that the right-hand side of (22) is exactly equal to $LC \cdot \sum_{t=1}^T \gamma^t$ and therefore:

$$E [P(w, \tilde{\epsilon}_t, K^*)] = LC. \quad (23)$$

The expected equilibrium price in period t is equal to:

$$E [P(w, \tilde{\epsilon}, K^*)] = \int_{\underline{\epsilon}}^{\epsilon(K^*, w)} w \cdot h(\epsilon) d\epsilon + \int_{\epsilon(K^*, w)}^{\bar{\epsilon}} \epsilon \cdot P^o(K^*) \cdot h(\epsilon) d\epsilon. \quad (24)$$

We next demonstrate that $E [P(w, \tilde{\epsilon}, \cdot)]$ is decreasing in K . For any $K_1 < K_2$, we can write $E [P(w, \tilde{\epsilon}, K_1)]$ as:

$$\int_{\underline{\epsilon}}^{\epsilon(K_1, w)} w \cdot h(\epsilon) d\epsilon + \int_{\epsilon(K_1, w)}^{\epsilon(K_2, w)} \epsilon \cdot P^o(K_1) h(\epsilon) d\epsilon + \int_{\epsilon(K_2, w)}^{\bar{\epsilon}} \epsilon \cdot P^o(K_1) h(\epsilon) d\epsilon \quad (25)$$

and similarly $E [P(w, \tilde{\epsilon}_t, K_2)]$ can be expressed as:

$$\int_{\underline{\epsilon}}^{\epsilon(K_1, w)} w h(\epsilon) d\epsilon + \int_{\epsilon(K_1, w)}^{\epsilon(K_2, w)} w h(\epsilon) d\epsilon + \int_{\epsilon(K_2, w)}^{\bar{\epsilon}} \epsilon \cdot P^o(K_2) h(\epsilon) d\epsilon. \quad (26)$$

Since $\epsilon \cdot P^o(K_1) \geq w$ for $\epsilon(K_1, w)$ and $P^o(K_1) > P^o(K_2)$, we conclude that each of the three integrals in (25) is respectively equal to or larger than its counterpart in (26).

Suppose next that the condition

$$\text{Prob} \left[\tilde{\epsilon} \geq \frac{w}{LC} \right] = 1 \quad (27)$$

is met. By construction, we then have $\epsilon(K^*, w) = \underline{\epsilon}$ and therefore

$$E [P(w, \tilde{\epsilon}, K^*)] = P^o(K^*).$$

Equation (23) yields that $K^* = K^o$, with K^o defined as $P^o(K^o) \equiv LC$. Conversely, if (27) is not met, then the equilibrium capacity level must satisfy $K^* > K^o$. To see this, assume to the contrary that $K^o \geq K^*$. That would lead to a contradiction since by (24) :

$$\begin{aligned} E[P(w, \tilde{\epsilon}_t, K^*)] &\geq E[P(w, \epsilon_t, K^o)] \\ &> P^o(K^o) \\ &\equiv LC. \end{aligned} \quad (28)$$

The strict inequality in (28) follows from the fact that if the condition in (27) is not met, then $\underline{\epsilon} < \epsilon(K^o, w) \leq \bar{\epsilon}$. We thus conclude that $K^* > K^o$.

To show the upper bound on K^* claimed in Proposition 2, we note that the zero-profit condition again stipulates that:

$$E[P(w, \tilde{\epsilon}_t, K^*)] = w + f + c \cdot \Delta.$$

It follows from (24) that

$$E[P(w, \tilde{\epsilon}_t, K^*)] \leq \bar{\epsilon} (1 - H(\epsilon(K^*, w))) \cdot P^o(K^*) + w \cdot H(\epsilon(K^*, w))$$

where $H(\cdot)$ is the cumulative distribution function of the density $h(\cdot)$. Thus

$$[\bar{\epsilon} \cdot P^o(K^*) - w] \cdot [1 - H(\epsilon(K^*, w))] \geq f + c \cdot \Delta$$

and therefore

$$P^o(K^*) \geq \frac{LC}{\bar{\epsilon}},$$

which is equivalent to $K^* \leq D^o\left(\frac{LC}{\bar{\epsilon}}\right)$.

To demonstrate that $K^* \leq D^o(LFC)$ is another bound on the aggregate capacity in equilibrium, let K^{**} denote the capacity level that would emerge if $w = 0$ and thus $LC = LFC$. Since firms would then always produce at capacity, we have:

$$E [P(w = 0, \tilde{\epsilon}_t, K^{**})] = P^o(K^{**}) = LFC.$$

At the same time,

$$E [P(w, \tilde{\epsilon}_t, K^*)] = LC,$$

or equivalently

$$E [P(w, \tilde{\epsilon}_t, K^*)] = \int_{\underline{\epsilon}}^{\epsilon(K^*, w)} w \cdot h(\epsilon) d\epsilon + \int_{\epsilon(K^*, w)}^{\bar{\epsilon}} \epsilon \cdot P^o(K^*) \cdot h(\epsilon) d\epsilon = w + LFC.$$

It follows that

$$P^o(K^{**}) = \int_{\epsilon(K^*, w)}^{\bar{\epsilon}} [\epsilon \cdot P^o(K^*) - w] h(\epsilon) d\epsilon. \quad (29)$$

If $w = 0$, (29) would require $K^* = K^{**}$. The right-hand side of (29) is decreasing in w since its derivative—by Leibniz' rule—is given by: $-[1 - H(\epsilon(K^*, w))]$. At the same time, the right-hand side of (29) is decreasing in K^* . Thus for any $w > 0$, (29) can only hold if $K^* < K^{**}$.

It remains to show that in equilibrium all investments will occur initially, that is, no firm will make capacity investments at a date $t \geq 1$. Assume to the contrary, that some firms will invest at date $t = 0$ and others (or the same firm) will possibly add additional capacity at subsequent dates. The resulting sequence of aggregate capacity levels $\mathbf{K} = (K_1, \dots, K_T)$ will therefore be weakly increasing. The condition of zero economic profits dictates that in equilibrium both an investment at date 0 and an investment at date 1 have expected future cash flows with zero present values. Without loss of generality, we can normalize each one of these investments to one unit of capacity. Thus:

$$\Gamma_0(1|\mathbf{K}) = \sum_{t=1}^T \left\{ E \left[P(w, \tilde{\epsilon}_t, K_t) - w - f - \alpha \cdot \tilde{I}_t \right] \right\} \gamma^t - v = 0, \quad (30)$$

with $\tilde{I}_t = P(w, \tilde{\epsilon}_t, K_t) - w - f - \hat{d}_t \cdot v$ and

$$\Gamma_1(1|\mathbf{K}) = \sum_{t=2}^T \left\{ E \left[P(w, \tilde{\epsilon}_t, K_t) - w - f - \alpha \cdot \tilde{I}_t \right] \right\} \gamma^{t-1} - v = 0, \quad (31)$$

with $\tilde{I}_t = P(w, \tilde{\epsilon}_t, K_t) - w - f - \hat{d}_{t-1} \cdot v$ for $2 \leq t \leq T - 1$ and

$$\tilde{I}_T = P(w, \tilde{\epsilon}_T, K_T) - w - f - (\hat{d}_{T-1} + \hat{d}_T) \cdot v$$

The investment undertaken at date $t = 1$ has a useful life of T periods and can thus be used until $t = T + 1$. However, since we effectively assume that the market closes at

$t = T$ and firms will earn no revenue beyond that date, the asset is effectively impaired and therefore can be written off at $t = T$. As shown above, the zero-profit condition for the investment at date 0 requires that:

$$\sum_{t=1}^T E[P(w, \tilde{\epsilon}_t, \mathbf{K})] \cdot \gamma^t = LC \cdot \sum_{t=1}^T \gamma^t.$$

Since $K_1 \leq K_2 \leq \dots \leq K_T$, the expected equilibrium prices are weekly decreasing over time and

$$\sum_{t=2}^T E[P(w, \tilde{\epsilon}_t, K_t)] \gamma^{t-1} \leq LC \cdot \sum_{t=2}^T \gamma^{t-1}.$$

Recalling that $LC = w + \gamma + c \cdot \Delta$, we find that:

$$\Gamma_1(1 \mid \mathbf{K}) \leq \sum_{t=2}^{T-1} \{(1-\alpha) \cdot \Delta \cdot c + \alpha \cdot \hat{d}_{t-1} \cdot v\} \cdot \gamma^{t-1} + \{(1-\alpha) \cdot \Delta \cdot c + \alpha (\hat{d}_{T-1} + \hat{d}_T) \cdot v\} \gamma^{T-1} - v. \quad (32)$$

By definition of Δ and c ,

$$\sum_{t=2}^{T+1} \{(1-\alpha) \cdot \Delta \cdot c + \alpha \cdot \hat{d}_{t-1} \cdot v\} \gamma^{t-1} = v.$$

Therefore (32) can be expressed as:

$$\Gamma_1(1 \mid \mathbf{K}) \leq \alpha \cdot \hat{d}_T \cdot \gamma^{T-1} - [(1-\alpha) \cdot \Delta \cdot c + \alpha \cdot \hat{d}_T \cdot v] \gamma^T. \quad (33)$$

Thus it remains to show that the right-hand side of (33) is strictly negative, or equivalently,

$$\alpha \cdot \hat{d}_T < [(1-\alpha) \cdot \Delta \cdot c + \alpha \cdot \hat{d}_T \cdot v] \cdot \gamma. \quad (34)$$

Since

$$\Delta = \frac{1 - \alpha \cdot \sum_{t=1}^T \hat{d}_t \cdot \gamma^t}{1 - \alpha},$$

$$c = \frac{v}{\sum_{i=1}^T \gamma^i},$$

and $(1 - \gamma) = r \cdot \gamma$, the inequality in (34) becomes:

$$\frac{1}{\sum \gamma^i} [1 - \alpha \sum_{i=1}^T \hat{d}_i \cdot \gamma^i] > \alpha \cdot \hat{d}_T \cdot r \cdot \sum_{i=1}^T \gamma^i,$$

or equivalently:

$$1 - \alpha [\hat{d}_1 \cdot \gamma + \dots + \hat{d}_T [\gamma^T + r \sum_{i=1}^T \gamma^i]] > 0.$$

This last inequality indeed holds because for any T ,

$$\gamma^T + r \cdot \sum_{i=1}^T \gamma^i = 1,$$

and $d_1 \cdot \gamma + d_2 \cdot \gamma^2 + \dots d_{T-1} \cdot \gamma^{T-1} + d_T < 1$. ■

Proof of Corollary 2. If fixed operating costs are avoidable and the aggregate capacity level is equal to K , the expected equilibrium price in period t is equal to:

$$E [P(w', \tilde{\epsilon}, K)] = \int_{\underline{\epsilon}}^{\epsilon(K, w')} w' \cdot h(\epsilon) d\epsilon + \int_{\epsilon(K, w')}^{\bar{\epsilon}} \epsilon \cdot P^o(K) \cdot h(\epsilon) d\epsilon.$$

where $w' \equiv w + f$. We note that $\epsilon(K, w')$ is weakly increasing in w' . Similar to the argument in the proof of Proposition 2, we next demonstrate that $E [P(w', \tilde{\epsilon}, \cdot)]$ is increasing in w' . For any $w'_1 < w'_2$, we can write $E [P(w'_1, \tilde{\epsilon}, K)]$ as:

$$\int_{\underline{\epsilon}}^{\epsilon(K, w'_1)} w'_1 \cdot h(\epsilon) d\epsilon + \int_{\epsilon(K, w'_1)}^{\epsilon(K, w'_2)} \epsilon \cdot P^o(K) h(\epsilon) d\epsilon + \int_{\epsilon(K, w'_2)}^{\bar{\epsilon}} \epsilon \cdot P^o(K) h(\epsilon) d\epsilon. \quad (35)$$

Similarly $E [P(w, \tilde{\epsilon}, K_2)]$ can be expressed as:

$$\int_{\underline{\epsilon}}^{\epsilon(K, w'_1)} w'_2 h(\epsilon) d\epsilon + \int_{\epsilon(K, w'_1)}^{\epsilon(K, w'_2)} w'_2 h(\epsilon) d\epsilon + \int_{\epsilon(K, w'_2)}^{\bar{\epsilon}} \epsilon \cdot P^o(K) h(\epsilon) d\epsilon \quad (36)$$

Since $\epsilon \cdot P^o(K) \geq w$ for $\epsilon > \epsilon(K, w)$ and $w'_1 < w'_2$, we conclude that each of the three integrals in (35) is respectively at least as large as its counterpart in (36). It should be noted that the claimed monotonicity here is strict whenever $\epsilon(K, w'_2) > \underline{\epsilon}$.

Denote the equilibrium capacity levels with avoidable and unavoidable fixed costs by K^* and K^{**} , respectively, it follows directly from the arguments in Proposition 2 that:

$$E [P(w, \tilde{\epsilon}, K^*)] = E [P(w + f, \tilde{\epsilon}, K^{**})] = LC.$$

Since $E [P(w, \tilde{\epsilon}, K)]$ is decreasing in K and strictly increasing in w whenever $\epsilon(K, w + f) > \underline{\epsilon}$, we conclude that $K^{**} > K^*$ whenever

$$\text{Prob}[\tilde{\epsilon} \geq \frac{w + f}{LC}] < 1,$$

as claimed in Corollary 2. ■

Proof of Corollary 3.

Under the assumptions of Proposition 2, the expected net income per unit of capacity in equilibrium is given by:

$$\pi_t = LC - w - f - d_t \cdot v - \alpha \cdot I_t,$$

where I_t denotes the expected taxable income in period t :

$$I_t = LC - w - f - \hat{d}_t \cdot v.$$

With straight-line depreciation, $d_t = \frac{1}{T}$. Direct substitution yields:

$$\pi_t = (1 - \alpha)(LC - w - f) - v\left(\frac{1}{T} - \alpha \cdot \hat{d}_t\right).$$

The expression for c in equation (1) yields:

$$\pi_t = v \left[(1 - \alpha) \cdot \frac{\Delta}{\sum_{i=1}^T \gamma^i} - \frac{1}{T} + \alpha \cdot \hat{d}_t \right].$$

Recalling the definition of the tax factor Δ , we have

$$\pi_t = v \left[\frac{1 - \alpha \cdot \sum_{i=1}^T \hat{d}_i \cdot \gamma^i}{\sum_{i=1}^T \gamma^i} - \frac{1}{T} + \alpha \cdot \hat{d}_t \right].$$

Thus $\pi_t \geq 0$ if and only if

$$1 - \frac{\sum_{i=1}^T \gamma^i}{T} \geq \alpha \left[\sum_{i=1}^T \hat{d}_i \cdot \gamma^i - \hat{d}_t \cdot \sum_{i=1}^T \gamma^i \right] \geq 0.$$

This last inequality indeed holds whenever (13) is met, since:

$$\sum_{i=1}^T \hat{d}_i \cdot \gamma^i - \hat{d}_t \cdot \sum_{i=1}^T \gamma^i \leq 1,$$

for any depreciation rule $(\hat{d}_1, \dots, \hat{d}_T)$. ■

Proof of Proposition 3.

Denote the aggregate capacity investment at date 0 by K . We characterize the equilibrium level of K^* and the expected equilibrium prices, $E[P(w_t, \tilde{\epsilon}_t, \lambda, K^*)]$, under the assumption that no firm makes additional capacity investments at any date t , $1 \leq t < T$.

Arguments similar to the ones provided in the proof of Proposition 2 show that indeed no late entrants could break even on their investments. This is a fortiori true if the product market is declining over time and the unit variable costs are (weakly) increasing.

If the aggregate capacity level is K at date 0, the equilibrium price in period t is given by:

$$P(w, \epsilon_t, \lambda_t, K) = \begin{cases} \epsilon_t \cdot \lambda_t \cdot P^o(K) & \text{if } \epsilon_t \geq \epsilon(K, w_t, \lambda_t) \\ w_t & \text{if } \epsilon_t < \epsilon(K, w_t, \lambda_t). \end{cases}$$

Here $\epsilon(K, w_t, \lambda_t)$ denotes the unique cut-off level defined as follows:

$$\epsilon(K, w_t, \lambda_t) = \begin{cases} \bar{\epsilon} & \text{if } \bar{\epsilon} \cdot \lambda_t \cdot P^o(K) \leq w_t \\ \frac{w_t}{\lambda_t \cdot P^o(K)} & \text{if } \bar{\epsilon} \cdot \lambda_t \cdot P^o(K) > w_t > \underline{\epsilon} \cdot \lambda_t \cdot P^o(K) \\ \underline{\epsilon} & \text{if } \underline{\epsilon} \cdot \lambda_t \cdot P^o(K) \geq w_t \end{cases}$$

Clearly, $\epsilon(K, w_t, \lambda_t)$ is increasing in w_t , weakly increasing in K and decreasing in λ_t .

For an individual firm that has invested a capacity level k , there is again no loss of generality in assuming that this firm will also produce k units of output since the product market price never falls below the short-run marginal cost w_t . The present value of expected future cash flows for the representative firm then becomes:

$$\Gamma(k|K) = \sum_{t=1}^T \left\{ E \left[[\lambda_t \cdot P_t(w_t, \tilde{\epsilon}_t, \lambda_t, K) - w_t] \cdot k - f \cdot k - \alpha \cdot \tilde{I}_t \right] \right\} \gamma^t - v \cdot k,$$

with taxable income given by:

$$\tilde{I}_t = [\gamma^t \cdot P_t(w_t, \tilde{\epsilon}_t, \lambda_t, K) - w_t] \cdot k - f \cdot k - \hat{d}_t \cdot v \cdot k.$$

Solving the equation $\Gamma(k|K^*) = 0$ yields:

$$\sum_{t=1}^T \lambda_t \cdot E [P_t(w_t, \tilde{\epsilon}_t, \lambda_t, K)] \cdot \gamma^t = \sum_{t=1}^T \gamma^t \cdot LC, \quad (37)$$

where

$$\lambda_t \cdot E [P_t(w_t, \tilde{\epsilon}_t, \lambda_t, K)] = \int_{\underline{\epsilon}}^{\epsilon(K, w_t, \epsilon_t)} w_t \cdot h(\epsilon) d\epsilon + \int_{\epsilon(K, w_t, \epsilon_t)}^{\bar{\epsilon}} \epsilon \cdot \lambda_t \cdot P^o(K) \cdot h(\epsilon) d\epsilon.$$

Suppose now that $K^* \leq K^o$. We obtain:

$$\begin{aligned}
\sum_{t=1}^T \lambda_t \cdot E [P^o(w_t, \tilde{\epsilon}_t, \lambda_t, K^*)] \cdot \gamma^t &\geq \sum_{t=1}^T \lambda_t \cdot E [P^o(w_t, \tilde{\epsilon}_t, \lambda_t, K^o)] \cdot \gamma^t \\
&\geq \sum_{t=1}^T \lambda_t \cdot P^o(K^o) \cdot \gamma^t \\
&= LC \cdot m(\boldsymbol{\lambda}).
\end{aligned}$$

It follows that $K^* \geq K^o$ for otherwise (37) could not hold. Furthermore, $K^* = K^o$ requires that

$$\text{Prob} \left[\tilde{\epsilon} \cdot \lambda_T \geq \frac{w_t}{LC \cdot m(\boldsymbol{\lambda})} \right] = 1,$$

otherwise the second inequality in (38) could not hold. To establish the claim on the upper bound for K^* , we note that

$$\sum_{t=1}^T \lambda_t \cdot E [P_t(w_t, \tilde{\epsilon}_t, \lambda_t, K^*)] \cdot \gamma^t \leq \sum_{t=1}^T \left\{ \lambda_t \cdot P^o(K^*) \cdot [\bar{\epsilon} \cdot (1 - H(\epsilon(K^*, w, \lambda_t)))] + w_t \cdot H(\epsilon(K^*, w_t, \lambda_t)) \right\} \gamma^t. \tag{38}$$

We recall that with time-variant variable costs:

$$LC \equiv LFC + \frac{\sum_{t=1}^T w_t \cdot \gamma^t}{\sum_{t=1}^T \gamma^t}.$$

In conjunction with the inequality in (37), (38) yields:

$$\begin{aligned}
\sum_{t=1}^T LFC \cdot \gamma^t &\leq \sum_{t=1}^T [\bar{\epsilon} \cdot \lambda_t \cdot P^o(K^*) - w_t] (1 - H(\epsilon(K^*, w_t, \lambda_t))) \cdot \gamma^t \\
&\leq \sum_{t=1}^T [\bar{\epsilon} \cdot \lambda_t \cdot P^o(K^*) - w_t] \cdot \gamma^t \\
&= \sum_{t=1}^T \lambda_t \cdot \gamma^t \cdot \bar{\epsilon} \cdot P^o(K^*) - \sum_{t=1}^T w_t \cdot \gamma^t.
\end{aligned}$$

Since

$$LC \cdot \sum_{t=1}^T \gamma^t = \sum_{t=1}^T w_t \cdot \gamma^t + LFC \cdot \sum_{t=1}^T \gamma^t,$$

we conclude that

$$LC \cdot m(\boldsymbol{\lambda}) \leq \bar{\epsilon} \cdot P^o(K^*),$$

or equivalently,

$$K^* \leq D^o \left(m(\boldsymbol{\lambda}) \cdot \frac{LC}{\bar{\epsilon}} \right).$$

The proof that K^* must also satisfy

$$K^* \leq D^o (m(\boldsymbol{\lambda}) \cdot LFC)$$

proceeds analogously to the one given in connection with Proposition 2. ■

Proof of Proposition 4.

Given any capacity investment k at date 0, the monopolist will choose the optimal quantity in period t as

$$q^*(w, \epsilon_t, k) = \begin{cases} k & \text{if } \epsilon_t \geq \epsilon(k, w) \\ q(w, \epsilon_t) & \text{if } \epsilon_t < \epsilon(k, w), \end{cases}$$

where $q(w, \epsilon)$ solves the equation: $\epsilon \cdot MR^o(q(w, \epsilon)) = w$, and

$$\epsilon(k, w) = \begin{cases} \bar{\epsilon} & \text{if } \bar{\epsilon} \cdot MR^o(k) \leq w \\ \frac{w}{MR^o(k)} & \text{if } \bar{\epsilon} \cdot MR^o(k) \geq w \geq \underline{\epsilon} \cdot MR^o(k) \\ \underline{\epsilon} & \text{if } \underline{\epsilon} \cdot MR^o(k) \geq w \end{cases}$$

We note that $\epsilon(k, w)$ is increasing in k and that $q(w, \epsilon) < k$ whenever $\epsilon < \epsilon(k, w)$.

For any capacity level k , the present value of expected cash flows becomes:

$$\Gamma(k) = \sum_{t=1}^T \left\{ E \left[\tilde{\epsilon}_t \cdot R^o(q^*(w, \tilde{\epsilon}_t, k)) - w \cdot q^*(w, \tilde{\epsilon}_t, k) - f \cdot k - \alpha \cdot \tilde{I}_t \right] \right\} \cdot \gamma^t - v \cdot k,$$

where

$$\tilde{I}_t = \tilde{\epsilon}_t \cdot R^o(q^*(\cdot)) - w \cdot q^*(\cdot) - f \cdot k - \hat{d}_t \cdot v \cdot k.$$

Define

$$\hat{\Gamma}(k) = \frac{\Gamma(k)}{(1 - \alpha) \cdot \sum_{t=1}^T \gamma^t}.$$

Recalling the definition of the unit capacity cost, c , the tax factor, Δ , and the levelized fixed cost, LFC , we obtain:

$$\hat{\Gamma}(k) = E[\tilde{\epsilon} \cdot R^o(q^*(w, \tilde{\epsilon}, k)) - w \cdot q^*(w, \tilde{\epsilon}, k)] - LFC \cdot k.$$

Thus,

$$\hat{\Gamma}'(k) = \int_{\epsilon(k,w)}^{\bar{\epsilon}} [\epsilon \cdot MR^o(k) - w] h(\epsilon) d\epsilon - LFC, \quad (39)$$

where $MR^o(k) = \frac{d}{dk}R^o(\cdot)$. Since for any random variable X , $E[f(X)] \leq E[\max\{0, f(X)\}]$, it follows that

$$\begin{aligned} \hat{\Gamma}'(k) &\geq \int_{\underline{\epsilon}}^{\bar{\epsilon}} [\epsilon \cdot MR^o(k) - w] h(\epsilon) d\epsilon - LFC \\ &= MR^o(k) - w - LFC \end{aligned} \quad (40)$$

$$= MR^o(k) - LC. \quad (41)$$

By definition $MR^o(k^o) \equiv LC$. Therefore $k^* \geq k^o$. We next demonstrate that $k^* = k^o$ if and only if the condition

$$\text{Prob} \left[\tilde{\epsilon} \geq \frac{w}{LC} \right] = 1 \quad (42)$$

holds. This condition is equivalent to $\epsilon(k^o, w) = \underline{\epsilon}$, because $\epsilon(k^o, w) > \underline{\epsilon}$ if and only if $\underline{\epsilon} \cdot MR^o(k^o) < w$. Referring back to (39), we obtain:

$$\hat{\Gamma}'(k^o) = MR^o(k^o) - LC = 0$$

if and only if condition (42) is met. To demonstrate the upper bound on k^* claimed in Proposition 4, we note that

$$\begin{aligned} \hat{\Gamma}'(k) &\leq [MR^o(k) \cdot \bar{\epsilon} - w] \cdot [1 - H(\epsilon(k, w))] - LFC \\ &\leq \bar{\epsilon} \cdot MR^o(k) - LC. \end{aligned}$$

Therefore $\hat{\Gamma}'(k) < 0$ for any $k > k^+$ with k^+ given by $MR^o(k^+) = \frac{LC}{\bar{\epsilon}}$.

To show that an upper bound on capacity is also provided by k^+ defined by the requirement that $MR^o(k^+) = LFC \equiv LC - w$, we first state an auxiliary result which ranks the maximizers of two functions, $\Gamma_1(k)$ and $\Gamma_2(k)$. Suppose both functions are differentiable on the positive real line.

Auxiliary Lemma: *If k_1^* is the unique maximizer of $\Gamma_1(k)$ and k_2^* is the unique maximizer of $\Gamma_2(k)$, then $k_2^* \geq k_1^*$ provided $\Gamma_2'(k) \geq \Gamma_1'(k)$ for all k .*

Proof of Auxiliary Lemma: Suppose to the contrary that $k_1^* > k_2^*$. By definition:

$$\Gamma_1(k_1^*) > \Gamma_1(k_2^*) \text{ and } \Gamma_2(k_2^*) > \Gamma_2(k_1^*).$$

Adding these two inequalities yields:

$$\Gamma_1(k_1^*) - \Gamma_1(k_2^*) > \Gamma_2(k_1^*) - \Gamma_2(k_2^*),$$

or

$$\int_{k_2^*}^{k_1^*} [\Gamma_1'(u) - \Gamma_2'(u)] du > 0.$$

Since $\Gamma_1'(u) \leq \Gamma_2'(u)$, we obtain a contradiction. ■

To complete the proof of Proposition 4, we identify $\Gamma_1(\cdot)$ in the Auxiliary Lemma with the monopolist's problem, that is, $\Gamma_1(k) = \hat{\Gamma}(k)$. Let $\Gamma_2(k)$ be given by the corresponding function when $w = 0$. Thus $\Gamma_2(k) = \hat{\Gamma}_2(k|w = 0)$. If $w = 0$, the firm always produces at capacity and therefore k_2^* is such that $MR^o(k_2^*) = LFC$. To invoke the Auxiliary Lemma, we verify that

$$\begin{aligned} \Gamma_2'(k) &= MR^o(k) - LFC \\ &\leq \Gamma_1'(k) = \int_{\epsilon(k,w)}^{\bar{\epsilon}} [\epsilon \cdot MR^o(k) - w] h(\epsilon) d\epsilon - LFC. \end{aligned}$$

This inequality indeed holds because

$$\frac{\partial}{\partial w} \int_{\epsilon(k,w)}^{\bar{\epsilon}} [\epsilon \cdot MR^o(k) - w] h(\epsilon) d\epsilon = -[1 - H(\epsilon(k,w))] \leq 0,$$

and

$$MR^o(k) = \int_{\epsilon(k,w=0)}^{\bar{\epsilon}} \epsilon \cdot MR^o(k) \cdot h(\epsilon) d\epsilon.$$

Thus the monopolist's optimal capacity choice satisfies $k^* \leq k_2^*$, where $MR^o(k_2^*) = LFC$.

Combining the two upper bounds on capacity, we conclude that $k^* \leq k^+$, where

$$MR^o(k^+) = LC^- = \max \left\{ LFC, \frac{LC}{\bar{\epsilon}} \right\}.$$

■

Proof of Corollary 4: To show that the monopolist will choose a higher level of capacity if market volatility increases in the sense of a mean-preserving spread, we again invoke the Auxiliary Lemma in the Proof of Proposition 4. Specifically, we show that

$$\hat{\Gamma}'(k|h_2(\cdot)) \geq \hat{\Gamma}'(k|h_1(\cdot))$$

for all k . As shown in connection with Proposition 4:

$$\begin{aligned}\Gamma'(k|h_1(\cdot)) &= \int_{\underline{\epsilon}(k,w)}^{\bar{\epsilon}} [\epsilon \cdot MR^o(k) - w] h_1(\epsilon_1) d\epsilon_1 - LFC \\ &= \int_{\underline{\epsilon}}^{\bar{\epsilon}} V(\epsilon_1) \cdot h_1(\epsilon_1) d\epsilon_1 - LFC\end{aligned}$$

where³⁰

$$V(\epsilon) \equiv \begin{cases} 0 & \text{if } \epsilon(k, w) \cdot MR^o(k) < w \\ \epsilon \cdot MR^o(k) - w & \text{if } \epsilon(k, w) \cdot MR^o(k) \geq w. \end{cases}$$

If $\tilde{\epsilon}_2$ is obtained by a mean-preserving spread of $\tilde{\epsilon}_1$, then $\tilde{\epsilon}_2 = \tilde{\epsilon}_1 + \tilde{\mu}$ where $\tilde{\mu}$ is an unbiased random variable such that

$$\int_{\underline{\mu}}^{\bar{\mu}} \mu \cdot g(\mu|\epsilon_1) d\mu = 0$$

for all ϵ_1 . Since $h_2(\epsilon_2) = g(\mu|\epsilon_1) \cdot h_1(\epsilon_1)$, we obtain:

$$\begin{aligned}\hat{\Gamma}'(k|h_2(\cdot)) &= \int_{\underline{\epsilon}_1}^{\bar{\epsilon}_1} \int_{\underline{\mu}}^{\bar{\mu}} V(\epsilon_1 + \mu) \cdot g(\mu|\epsilon_1) \cdot h_1(\epsilon_1) d\mu d\epsilon_1 - LFC \\ &\geq \int_{\underline{\epsilon}_1}^{\bar{\epsilon}_1} V\left(\epsilon_1 + \int_{\underline{\mu}}^{\bar{\mu}} \mu \cdot g(\mu|\epsilon_1) d\mu\right) h_1(\epsilon_1) d\epsilon_1 - LFC \\ &= \int_{\underline{\epsilon}_1}^{\bar{\epsilon}_1} V(\epsilon_1) h_1(\epsilon_1) d\epsilon_1 - LFC \\ &= \hat{\Gamma}'(k|h_1(\cdot)).\end{aligned}$$

The preceding inequality follows from Jensen's inequality and the convexity of $V(\cdot)$ in ϵ . Thus, $k_2^* \geq k_1^*$.³¹ ■

Proof of Proposition 5.

Claim 1: For every pair of capacity investments (k^1, k^2) and every realization of ϵ_t , there exists a unique Nash equilibrium

$$(q_t^1(k^1, k^2, \epsilon_t), q_t^2(k^1, k^2, \epsilon_t))$$

in period t .

Proof of Claim 1: Define $q(\epsilon_t)$ by the requirement that

$$\epsilon_t \cdot MR^o(q(\epsilon_t)|q(\epsilon_t)) = w.$$

³⁰For notational simplicity, we suppress the dependence of $V(\cdot)$ on k and w .

³¹The Auxiliary Lemma requires unique maximizers which follows from the fact that $\Gamma'(\cdot|h_i(\cdot))$ is strictly decreasing.

Thus $q(\epsilon_t)$ denotes the equilibrium quantities the parties would supply in a symmetric, capacity-unconstrained equilibrium. We distinguish three scenarios.

Scenario 1: $k^i > q(\epsilon_t)$ for $1 \leq i \leq 2$. In this scenario, both firms will supply $q_t = q(\epsilon_t)$ in equilibrium. To see that this equilibrium is unique, suppose, to the contrary, that another pair (q^1, q^2) constitutes a Nash equilibrium in period t . Then, if $q^i < k^i$, it must be that

$$\epsilon_t \cdot MR^o(q^1|q^2) = w = \epsilon_t \cdot MR^o(q^2|q^1)$$

which implies $q^i = q(\epsilon_t)$. Similarly, there cannot be an equilibrium in which either $q^i = k^i$ or both.

Scenario 2: $k^i \leq q(\epsilon_t)$ for $1 \leq i \leq 2$. In this case, the unique equilibrium is for both firms to supply $q_t^i = k^i$ in period t because

$$\epsilon_t \cdot MR^o(q^i|k^j) \geq w$$

for all $q^i \leq k^i$.

Scenario 3: $k^i > q(\epsilon_t)$ and $k^j \leq q(\epsilon_t)$. It is then a Nash equilibrium for firm j to supply k^j and for firm i to choose:

$$q_t^i \in \operatorname{argmax}_{q \leq k^i} \{ \epsilon_t \cdot R^o(q|k^j) - w \cdot q \}.$$

Clearly, there cannot be an equilibrium in which $q^j < k^j$, regardless of whether firm i exhausts its available capacity or not.

Claim 2: *Given the limited price volatility condition, suppose that firm j chooses the capacity level $k^j = k^*$. In equilibrium, firm j will then supply $q_t^j = k^*$ for all ϵ_t and any capacity level k^i .*

Proof of Claim 2: If $k^i \leq k^*$, the condition of limited price volatility ensures that

$$\begin{aligned} & \epsilon_t \cdot MR^o(k^*|k^i) - w \\ & \geq \epsilon_t \cdot MR^o(k^*|k^*) - w \\ & = \epsilon_t \cdot LC - w \geq 0. \end{aligned}$$

Thus firm j will supply $k^j = k^*$ for any ϵ_t and any q^i which satisfies $q^i \leq k^i \leq k^*$.

If $k^i > k^*$ and $k^j = k^*$, the unique Nash equilibrium $(q_t^1(k^1, k^2, \epsilon_t), q_t^2(k^1, k^2, \epsilon_t))$, given ϵ_t , must correspond to either Scenario 2 or 3 in the proof of Claim 1. To see this, we note that in Scenario 1:

$$k^i > k^j = k^* > q(\epsilon_t),$$

and

$$\epsilon_t \cdot MR^o(q(\epsilon_t)|q(\epsilon_t)) = w.$$

Yet that would lead to a contradiction on account of the inequalities:

$$\epsilon_t \cdot MR^o(q(\epsilon_t)|q(\epsilon_t)) > \epsilon_t \cdot MR^o(k^*|k^*) = \epsilon_t \cdot LC > w.$$

Thus, either Scenario 2 or 3 applies if $k^j = k^*$ and, as shown in Claim 1, in either one of these $q_t^j(k^i, k^j = k^*, \epsilon_t) = k^*$.

Claim 3: *Given $k^j = k^*$, it is a best response for firm i to choose the initial capacity level $k^i = k^*$.*

Proof of Claim 3: The argument adapts the notation in the proof of Proposition 4. Given any k^i and $q^i(k^i, k^j = k^*, \epsilon_t)$, firm i 's Nash equilibrium quantity in period t is given by

$$q^i(k^i, k^*, \epsilon_t) = \begin{cases} k^i & \text{if } \epsilon_t \geq \epsilon(k^i|k^*) \\ q(k^*, \epsilon_t) & \text{if } \epsilon_t \leq \epsilon(k^i|k^*) \end{cases}$$

where $q(k^*, \epsilon_t)$ is uniquely determined by

$$\epsilon_t \cdot MR^o(q(k^*, \epsilon_t)|k^*) = w,$$

and

$$\epsilon(k^i|k^*) = \begin{cases} \bar{\epsilon} & \text{if } \bar{\epsilon} \cdot MR^o(k^i|k^*) \leq w \\ \frac{w}{MR^o(k^i|k^*)} & \text{if } \bar{\epsilon} \cdot MR^o(k^i|k^*) \geq w \geq \underline{\epsilon} \cdot MR^o(k^i|k^*) \\ \underline{\epsilon} & \text{if } \underline{\epsilon} \cdot MR^o(k^i|k^*) \geq w. \end{cases}$$

It follows that $\epsilon(k^i|k^*)$ is increasing in k^i and $q(k^*, \epsilon_t) < k^i$ if and only if $\epsilon_t < \epsilon(k^i|k^*)$.

Given $k^j = k^*$, firm i 's objective function can then be expressed as

$$\Gamma(k^i|k^*) = \sum_{t=1}^T \left\{ E \left[\tilde{\epsilon}_t \cdot R^o(q^i(k^i, k^*, \epsilon_t)|k^*) - w \cdot q^i(k^i, k^*, \epsilon_t) - f \cdot k^i - \alpha \cdot \tilde{I}_t \right] \right\} \gamma^t - v \cdot k^i,$$

where

$$\tilde{I}_t = \epsilon_t \cdot R^o(q^i(\cdot)|k^*) - w \cdot q^i(\cdot) - f \cdot k^i - d_t \cdot v \cdot k^i.$$

Using the same normalization as in the proof of Proposition 4, we define:

$$\hat{\Gamma}(k^i|k^*) = \Gamma(k^i|k^*) \cdot \frac{1}{(1 - \alpha) \cdot \sum_{t=1}^T \gamma^t},$$

with

$$\hat{\Gamma}(k^i|k^*) = E [\tilde{\epsilon} \cdot R^o(q^i(k^i, k^*, \tilde{\epsilon})|k^*) - w \cdot q^i(k^i, k^*, \tilde{\epsilon})] - LFC \cdot k^i.$$

It follows that

$$\frac{\partial}{\partial k^i} \hat{\Gamma}(k^i|k^*) = \int_{\epsilon(k^i|k^*)}^{\bar{\epsilon}} [\epsilon \cdot MR^o(k^i|k^*) - w] h(\epsilon) d\epsilon - LFC. \quad (43)$$

The right-hand side of (43) is decreasing in k^i . Furthermore, the limited price volatility condition implies $\epsilon(k^*|k^*) = \underline{\epsilon}$. Since, by definition, $MR^o(k^*|k^*) = LC$, it follows that:

$$\left. \frac{\partial}{\partial k^i} \hat{\Gamma}(k^i|k^*) \right|_{k^i=k^*} = 0,$$

which shows that $k^i = k^*$ is a best response to $k^j = k^*$. Thus, $(k^1, k^2) = (k^*, k^*)$ followed by $(q_t^1, q_t^2) = (k^*, k^*)$ in each period is a subgame perfect equilibrium outcome. ■

References

- Al-Najjar, N., S. Baliga, and D. Besanko (2008), “Market Forces Meet Behavioral Biases: Cost Misallocation and Irrational Pricing,” *RAND Journal of Economics*, 39, 214–237.
- Atkinson, A. A., R. S. Kaplan, E. M. Matsumura, and S. M. Young (2011), *Management Accounting: Information for Decision-Making and Strategy Execution*, 6th ed., Prentice Hall, New York.
- Balakrishnan, R., and K. Sivaramakrishnan (2001), “Sequential Solutions to Capacity Planning and Pricing Decisions,” *Contemporary Accounting Research*, 1–26.
- Balakrishnan, R., and K. Sivaramakrishnan (2002), “A Critical Overview of Full-Cost Data for Planning and Pricing,” *Journal of Management Accounting Research*, 14, 3–31.
- Baldenius, T., S. Reichelstein, and S. A. Sahay (1999), “Negotiated versus Cost-Based Transfer Pricing,” *Review of Accounting Studies*, 6, 67–91.
- Banker, R. D., and J. S. Hughes (1994), “Product Costing and Pricing,” *The Accounting Review*, 69, 479–494.
- Banker, R. D., I. Hwang, and B. K. Mishra (2002), “Product Costing and Pricing under Long-Term Capacity Commitment,” *Journal of Management Accounting Research*, 14, 79–97.
- Borenstein, S. (2000), “Understanding Competitive Pricing and Market Power in Wholesale Electricity Markets,” *The Electricity Journal*, 13, 49–57.
- Borenstein, S. (2008), “The Market Value and Cost of Solar Photovoltaic Electricity Production,” *CSEM Working Paper No. 176*.
- Bouwens, J., and B. Steens (2008), “The Economics of Full Cost Transfer Pricing,” *Working Paper, Tilburg University*.
- Campbell, M. (2008), “The Drivers of the Levelized Cost of Electricity for Utility-Scale Photovoltaics,” *White Paper: SunPower Corporation*.

- Campbell, M. (2011), “Charting the progress of PV Power Plant Energy Generating Costs to Unsubsidized Levels, Introducing the PV-LCOE Framework,” *White Paper: Sun Power Corporation*.
- Carlton, D. W., and J. M. Perloff (2005), *Modern Industrial Organization*, 4th ed., Pearson/Addison Wesley, New York.
- Coenenberg, A. G., T. M. Fischer, and T. Günther (2012), *Kostenrechnung und Kostenanalyse*, 8th ed., Schäffer-Poeschel, Stuttgart.
- Cooper, R., and R. S. Kaplan (1988), “How Cost Accounting Distorts Product Costs,” *Management Accounting*, 69, 20–28.
- Davidson, C., and R. Deneckere (1986), “Long-Run Competition in Capacity, Short-Run Competition in Price, and the Cournot Model,” *RAND Journal of Economics*, 17, 404–415.
- Drury, C. S., B. P. Osborne, and M. Tayles (1993), *A Survey of Management Accounting Practices in the U.K. Manufacturing Companies*, ACCA Publications, London, U.K.
- Dutta, S., and S. Reichelstein (2010), “Decentralized Capacity Management and Internal Pricing,” *Review of Accounting Studies*, 15, 442–478.
- Eccles, R. (1985), *The Transfer Pricing Problem: A Theory for Practice*, Lexington, KY: Lexington Books.
- EPIA (September 2011), “Solar Photovoltaics Competing in the Energy Sector - On the Road to Competitiveness,” .
- Ewert, R., and A. Wagenhofer (2008), *Interne Unternehmensrechnung*, 7th ed., Springer, Heidelberg.
- Govindarajan, V., and R. Anthony (1983), “How Firms Use Cost Data in Price Decisions,” *Management Accounting*, 64, 30–36.
- Göx, R. (2002), “Capacity Planning and Pricing under Uncertainty,” *Journal of Management Accounting Research*, 14, 59–78.

- Grant, S., and J. Quiggin (1996), "Capital Precommitment and Competition in Supply Schedules," *Journal of Industrial Economics*, XLIV, 427–441.
- Horngren, C. T., S. M. Datar, and M. V. Rajan (2012), *Cost Accounting - A Managerial Emphasis*, 14th ed., Prentice Hall, Boston New York San Francisco.
- Hviid, M. (1991), "Capacity Constrained Duopolies, Uncertain Demand, and Non-Existence of Pure Strategy Equilibria," *European Journal of Political Economy*, 7, 183–190.
- Johnson, N. B. (2006), "Divisional Performance Management and Transfer Pricing for Intangible Assets," *Review of Accounting Studies*, 11, 339–365.
- Kreps, D. M., and J. A. Scheinkman (1983), "Quantity Precommitment and Bertrand Competition Yield Cournot Outcomes," *Bell Journal of Economics*, 14, 326–337.
- Küpper, H.-U. (1985), "Investitionstheoretische Fundierung der Kostenrechnung," *Zeitschrift für betriebswirtschaftliche Forschung*, 37, 26–46.
- Küpper, H.-U. (2009), "Investment-Based Cost Accounting as a Fundamental Basis of Decision-Oriented Management Accounting," *Abacus*, 45, 249–274.
- Laffont, J. J., and J. Tirole (2000), *Competition in Telecommunications*, MIT Press, Cambridge, MA.
- Mas-Colell, A., M. D. Whinston, and J. R. Green (1995), *Microeconomic Theory*, Oxford University Press.
- MIT (2007), "The Future of Coal," <http://web.mit.edu/coal>.
- Nezlobin, A. A. (2012), "Accrual Accounting, Informational Sufficiency, and Equity Valuation," *Journal of Accounting Research*, 50, 233–273.
- Nezlobin, A. A., M. V. Rajan, and S. Reichelstein (2012), "Dynamics of Rate-of-Return Regulation," *Management Science*, 58, 980–995.
- Osborne, M., and C. Pitchik (1986), "Price Competition in a Capacity-Constrained Duopoly," *Journal of Economic Theory*, 38, 238–260.
- Pfeiffer, T., U. Schiller, and J. Wagner (2011), "Cost-based Transfer Pricing," *Review of Accounting Studies*, 16, 219–246.

- Pittman, B. (2009), “Whom Are You Calling Irrational? Marginal Costs, Variable Costs, and the Pricing Practices of Firms,” *Working Paper, Department Of Justice, Washington DC*.
- Preinreich, G. A. D. (1938), “Annual Survey of Economic Theory: The Theory of Depreciation,” *Econometrica*, 6, 219–241.
- Rajan, M., and S. Reichelstein (2009), “Depreciation Rules and The Relation Between Marginal and Historical Cost,” *Journal of Accounting Research*, 47, 1–43.
- Reichelstein, S., and M. Yorston (2013), “The Prospects for Cost Competitive Solar PV Power,” *Energy Policy*, 55, 117–127.
- 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, Article 2.
- Ross, S., R. Westerfield, and J. Jaffe (2005), *Corporate Finance*, McGrawHill Publishers, New York, N.Y.
- Shim, E., and E. F. Sudit (1995), “How Manufacturers Price Products,” *Management Accounting*, 76, 37–39.
- Sutton, J. (1991), *Sunk Costs and Market Structure*, MIT Press, Cambridge, MA.
- Werner, T. (2012), “Solar Power and Your Future,” Presentation at the Stanford Graduate School of Business.
- Zimmerman, J. (2010), *Accounting for Decision Making and Control*, 7th ed., McGraw-Hill/Irwin, New York.