The Economic Significance of Historic Cost

Madhav V. Rajan*  Stefan Reichelstein*

February 11, 2008

Abstract: Inferences about the economic cost of a product are frequently based on measures of historic cost. We examine a model in which a firm makes a sequence of overlapping capacity investments. Earlier research has identified particular accrual accounting (depreciation) rules which have the property that on a per unit basis the historic cost of a product captures precisely its marginal cost. Relative to this benchmark, we investigate the direction and magnitude of the bias in reported historic cost that results from alternative depreciation rules, including in particular straight-line depreciation for tangible assets and direct expensing for intangible assets. Our analysis shows that for a reasonable range of parameter specifications the resulting bias is rather small. Our framework lends itself to an analysis of the Accounting Profit Margin and the extent to which this financial ratio is indicative of market power, as measured by the Lange-Lerner index.

*Graduate School of Business, Stanford University. Email: mrajan@gsb.stanford.edu; reichelstein@stanford.edu. We would like to thank Julia Nasev and Jim Ohlson for detailed comments on this manuscript. We also acknowledge helpful comments by seminar participants at the following institutions: Chicago-Minnesota Theory Conference, Interdisciplinary Accounting Conference in Copenhagen, Harvard Business School, Stockholm School of Economics, European School of Management and Technology in Berlin and City University, London. Finally, we received valuable research assistance from Ian Gow, Alexander Nezlobin and Yanruo Wang.
1 Introduction

The historic cost principle is ubiquitous in accounting. Under Generally Accepted Accounting Principles (GAAP), most tangible assets are recorded at historic cost on the balance sheet. Subsequent depreciation and amortization expenses in the income statement then reflect the initial expenditure of past asset acquisitions. Economists understandably worry that such expenses reflect largely costs which are sunk and therefore have little economic relevance with regard to current economic decisions.\(^1\) Yet, in numerous contexts economic inferences and decisions are based on measures of historic cost. In the industrial organization and antitrust literature, reported Profit-Sales ratios have long been viewed as indicators of a firm’s market power.\(^2\) In particular, it has been suggested that the Profit-Sales ratio can be regarded as a proxy for the Lange-Lerner index of monopoly power.\(^3\) However, since the Lange-Lerner index is principally based on the concept of marginal cost, some economists have argued that the use of reported Profit-Sales ratios is fundamentally flawed in this context, precisely because accounting profit is typically based on the historic cost of fixed production assets (Fisher, 1987).

Cost-based pricing is another context where valuations based on historic cost determine current resource allocations. In many regulated industries, including utilities, telecommunications and defense contracting, prices are set so as to cover the firm’s cost of production, measured on a historic cost basis, and to provide the regulated firm with an appropriate return on its invested capital. A final example of where historic cost accounting guides current economic decisions is managerial performance measurement. It is common practice to measure the financial performance of business units by metrics such as Income, Return-on-Assets (ROA), or Economic Value Added (EVA). While the practitioner oriented literature on value-based management has advocated select “accounting adjustments” for the purpose of internal performance measurement, the proposed metrics remain for the most part grounded in the historic cost principle.\(^4\)

---

\(^1\)For instance, Kreps (2004) articulates this view as follows: “…the strange accounting practice of depreciating durable assets, where firms… take a charge against current income for depreciation of assets, using such rules as straight line depreciation.” Without offering detailed answers, Kreps then poses the following questions: “What, if anything, does accounting income have to do with economic income? Why do accountants depreciate assets? From an economic perspective, what sins of omission or commission do accountants commit?”

\(^2\)According to Long et al. (1982), the Federal Trade Commission’s line of business program proposed that Profit-Sales ratio data be used as screening indicators for possible antitrust investigations.

\(^3\)See, for instance, Domowitz et al. (1986), Mueller (1986) or Martin (2002).

\(^4\)Young and O’Byrne (2001) survey different such metrics. The literature on managerial incentives has pointed out that proper matching of revenues and expenses is essential in order to motivate managers with
To examine the economic significance of historic cost measures, we consider a setting in which a firm makes a sequence of overlapping investments in production capacity. The capacity generated by new investments may decline in subsequent periods. At any given point in time, the overall capacity available is the aggregate of the practical capacity levels left from past investments. For such a setting, Arrow (1964) has shown how to calculate the marginal cost of procuring another unit of capacity for the next period. The main insight in Arrow’s analysis is that, even though capacity related expenditures are by their very nature common costs that generate benefits over multiple periods, the forward looking long-run cost is separable and linear and therefore permits a unit cost calculation for one unit of capacity made available for one period.

Recent work by Rogerson (2005, 2007) provides an important insight on the economic significance of historic cost by showing that for a suitably chosen depreciation policy the firm’s average historic cost of capacity will be equal to the marginal cost of another unit of capacity. The measure of historic cost includes both current depreciation and interest charges, the latter imputed on the remaining book value of past investment expenditures. The depreciation method that aligns historic with marginal cost will be referred to as the Relative Practical Capacity (R.P.C.) rule. It has a transparent interpretation in the context of capacity investments: the expenditure required to create one unit of capacity today is allocated to future periods such that the historic cost charge to a particular period is in proportion to that period’s capacity relative to the discounted total capacity available over the useful life of the asset.5

This paper investigates how the historic cost of capacity relates to the underlying marginal cost for a range of possible accounting rules. In particular, we are mindful of the fact that for external reporting purposes most firms may adhere to straight-line depreciation for tangible assets like plant property and equipment and that investment expenditures for many intangible assets are expensed under GAAP. Accordingly, we examine depreciation schedules that are either more accelerated or decelerated than the Relative Practical Capacity rule. While one would no longer expect the average historic of capacity to align with the marginal cost, it is essential to understand both the direction and magnitude of the resulting deviations.

---

5The R.P.C. rule is analogous the so called Relative Benefit Depreciation rule that has played a prominent role in earlier work on goal congruent performance measures; see, for example, Rogerson (1997) and Reichelstein (1997). It should be noted, however, that in contrast to the R.P.C. rule, the relative benefit rule requires an understanding of the expected future cash benefits resulting from an investment.
We find that the behavior of the average historic cost is dependent on both the accrual accounting rules and the rate at which new capacity investments have been growing. Specifically, we obtain the following quadrant result: the average historic cost per unit of capacity is lower than the marginal cost if the depreciation schedule is more accelerated than that corresponding to the Relative Practical Capacity rule and, at the same time, new investments have been growing “moderately,” i.e., at a rate less than the firm’s cost of capital, in each period. In contrast, if past growth rates consistently exceeded the cost of capital (“aggressive growth”), the average historic cost will exceed the marginal cost of capacity. On the other hand, a depreciation schedule that is decelerated relative to the R.P.C. rule has the effect of precisely reversing these relations.

The quadrant result we obtain with regard to the average historic cost has distinct parallels to the one obtained in connection with a firm’s Return-on-Investment (ROI) in relation to the Internal Rate of Return (IRR) of its investment projects. In particular, it has been shown that if the accounting is neutral or operations have been growing exponentially at a particular rate, the two rates of return will coincide. It should be noted, however, that the two quadrants have different “boundaries.” First, the R.P.C. rule generally does not amount to neutral accounting. In addition, the separation point between moderate and aggressive growth is the firm’s cost of capital for the average historic cost function, while it is the underlying internal rate of return for the ROI metric.

Assuming that the accounting rules call for direct expensing of a share of new investments and straight-line depreciation for the remaining share of investments that are capitalized in the first place, we provide a comprehensive analysis of the percentage error that results from using the average historic cost as a “proxy” for the marginal cost of capacity. Our findings indicate that the resulting percentage error is remarkably small for a wide range of parameter values in terms of the useful life of assets, the growth rate in new investments and the physical decay of productive capacity. At the same time, we find that sharp increases in the percentage of new investments directly expensed can cause a substantial wedge between marginal- and average historic cost.

Our model framework lends itself to an analysis of the Accounting Profit Margin. To that end, we introduce sales revenues which the firm obtains in any given period from the presently available capacity. The quadrant result relating average historic cost to marginal cost then

\[ \text{error} = \frac{\text{average historic cost} - \text{marginal cost}}{\text{marginal cost}} \]


An intuitive way of seeing this is to note that neutral depreciation must reflect the underlying internal rate of return of a firm’s investment projects.
maps in a one-to-one fashion to a relation between the Accounting Profit margin and the Lange-Lerner index, i.e., the ratio whose numerator is the difference between average sales price and marginal cost and whose denominator is the average sales price. In particular, if the firm were to follow the R.P.C. rule, the Accounting Profit Margin for a single product is predicted to be equal to the price elasticity of demand, because a profit maximizing monopolist sets product prices so that the Lange-Lerner index is equal to the price elasticity of demand.

Most of our analysis treats growth as an exogenous variable. The final part of the paper expands the analysis to settings where the firm is a profit-maximizing monopolist whose capacity investments are made optimally in response to changes in the underlying product market. The growth rate in this market and any biases in the accounting rules are shown to determine jointly the behavior of the Accounting Profit Margin relative to the Lange-Lerner index of monopoly power. In contrast to earlier studies, such as those by Fisher (1987), our results again indicate that there is a systematic relation between these two ratios which pivots on the relation between historic and marginal cost.

The remainder of the paper is organized as follows. Section 2 lays out the model of overlapping capacity investments and reviews the benchmark results by Arrow (1964) and Rogerson (2005, 2007). Section 3 derives our main results regarding the behavior of the average cost of capacity. The relation between the Lange-Lerner ratio and the Accounting Profit Margin is analyzed in Section 4. We conclude in Section 5.

### 2 Historic- and Marginal Cost

We consider a firm that undertakes a sequence of investments in productive capacity. Each investment has a useful life of $T$ periods, though capacity derived from new investments may decline over time. At any given point in time, the total capacity currently available is determined by the investments over the past $T$ periods. New investments come “on-line” with a lag of one period and require a cash outlay of $v$ dollars for one unit of capacity. Specifically, an expenditure of $v \cdot I_i$ dollars at date $i$ will add productive capacity to produce $x_{t} \cdot I_i$ units of output at date $i + t$ for $1 \leq t \leq T$. To allow for the possibility of decaying capacity, possibly to reflect the need for increased maintenance and repair time, we specify $x_T \leq x_{T-1} \leq \ldots \leq x_1 = 1$.

Our analysis initially treats growth as an exogenous variable. It will be convenient to normalize the initial capacity investment at date 0 to $I_0 = 1$ and to denote the growth rate

---

8Our definition of the Accounting Profit Margin differs from of the familiar textbook “Return-on-Sales” metric in that we take the numerator to be Residual Income, rather than Income.
in new investments at date $t$ by $\lambda_t$. For a given history of growth rates, $\lambda = (\lambda_1, \ldots, \lambda_T)$, the overall productive capacity at date $T$ becomes:

$$K_T(\lambda) = x_T + x_{T-1} \cdot (1 + \lambda_1) + \ldots + x_2 \cdot \prod_{i=1}^{T-2} (1 + \lambda_i) + x_1 \cdot \prod_{i=1}^{T-1} (1 + \lambda_i). \quad (1)$$

To illustrate some of our results, it will be useful to consider particular decay patterns. Specifically, we will consider scenarios where capacity declines linearly, that is, $x_t = 1 - \beta \cdot (t - 1)$, with $\beta = 0$ reflecting the common “one hoss shay” scenario. Alternatively, a pattern of geometric decline would set $x_t = x^{t-1}$ for some $x \leq 1$.

For the model setup described here, the firm only incurs capacity related costs in producing its output. However, this specification is entirely one of model parsimony. None of our results derived below would be affected by the presence of variable costs incurred on a cash basis. A maintained hypothesis in our analysis is that the accounting rules do not change over time as the firm undertakes new capacity investments. The depreciation schedule for any capacity related expenditure will be denoted by $d = (d_0, d_1, \ldots, d_T)$. A positive $d_0$ corresponds to partial direct expensing, while a negative value for $d_0$ can be interpreted as an initial write-up of the asset.\footnote{The superscript “o” refers to the accounting variables for a particular investment decision rather than the aggregate variables reflecting the entire relevant history.}

Since an investment in one unit of capacity requires a cash outlay of $v$ dollars and the initial investment is normalized to one unit of capacity, the initial asset value is given by:

$$BV_0^o = v \cdot (1 - d_0^o).$$

while the depreciation charge in period $t$ is:

$$d_t^o \cdot BV_0^o,$$

with $\sum_{t=1}^{T} d_t^o = 1$. As a consequence, $BV_t^o = BV_0^o \cdot (1 - \sum_{i=1}^{t} d_i^o)$. The firm’s aggregate book value at the beginning of period $T$ (i.e., at date $T - 1$) is given by:

$$BV_{T-1}(\lambda, d) = BV_{T-1}^o + BV_{T-2}^o \cdot (1 + \lambda_1) + \ldots + BV_0^o \cdot \prod_{i=1}^{T-1} (1 + \lambda_i).$$

For a unit of capacity acquired at date 0, the sum of depreciation and interest charges at date $t$ is:

$$z_t^o \equiv d_t^o \cdot BV_0^o + r \cdot BV_{t-1}^o \quad (2)$$
where $z^o_t \equiv d^o_t \cdot v$ and $r$ denotes the exogenously given cost of capital. We refer to $z^o_t$ as the historic cost charge in period $t$ for a unit of capacity acquired at date 0. The aggregate of all depreciation and interest charges corresponding to past investments will be called the historic cost at date $T$. Denoting this cost by $H_T(\lambda, d)$, we obtain:

$$H_T(\lambda, d) \equiv z^o_T(d) + z^o_{T-1}(d) \cdot (1 + \lambda_1) + \ldots + z^o_t(d) \cdot \prod_{i=1}^{T-1} (1 + \lambda_i) + z^o_0(d) \cdot \prod_{i=1}^T (1 + \lambda_i).$$  (3)

The first term in the expression for the aggregate historic cost at date $T$ reflects that the initial investment at date 0 was for one unit of capacity and the remaining historic cost charge for this investment is $z^o_T(d)$. Conversely, the last term $z^o_0(d) \cdot \prod_{i=1}^T (1 + \lambda_i)$ reflects the direct expense charge for the expenditure made at the current date $T$.

Rogerson (2007) considers a depreciation schedule that has a particularly transparent interpretation in the current framework of capacity investments: the cash expenditure for a capacity investment is apportioned over the next $T$ periods so that the historic cost charges for a particular period are proportional to the current capacity relative to the total discounted capacity levels created by the investment. Formally, $z^o_0 = 0$ and:

$$z^o_t = v \cdot \frac{x_t}{\sum_{i=1}^T x_i \cdot \gamma^i},$$  (4)

for $1 \leq t \leq T$, where $\gamma \equiv \frac{1}{1+r}$ and $r$ denotes the firm’s discount rate. The historic cost charges in (4) will be referred to as the Relative Practical Capacity (R.P.C.) rule.\footnote{In Rogerson (2007), the cost of new investments may decrease geometrically over time. As this decline in input prices is anticipated, it impacts the choice of the current depreciation schedule. Accordingly, Rogerson coins the term Relative Replacement Cost rule. This method reduces to the R.P.C. rule if the acquisition cost of new investments remains constant over time.}

Earlier literature has shown that there is a one-to-one relation between depreciation schedules $d = (d^o_0, d^o_1, \ldots, d^o_T)$ and the historic cost charges $(z^o_0, z^o_1, \ldots z^o_T)$ under residual income (Rogerson, 1997). Formally, the linear mapping defined by (2) is a one-to-one mapping: for any vector of cost charges $(z^o_0, z^o_1, \ldots, z^o_T)$, there exists a unique depreciation schedule $d$ such that (2) is satisfied. For future reference, we will denote by $d^*$ the depreciation schedule corresponding to the R.P.C. cost allocation rule. It is readily verified that if there is no decay in practical capacity, i.e., $x_t = 1$, the depreciation schedule corresponding to the R.P.C. rule is simply the Annuity Depreciation method. On the other hand, the R.P.C. rule coincides with straight-line depreciation if practical capacity declines linearly over time such that $x_t = 1 - \frac{r}{1+r \cdot T} \cdot (t - 1)$.\footnote{In Rogerson (2007), the cost of new investments may decrease geometrically over time. As this decline in input prices is anticipated, it impacts the choice of the current depreciation schedule. Accordingly, Rogerson coins the term Relative Replacement Cost rule. This method reduces to the R.P.C. rule if the acquisition cost of new investments remains constant over time.}
The Relative Practical Capacity cost allocation rule is a close variant of the so-called *Relative Benefit Depreciation* rule which has played a central role in earlier studies on managerial incentives for investment decisions.\(^{11}\) Beginning with the pioneering work of Solomons (1965), subsequent studies have studied a variety of incentive problems in which a manager is given authority to make an investment decision because s/he is better informed about the proposed project.

For a single investment decision, the initial expenditure is inherently a joint cost for the additional capacity created over the next \(T\) periods. Yet, for a sequence of overlapping capacity investments, it becomes possible to characterize the marginal cost of another unit of capacity for *one* period of time. By definition, this cost must depend not only on past investments but also on the entire stream of anticipated future capacity levels. Let \(K_T\) denote a sequence of targeted capacity levels \((K_T, K_{T+1}, \ldots)\). Holding these targets fixed, an additional unit of capacity to be available at date \(T\) will cause a cash outflow of \(v\) at date \(T - 1\) and at the same time result in a “re-shuffling” of future investment decisions required to sustain the targeted capacity levels \(K_T\).\(^{12}\)

While it is generally difficult to calculate the long-run marginal cost of capacity explicitly, this function exhibits considerable separability when the sequence of future targeted capacity levels never results in a situation in which the firm expects to have excess capacity. The following condition dates back to earlier work by Arrow (1964). Rogerson (2005) refers to it as the “fully utilized investment” or FUI condition.

**Definition 1** For a history of investment decisions \((I_0, I_1, \ldots, I_{T-2})\), the future capacity levels \(K_T\) satisfy the No Excess Capacity (NEC) Condition provided (i) \(x_2 \cdot I_{T-2} + x_3 \cdot I_{T-3} + \ldots + x_T \cdot I_0 \leq K_T\) and (ii) \(K_{t+1} \geq K_t\), for all \(t \geq T\).

As the name suggests, the NEC condition requires the sequence of targeted future capacity levels, \(K_T\), to have the property that the firm never ends up with excess capacity, given the acquisitions it had to make in order to obtain the earlier targeted capacity levels. The present value of future cash flows required to obtain \(K_T\), given past investments, defines the *long-run* cost of the targeted capacity levels \(K_T\) at date \(T - 1\). Formally, we have:

\(^{11}\)See, for instance, Rogerson (1997), Reichelstein (1997), Baldenius and Ziv (2003), Bareket and Mohnen (2007) and Pfeiffer and Schneider (2007), among others.

\(^{12}\)To illustrate this idea in the “one hoss shay” scenario \((x_t = 1)\), an additional unit of capacity at date \(T\) requires an expenditure of \(v\) dollars at date \(T - 1\). However, the firm can reduce its investment by \(v\) dollars at date \(T\). This saving in turn is temporary since an additional \(v\) will be spent at date \(2 \cdot T - 1\), as the additional capacity acquired at date \(T\) becomes obsolete at date \(2 \cdot T\), and so forth.
subject to the constraint that the sequence \((I_{T-1}, I_T, \ldots)\) generates at least the capacity levels prescribed by the sequence \(K_T\), given the history \((I_0, \ldots, I_{T-2})\).

**Lemma 1** Given the No Excess Capacity condition, there exist coefficients \(\{u_t\}_{t=0}^{T-2}\) such that the long-run cost \(C_{T-1}(K_T|I_0, \ldots, I_{T-2})\) takes the form:

\[
C_{T-1}(K_T|I_0, \ldots, I_{T-2}) = \sum_{t=0}^{T-2} u_t \cdot I_t + \sum_{t=T}^{\infty} c \cdot K_t \cdot \gamma^{t-T+1}
\]

with

\[
c \equiv \frac{v}{\sum_{i=1}^{T} x_i \cdot \gamma^i}.
\]

This result shows that the long-run cost of capacity is additively separable in the desired capacity levels. Furthermore, the marginal cost at date \(T - 1\) of another unit of capacity for date \(T\) is given by the constant \(\gamma \cdot c\). Thus, \((1 + r) \cdot \gamma \cdot c = c\) becomes the marginal cost of another unit of capacity made available at date \(T\). This cost is decreasing in \(T\), but increasing in \(r\). As one would expect, \(\gamma \cdot c < v\), that is, the date \(T - 1\) marginal cost of capacity is less than the joint cost of one unit of capacity available for the next \(T\) periods.

A central insight in the papers of Rogerson (2005, 2007) is that the R.P.C. rule aligns marginal and average historic cost. Formally, the average historic cost of capacity is the aggregate historical cost, \(H_T(\lambda, d)\) divided by the total practical capacity at date \(T\):

\[
AHC_T(\lambda, d) = \frac{H_T(\lambda, d)}{K_T(\lambda)}.
\]
Proposition 0: Suppose depreciation is calculated according to the Relative Practical Capacity depreciation rule, \(d^*\). Then

\[
AHC_T(\lambda, d^*) = c,
\]

irrespective of the past growth levels \(\lambda\).

The algebra underlying this result simply exploits that, provided new investments are amortized according to the R.P.C. rule and therefore \(z_0^\beta(d)\) conforms to (4), the aggregate historic cost in (3) can be written as:

\[
H_T(\lambda, d) = \frac{v \cdot x_T}{\sum_{i=1}^{T} x_i \cdot \gamma^i} + \frac{v \cdot x_{T-1}}{\sum_{i=1}^{T} x_i \cdot \gamma^i} \cdot (1 + \lambda_1) + \ldots + \frac{v}{\sum_{i=1}^{T} x_i \cdot \gamma^i} \cdot \prod_{i=1}^{T-1} (1 + \lambda_i)
\]

\[
= \frac{v}{\sum_{i=1}^{T} x_i \cdot \gamma^i} \left[ x_T + x_{T-1} \cdot (1 + \lambda_1) + \ldots + \prod_{i=1}^{T-1} (1 + \lambda_i) \right]
\]

\[
= \frac{v}{\sum_{i=1}^{T} x_i \cdot \gamma^i} \cdot K_T(\lambda)
\]

\[
= c \cdot K_T(\lambda).
\]

We conclude that even though the aggregate historic cost \(H_T(\lambda, d)\) indeed represents costs that are sunk at date \(T\), it is nonetheless true that the date \(T\) marginal cost of capacity is captured precisely by the average historical cost, \(AHC_T(\lambda, d)\), provided past investments have consistently been depreciated according to the Relative Practical Capacity rule. As pointed out in Rogerson (2005), this insight has immediate implications for the performance of rate-of-return regulation. Suppose regulators set a firm’s product price so as to give the firm a return on its invested capital (book value) equal to its cost of capital, \(r\). Equivalently, the product price is set so that the firm’s residual income is zero in each period, that is, revenue is set equal to historic cost (plus any applicable variable operating expenses). One implication of the benchmark result in Proposition 0 is that if investments are consistently depreciated according to the R.P.C. rule, this regulatory scheme will indeed result in marginal cost pricing. Thus in a setting with an overlapping sequence of investments proper accrual accounting makes it possible to satisfy two simultaneous goals: (i) marginal cost pricing and (ii) allowing the regulated firm to break even.\(^{16}\)

\(^{16}\)Rogerson (2007) considers an incentive problem in which a principal delegates authority for capacity investments to an agent who has superior information about expected future revenues attainable from the resulting capacity stocks. If the cash expenditures required for new investments are allocated over time
3 The Behavior of Historic Cost

This section examines how the average historic cost relates to marginal cost when the depreciation schedule differs from the Relative Practical Capacity rule. Specifically, which depreciation schedules will cause the average historic cost to fall below (exceed) the marginal cost, and what factors drive the difference between the two cost measures? Under what conditions is the average historic cost a reasonably good approximation for the marginal cost of capacity?

To address these questions, we first introduce a partial ordering on the class of depreciation schedules. This condition is stated in terms of the book values associated with an investment in one unit of capacity at date 0. We recall that any depreciation schedule \(d = (d_0, d_1, \ldots, d_T)\) defines a sequence of book values according to:

\[ BV_0^a(d) = v \cdot (1 - d_0) \]

and

\[ BV_t^a(d) = BV_0^a(d) \cdot (1 - \sum_{i=1}^{t} d_i) \].

**Definition 2** A depreciation schedule \(d\) is (weakly) more accelerated than \(d'\) if

\[ BV_t^a(d) \leq BV_t^a(d') \]

for all \(0 \leq t \leq T\).\(^{17}\)

The criterion of accelerated depreciation is a cumulative (second-order) condition pertaining to the total amount depreciated in the past. It will be useful to recast this criterion in terms of discounted historic cost charges.

**Lemma 2** The depreciation schedule \(d\) is (weakly) more accelerated than \(d'\) if and only if

\[ \sum_{i=0}^{t} z_i^0(d) \cdot \gamma^i \geq \sum_{i=0}^{t} z_i^0(d') \cdot \gamma^i \]

for all \(0 \leq t \leq T - 1\).

The inequality in Lemma 2 is a direct consequence of the fundamental accounting identity relating discounted future cash flows to the sum of current book value and future according to the R.P.C. rule, the manager will have a robust incentive to implement the first-best investment levels, that is, the manager’s incentive holds irrespective of his planning horizon and irrespective of the compensation scheme in place.

\(^{17}\)Of course, the restriction to tidy depreciation schedules implies that \(BV_T^a(d') = BV_T^a(d) = 0\).
residual incomes.\footnote{This identity, which dates back to Preinreich (1938), has been the basis of so-called Residual Income Valuation models; see, for instance, Feltham and Ohlson (1996) or Penman (2006).} For any (and as of now unexplained) sequence of operating cash flows \((CF_1^o, \ldots, CF_T^o)\), we take residual income to be the difference between cash flow and historic cost, i.e., \(RI_t^o \equiv CF_t^o - z_t^o(d)\). The fundamental identity relating discounted future cash flows to the stream of future residual incomes says that independent of the accounting rules:

\[
BV_t^o(d) + \sum_{i=t+1}^{T} RI_i^o \cdot \gamma^{i-t} = \sum_{i=t+1}^{T} CF_i^o \cdot \gamma^{i-t}.
\] (9)

Substituting \(RI_t^o = CF_t^o - z_t^o(d)\) on the left-hand side of this equation yields:

\[
BV_t^o(d) = \sum_{i=t+1}^{T} z_i^o(d) \cdot \gamma^{i-t}.
\] (10)

The claim in Lemma 2 now follows because for any two depreciation schedules the total discounted historic cost charges must be equal to the initial investment expenditure, that is, \(\sum_{i=0}^{T} [z_i^o(d) - z_i^o(d')] \gamma^i = 0\).

We are now in a position to examine how growth and the choice of accounting rules jointly impact the aggregate historic cost. The following result shows that the impact of more accelerated depreciation depends crucially on the magnitude of past growth rates.

**Proposition 1** If the depreciation schedule \(d\) is more accelerated than \(d'\), then

\[
H_T(\lambda, d) \leq H_T(\lambda, d')
\]

provided \(\lambda_t \leq r\) for all \(1 \leq t \leq T\). The reverse inequality holds whenever \(\lambda_t \geq r\) for all \(1 \leq t \leq T\).

According to Lemma 2, more accelerated depreciation results in front-loaded historic cost charges for an individual investment project. However, with a sequence of \(T\) investments, more accelerated depreciation will lower the aggregate historic cost provided the past rates of growth have been moderate, i.e., \(\lambda_t \leq r\). This finding is most easily seen in the special case of zero growth: the aggregate depreciation charge across all projects is unaffected by a more accelerated schedule because the initial expenditure is fully amortized under any depreciation schedule. Yet book values and the corresponding interest charges are uniformly lower. On the other hand, if all past growth rates \(\lambda_t\) exceed the cost of capital \(r\), then
the relatively large historic cost charges for the latest (and largest!) investments tend to
dominate and result in a higher aggregate historic cost.19

Proposition 2 Suppose depreciation is more accelerated than the R.P.C. depreciation rule,
\(d^*\). Then,

\[
AHC_T(\lambda, d) \begin{cases}
\leq c & \text{if } \lambda_t \leq r \text{ for all } 1 \leq t \leq T \\
\geq c & \text{if } \lambda_t \geq r \text{ for all } 1 \leq t \leq T.
\end{cases}
\]  \hspace{1cm} (11)

Conversely, when the depreciation schedule is more decelerated than the R.P.C. depreciation
rule, \(AHC_T(\lambda, d) \geq (\leq) c \text{ if } \lambda_t \leq (\geq) r \text{ for all } 1 \leq t \leq T.\)

The location of the \(AHC_T(\cdot, d)\) function among the four possible “quadrants” depending
on whether (i) depreciation is accelerated or decelerated relative to R.P.C. rule and (ii) the
rate of growth is moderate (\(\lambda_t \leq r\)), or aggressive, i.e., \(\lambda_t \geq r\). For moderate growth rates,
an accelerated depreciation schedule tends to understate the average historic cost relative
to the long-run marginal cost, while the opposite is true for growth rates that consistently
exceed the firm’s cost of capital.

Figure 1 illustrates the above finding for a constant growth scenario. The unbroken curve
represents the average historic cost as a function of growth under a baseline depreciation
schedule, \(d\). According to Proposition 2, this schedule must be accelerated relative to the
R.P.C. rule. In contrast, the dashed curve in Figure 1 depicts the impact of an even more
accelerated depreciation scheme, \(d'\). The center point of the rotations is the Pivot Point
\((\lambda, AHC_T) = (r, c)\) through which the functions \(AHC_T(\cdot, d)\) must always pass, regardless of
the choice of depreciation schedule. More accelerated depreciation schedules thus have the
effect of a counter-clockwise rotation of the \(AHC_T(\lambda, \cdot)\) function.

Finally, it should not come as a surprise that the choice of depreciation schedule is
irrelevant for the magnitude of the average historic cost when the growth rate happens to
equal \(r\) in each period. The fundamental identity:

\[
\sum_{i=0}^{T} z^0_i(d) \cdot \gamma^i = v
\]

\(\text{This finding generates a prediction for comparing the residual income of two hypothetical firms that}
\text{obtain identical cash flows but differ with regard to the accounting rules used, possibly because one firm}
\text{directly expenses a larger share of its new investments. We predict that more accelerated depreciation will}
\text{result in higher (lower) residual income at any point in time provided past investments have been growing}
\text{consistently at a rate less (greater) than the cost of capital, } r. \text{ To calibrate this observation, for the sample}
\text{studied in Rajan et al. (2007), the median past growth rate for new investments was close to 13 percent, a}
\text{number that is within the range of “normal” cost of capital estimates.}\)
implies that the aggregate historical cost $H_T(\lambda, d) \equiv \gamma^{-T} \cdot v$ whenever $\lambda = (r, ..., r)$.

The “quadrant” result in Proposition 2 exhibits some analogies to a result relating Return-on-Investment (ROI) to the internal rate of return of a firm’s investment projects. These characterizations are based on models in which a firm conducts the same “representative” project in each period, though the scale of these projects may increase due to growth. ROI will be equal to the internal rate of return of the representative project (IRR) if either the growth rate is identically equal to the IRR, or alternatively, the depreciation schedule is neutral. For a depreciation schedule that is accelerated relative to the neutral schedule, ROI will exceed (fall below) the internal rate of return if growth has been consistently below (above) the IRR.

Neutral depreciation and the relative practical capacity rule are both forms of unbiased accounting in the sense that one equates ROI with the underlying internal rate of return,

\[ \frac{Inc_t^i}{BV_{t-1}} \]

must be constant over time.

\[ \text{Stauffer (1971) and Beaver and Dukes (1974) introduced the notion of neutral accounting by the criterion that the ratios:} \]

\[ \text{Variants of this result have been obtained in different frameworks, and at different levels of generality, by Fisher and McGowan (1983), Stark (1994), Danielson and Press (2003), Gjesdal (2004) and Rajan et al. (2007).} \]
while the other equates average historic with marginal cost. The two concepts are not the same, however, since the criterion of neutrality is calibrated to the stream of cash flows derived from the representative project and its IRR. In the special case where all cash flows from the representative project are identical, both methods reduce to annuity depreciation, that is, \( d^*_t = (1 + \bar{r})^{t-1} \cdot d^*_1 \), where \( \bar{r} \) is given either by \( r \) or IRR. If the representative project has a positive NPV and thus \( IRR > r \), these schedules cross only once and neutral depreciation results in lower depreciation charges early on. We conclude that in this special case the Relative Practical Capacity rule is more accelerated than neutral accounting by the criterion given in Definition 1.\(^{22}\)

Hotelling (1925) suggested that depreciation in each period be calculated so that book value is equal to intrinsic value, i.e., the present value of future cash flows from the representative project. In the industrial organization literature, this method is referred to as “economic depreciation”; see, for instance, Fisher and McGowan (1983) and Fisher (1987). It is important to recall, however, that Hotelling confined attention to zero NPV projects. Since the Relative Practical Capacity rule coincides with neutral depreciation in the special case of zero NPV projects, both methods can be seen as extensions of Hotelling’s (1925) concept of “economic depreciation.”\(^{23}\)

The quadrant result in Proposition 2 suggests that the average historic cost is increasing in growth provided depreciation is more accelerated than the R.P.C. rule. In the special case of constant growth, this must be true at least in a neighborhood of the Pivot Point, \((\lambda, AHC_T) = (r, c)\), i.e., for growth rates close to \( r \). To state a general result about the monotonicity of \( AHC_T(\lambda, d) \), we introduce the following stronger notion of accelerated depreciation.

**Definition 4** A depreciation schedule \( d \) is said to be uniformly more accelerated than the R.P.C. depreciation rule if:

\[
\begin{align*}
    z^0_0(d) &\geq 0 \\
    \frac{z^0_t(d)}{x_t} &\text{ is monotonically decreasing in } t.
\end{align*}
\]

For the R.P.C. depreciation schedule, \( d^* \), the inequalities in Definition 4 are met as equalities, i.e., \( z^0_0 = 0 \) and \( x_{t+1} \cdot z^0_t(d^*) = x_t \cdot z^0_{t+1}(d^*) \). It is also straightforward to verify that the criterion of uniformly accelerated depreciation is indeed stronger than of accelerated

\(^{22}\)It can be shown via example that this ordering does not hold in general, that is, neutral accounting may be more accelerated than the R.P.C. rule.

\(^{23}\)Analytical work on equity valuation has focused on the effects of conservative accounting as defined by the criterion that at each point in time intrinsic value exceeds current book value, see, for example, Feltham and Ohlson (1996), Ohlson and Zhang (1998), Zhang (2000) and Ohlson and Gao (2006). Some authors, including Beaver and Ryan (2005), refer to this criterion as unconditional conservatism, to distinguish it from the conditional conservatism emphasized by Basu (1997), Watts (2003), Ball and Kothari (2007) and others.
depreciation: if \( d \) satisfies the inequalities in Definition 4, it will also satisfy the “cumulative” (second-order) inequalities in Definition 2. Finally, a depreciation schedule is said to exhibit uniform deceleration, relative to R.P.C., if it satisfies the reverse of the inequalities in Definition 4. With these definitions in place, we can now address the monotonicity of \( AHC_T(\lambda, d) \) in past growth rates.

**Proposition 3** The average historic cost, \( AHC_T(\lambda, d) \), is everywhere increasing (decreasing) in each \( \lambda_t \) if and only if \( d \) is uniformly more accelerated (decelerated) than the R.P.C. depreciation rule \( d^* \).

Proposition 2 showed that if past growth rates are below the cost of capital and the chosen depreciation schedule is uniformly accelerated relative to the R.P.C. rule, then the firm will find itself in the South-West quadrant of Figure 1. From this starting point, a move towards either a less accelerated schedule or to higher levels of growth (Propositions 1 and 3) will shift the average historic cost towards the long-run marginal cost. To examine how well historic cost approximates marginal cost, the remainder of this section focuses on accounting rules that are most common for external financial reporting purposes. Specifically, an \( \alpha \)-fraction of new investment is expensed immediately, while the remaining portion, \((1 - \alpha)\), is depreciated straight line over the \( T \)-year useful life of the investment. This specification is consistent with the observation that under current financial accounting rules many intangible assets, such as expenditures for training and Research & Development are directly expensed. For further simplification, we assume that the capacity derived from new investments declines linearly over time, i.e., \( x_t = 1 - \beta \cdot (t - 1) \). The slope parameter \( \beta > 0 \) then captures the annual decline in practical capacity. As observed earlier, straight line depreciation corresponds to the R.P.C. depreciation rule for such a linear decay pattern precisely when \( \beta \) equals \( \beta^* \equiv \frac{r}{1+r\cdot T} \).  

**Lemma 3** Partial direct expensing (\( \alpha \geq 0 \)) combined with straight line depreciation is uniformly more accelerated than the the R.P.C. depreciation rule if and only if \( \beta \leq \beta^* \).

For the remainder of this section, we restrict attention to settings where past growth in new investments has been constant and equal to \( \lambda \) over the past \( T \) years. Given straight line depreciation, we seek to further characterize the behavior of the average cost function. Proposition 3 together with Lemma 3 show that the average historic cost is an increasing function of growth, \( \lambda \), provided \( \beta < \beta^* \). Furthermore, higher values of \( \alpha \) result in depreciation schedules that are more accelerated. We refer to greater proportions of directly expensed

---

\(^{24}\)Friedl (2007) also makes this observation in connection with product pricing for regulated utilities.
investments, i.e., higher values of $\alpha$, as more conservative accounting. Growth and conservatism are said to be complements for the average historic cost if the cross partial derivative of $AHC_T(\lambda, \alpha)$ in $\alpha$ and $\lambda$ is positive.

**Proposition 4** Suppose that $\beta < \beta^*$ and past growth, $\lambda$, has been constant and non-negative. Then conservatism and growth are complements everywhere for the average historic cost, $AHC_T(\lambda, \alpha)$, if and only if $r < \frac{6}{T-1}$.

Thus higher levels of growth push up the average historic cost and do so at faster rates when the accounting is more conservative. The necessary and sufficient stipulation that the cost of capital not be too high is quite permissive. Even if the useful life of assets is taken to be a relatively high 25 years, the cost of capital must only satisfy the constraint of $r \leq 25\%$.\(^\text{25}\)

We finally investigate the magnitude of the error that arises from using average historic cost as a proxy for the marginal cost. In the current parametric setting this error depends on the five remaining variables: $(\lambda, \alpha, \beta, r, T)$. We seek to characterize the percentage error:

$$\Delta_1(\lambda, \alpha, \beta, r, T) = \frac{|AHC_T(\lambda, \alpha, \beta, r, T) - c(\beta, r, T)|}{AHC_T(\lambda, \alpha, \beta, r, T)}$$

The above findings show that $\Delta_1 = 0$ if either $\lambda = r$ or the accounting corresponds to the R.P.C. rule, that is, $\beta = \beta^*$ and $\alpha = 0$. Holding the remaining parameters $(\alpha, r, T)$ fixed, Figure 2 displays the level sets (isoquants) of the function $\Delta_1$ at the 1%, 3% and 5% levels for different combinations of growth, $\lambda$ and physical decay $\beta$. For these parameter values, the average historic cost provides a remarkably good approximation to the marginal cost. One would need to have rather high growth rates and/or fairly extreme decay factors in order for the average historic cost to misstate the marginal cost by more than 5%. Figure 2 suggests that the error function is a “star-shaped bowl” with its bottom at the point $(\lambda, \beta) = (r, \beta^*)$. In conjunction with Lemma 3, Proposition 3 implies that $\Delta_1(\cdot)$ is decreasing (increasing) in $\lambda$ provided $\lambda < (>) r$. Similarly, Figure 2 also suggests a monotonicity result with regard to the productivity parameter $\beta$.

**Proposition 5** The error function $\Delta_1(\cdot)$ is decreasing in $\beta$ provided $\beta < \beta^*$ and increasing in $\beta$ provided $\beta > \beta^*$.

\(^{25}\)As shown in the proof of Proposition 4, the number 6 in the sufficient condition arises from the formula for summing the squares of the first $T$ integers, that is, $\sum_{t=1}^{T} t^2 = \frac{1}{6} \cdot T \cdot (T+1) \cdot (2T+1)$. We also note that the result can be extended to negative levels of growth by appropriately tightening the upper bound on $r$.\[^{17}\]
This result is intuitively plausible because a capacity decay parameter closer to $\beta^*$ makes straight-line depreciation a better approximation to the R.P.C. rule, which would result in no error. Of course, this intuition is subject to the qualification that the decay parameter affects both the average historic- and the marginal cost, i.e., both the numerator and the denominator in the ratio of $\Delta_1 = \left| 1 - \frac{c}{AHC_T} \right|$. The proof of Proposition 5 in the Appendix provides a closed form expression for the ratio of the marginal- and the average historic cost. This expression shows, for example, that for the South-West corner point in Figure 2, corresponding to no growth and one hoss shay, the percentage error is given by:

$$\Delta_1(0, 0, 0, r, T) = \frac{2 \cdot T \cdot h(r)}{2 + r(1 + T)} - 1,$$

where $h(r)$ is the dollar amount which, when paid as an annuity over $T$ periods, has a present value of one dollar, given the discount rate $r$.\footnote{Thus, $h(r) = r \cdot \left[ 1 - (1 + r)^{-T} \right]^{-1}$ for $r \geq -1$. It can be verified that $h(\cdot)$ is increasing and convex with $h(0) = \frac{1}{T}$ and $\lim_{r \to -1} h(r) = 0$.} Thus if $r = .12$ and $T = 10$, as in Figure 2, the error in the South-West corner amounts to 0.066 or 6.6%
Figure 3: Minimum growth rate in investment needed for straight line depreciation to achieve distortion levels below 5% and 10% for various combinations of useful life of assets and cost of capital. For the sake of realism, the minimum growth rate is capped at an annual decline of 32%.
The charts in Figure 3 provide a rather comprehensive image of the error function \( \Delta_1(\cdot) \) when there is no direct expensing \((\alpha = 0)\). Each chart depicts the minimum growth rate required to keep the percentage error \( \Delta_1(\cdot) \) either below 5% or 10% for different combinations of the remaining parameters. The maximum growth rate in each chart is the cost of capital since that would result always in a zero error. The line marked at the bottom of each chart, marked A, sets the decay rate, \( \beta \), equal to \( \beta^* \) implying that \( \Delta_1(\cdot) = 0 \) at any growth rate. Curve B represents \( \beta = \frac{1}{2} \cdot \beta^* \), while curve C represents the one-hoss shay scenario \((\beta = 0)\). Collectively, these charts lend strong support to the earlier impression in Figure 2 that for a wide range of “reasonable” parameter values the average historic cost is a remarkably good proxy for the underlying marginal cost of capacity.27

To conclude the analysis of the error function \( \Delta_1(\cdot) \), we turn to the impact of more conservative accounting in the form of higher values of \( \alpha \). Using the closed-form expression for \( \Delta_1(\cdot) \) from the proof of Proposition 5, we obtain, for \( \beta = 0 \) and \( \lambda < r \):

\[
\Delta_1(\lambda, \alpha, 0, r, T) = \frac{\lambda \cdot T \cdot h(r)}{\alpha \cdot \lambda \cdot T \cdot h(\lambda) + (1 - \alpha) \cdot (\lambda - r + r \cdot T \cdot h(\lambda))} - 1. \tag{13}
\]

Plugging in values of \( \lambda = 0 \), \( r = 12\% \) and \( T = 10 \), we obtain, consistent with Figure 3, a percentage error of 6.62% if all investment expenditures are fully capitalized. Yet, as \( \alpha \) increases the percentage error \( \Delta_1(0, \alpha, 0, .12, 10) \) escalates fairly rapidly. To witness, if 20% of investment expenditures are expensed immediately, then, for the same set of parameters, the percentage error \( \Delta_1(\cdot) \) rises to 15.83%. For an \( \alpha \) of 0.4, \( \Delta_1(\cdot) \) increases steeply to 26.78%.

Our conclusion here can be contrasted to Lev (2001) who deplores the accounting practice of expensing expenditures for intangible assets, yet asserts that this practice is just as ad-hoc and arbitrary as the practice of using straight line depreciation for tangible assets.28

For the error metric analyzed in this section, we conclude that direct expensing of intangible investments leads generally to more pronounced distortions than the seemingly ad-hoc practice of straight line depreciation.

---

27Figure 3 suggests that the error \( \Delta_1(\cdot) \) is increasing in the cost of capital, \( r \). This can be shown to be true provided \( \lambda < r \) and \( \beta < \beta^* \). For all charts in Figure 3, the error is increasing in the useful life of the investment, \( T \). This property turns out not to be true globally if one pushes \( T \) to very large values.

28Specifically, Lev (2001, p. 18) states: “After all, the accounting for physical assets in financial statements is as deficient as the accounting for intangibles. True, physical assets are capitalized, but they are recorded at historical cost and depreciated by ad hoc, unrealistic schemes (for example ten-year straight-line depreciation). What economic inferences about value and performance of physical performance of physical assets can be drawn from their balance sheet values? Essentially none.”
4 Accounting Profit Margins

This section broadens the scope of our model by making growth an endogenous variable and introducing cash revenues into the model. In particular, we examine the behavior of accounting profit margins for a firm that makes a sequence of overlapping capacity investments. Suppose that net-revenues attainable from current capacity $K_T$ are given by some revenue function $R_T(\cdot)$, such that $R_T(\cdot)$ is monotonically increasing in $K_T$. The average sales price, $P_T(K_T)$ is then given by total net-revenue per unit of capacity:

$$P_T(K_T) = \frac{R_T(K_T)}{K_T}$$

Most textbooks on financial statement analysis define the Accounting Profit Margin, APM, or “Return-on-Sales”, as the ratio of income to sales. Consistent with the above measure of historic cost, our definition of the APM will include imputed interest charges in the definition of income; that is, we replace Income by Residual Income in the numerator. Formally, we examine the following measure of the accounting profit margin at date $T$:

$$APM_T(\lambda, d) = \frac{Inc_T(\lambda, d) - r \cdot BV_{T-1}(\lambda, d)}{R_T(K_T(\lambda))},$$ (14)

where Income is simply the difference between net-revenues and all depreciation charges:

$$Inc_T(\lambda, d) = R_T(K_T(\lambda)) - [d_{T}^o \cdot BV_0^o + d_{T-1}^o \cdot BV_0^o \cdot (1 + \lambda_1) + \ldots$$

$$+ d_{1}^o \cdot BV_0^o \cdot \prod_{i=1}^{T-1}(1 + \lambda_i) + d_{0}^o \cdot v \cdot \prod_{i=1}^{T}(1 + \lambda_i)].$$

As before, it is simply for parsimony that we set the variable costs of production equal to zero.\textsuperscript{29} Since the definition of historic cost in (3) includes both depreciation and interest charges, the APM at date $T$ can be expressed as:

$$APM_T(\lambda, d) = \frac{R_T(K_T(\lambda)) - H_T(\lambda, d)}{R_T(K_T(\lambda))}.$$ (15)

There is a long tradition in the economics literature of viewing the Lange-Lerner ratio as an index of the monopoly power a firm has in a particular product market. The Lange-Lerner ratio is usually defined as the difference between the product price and the marginal cost, divided by the product price. In the context of our model, it is natural to define the Lange-Lerner ratio as:

\textsuperscript{29}For the results derived below, any positive variable costs of production would be included equally in both the historic cost and the measure of marginal cost.
where $c$ denotes the marginal cost of producing another unit, that is, of providing another unit of capacity in period $T$. Various empirical studies in industrial organization have taken accounting-based measures of operating profit margins as proxies for the Lange-Lerner ratio.\footnote{30} A scathing critique of this approach has been provided by Fisher (1987), who concludes that the APM is generally not a valid proxy for the Lange-Lerner ratio and that furthermore there is no systematic relation between these two ratios. Fisher does claim that the APM will be equal to the Lange-Lerner index if either the firm’s investments have been growing exponentially at the firm’s cost of capital $r$, or the accounting conforms to what he terms “Hotelling” valuation.

As mentioned above, Hotelling (1925) suggested that depreciation be calculated so that for an individual investment project, the intrinsic value (future discounted cash flows) be equal to the current book value at each point in time. Hotelling confined attention to zero NPV projects and this restriction is essential because with positive NPV projects it would be impossible to find a depreciation schedule that equates book values consistently with intrinsic values. It should also be noted that, in our framework of overlapping capacity investments, there is no notion of project cash flows associated with an individual investment. Cash flows derive from the currently available capacity which in turn reflects the entire history of past investments.\footnote{31}

Dividing the Accounting Profit Margin in (15) by total capacity available, it becomes clear that the relation between the Lange-Lerner ratio and the Accounting Profit Margin hinges entirely on how the historic cost of capacity, $H_T(\lambda, d)$ relates to the long-run marginal cost of capacity, $c$. In particular, we obtain the following corollary to Proposition 2.

**Corollary 1** Given a depreciation schedule that is more accelerated than the R.P.C. depreciation rule, the Accounting Profit Margin in (15) exceeds (falls below) the Lange-Lerner ratio:

$$LL_T(\lambda) = \frac{P_T(K_T(\lambda)) - c}{P_T(K_T(\lambda))}$$

\footnote{30}See, for instance, Long et al. (1982), Domowitz et al. (1986) and Martin (2002).

\footnote{31}One way of extending Hotelling’s concept of economic depreciation to settings with positive NPV projects is to calculate depreciation in a manner that may be termed *fair value accounting*. Accordingly, the project’s net present value is recognized upfront by an asset write-up, i.e., a negative depreciation charge (Bierman, 1961; Bodenhorn, 1961). However, even if one were to modify our model so that each capacity investment had its own separate stream of cash revenues (possibly because $R_T(\cdot)$ is linear in $K_T$), this extended form of Hotelling valuation would still not coincide with the R.P.C. rule.
provided $\lambda_t \leq r \,(\lambda_t \geq r)$ for all $1 \leq t \leq T$. These inequalities reverse if the depreciation schedule is more decelerated than the R.P.C. rule.

It is natural explore the relation between the Accounting Profit Margin and the Lange-Lerner index in the context of a profit maximizing monopolist. In particular, suppose a monopolist makes capacity investments in response to changes in the external product market. It will be convenient to specify the demand curve so that the Lange-Lerner measure of monopoly power can be separated from changes in the size of the market. In effect, this leads us to a constant elasticity demand curve of the form:

$$K_t = a_t \cdot p^{-\epsilon},$$

where $\epsilon > 1$ denotes the demand elasticity of price. It is well known from the static theory of monopoly pricing that a profit-maximizing monopolist will set the product price so the Lange-Lerner ratio is equal to $\frac{1}{\epsilon}$. Thus the Lange-Lerner index is independent of the size of the market, represented by the parameter $a_t$.

To further calibrate the model, we assume that the firm enters the product market in question at date zero (or, equivalently at any date thereafter). In each period, the size of the market grows at some rate $\mu_t$, so that:

$$a_t = a_{t-1} \cdot (1 + \mu_t),$$

for $1 \leq t \leq T$. The entire history of growth rates will be denoted by $\mu = (\mu_1, \ldots, \mu_T)$.

**Proposition 6** If the product market exhibits a constant price elasticity of demand and consistently positive growth rates $\mu_t \geq 0$, the difference

$$LL_T(\mu) - APM_T(\mu, d) = \frac{AHC_T(\mu, d) - c}{P_T(K_T(\mu))}$$

is increasing in each $\mu_t$ provided depreciation is uniformly more accelerated than the R.P.C. depreciation rule.

As noted above, both the Accounting Profit Margin and the Lange-Lerner index are driven by past investment growth which is now determined endogenously by changes in the size of the product market. For the constant price elasticity setting considered in this

---

32If one assumes that the firm was already in the product market in question before date 0, it becomes necessary to specify the capacity investments that were made prior to date 0, since the firm replaces the assets that then become obsolete sometime between date 0 and date T.
section, the monopoly price is unaffected by the changes in the size of the market and equals \( P_T(K_T(\mu)) = c \cdot \frac{\epsilon}{\epsilon - 1} \). The proof of Proposition 6 demonstrates first that the \emph{induced} growth rate in new investments \( \lambda_t(\mu_1, ..., \mu_t) \) is increasing in each \( \mu_t \). We then invoke Proposition 3 to conclude that higher growth rates in the product market tend to increase the reported historic cost, \( AHC_T(\mu, d) \).

We conclude that the APM becomes a valid proxy for the Lange-Lerner index to the extent that the average historic cost provides a good approximation of the marginal cost. Thus, our analysis at the end of the previous section applies.\(^{33}\) With reference to (18), however, we note that the difference between APM\(_T\) and LL\(_T\) is also affected by the underlying price elasticity, \( \epsilon \). For firms enjoying a higher degree of pricing power (lower values of \( \epsilon \)), the absolute value of the difference in (18) will ceteris paribus be smaller, simply because of a higher monopoly price in the denominator.

In sum, our findings here stand in contrast to those obtained by Fisher (1987). First, Fisher asserts that “Hotelling valuation” creates unbiased accounting in the sense that the Accounting Profit Margin will be equal to the Lange-Lerner monopoly index. Aside from the conceptual issue of extending Hotelling’s concept of economic depreciation to environments with positive NPV projects, we submit that for the capacity investment model studied in this paper, the Relative Practical Capacity depreciation rule is the unique method for aligning the two ratios.\(^{34}\) Second, our results in Corollary 1 and Propositions 2 and 6 identify the directional bias that results from taking the Accounting Profit margin as a proxy for the Lange-Lerner ratio, depending on the depreciation policy and the rate at which market demand grows.\(^ {35}\) Finally, the validity of the Accounting Profit Margin as a proxy for the Lange-Lerner monopoly index hinges essentially on how well the average historic cost approximates the product’s marginal cost.

\(^{33}\)Similar to our characterization of the loss function \( \Delta_1(\cdot) \), one can again plot the resulting percentage errors for alternative parameterizations, such as linear or geometric decay in productive capacity. Of course, in general constant growth in the size of the product market does not translate into constant investment growth, though this will be the case in the special case of geometric decline, i.e., \( x_t = x^{t-1} \).

\(^{34}\)While we did not formally prove uniqueness, it is relatively straightforward to show that in order for the historic cost to be equal to the marginal cost \emph{irrespective} of past growth, depreciation must conform to the R.P.C. rule.

\(^{35}\)In that sense, we subscribe to the perspective in Martin (2002, p. 172) who, in response to Fisher and McGowan’s criticism of ROI as a proxy for the IRR, states: “The basic question is not whether some or any accounting rates of return equal the internal rate of return....The basic question is how accounting rates of return should be analyzed to see whether they signal the exercise of market power. The Fisher and Fisher-McGowan arguments do not address this issue.”
5 Concluding Remarks (not quite complete ..)

Summary of paper
Implications for Rate of Return Regulation. Capacity investments are driven by demand and the break-even constraint.

The results of this paper also have implications for other questions studies in both the industrial organization and the management literature. For instance, what measures of cost should a firm’s product price cover? In connection with predatory pricing, it has become common practice to view the product price that a firm charges as lawful if that price covers the product’s average variable cost (Areeda and Turner, 1975). By implication, there is no presumption that product prices must cover capacity related fixed costs, the usual argument being that such costs are sunk at the time the firm sets the product price. The preceding analysis suggests that this standard may be too lenient.

There were no variable costs in our model (without loss of generality) and the capacity related historic cost $H_T(\cdot)$ is indeed sunk at date $T$. Yet, in making the capacity investment decision, the firm imputes the average historic cost $AHC_T(\cdot)$ as the relevant cost (provided depreciation is calculated according to the R.P.C. rule). Barring shocks to market demand, one would therefore expect a profit maximizing firm to set a price that covers the capacity related sunk costs. To be sure, as the firm takes only the variable costs into consideration at date $T$, it might at that point in time prefer a lower price in exchange for a higher level of output and sales. However, the presence of the capacity constraint makes this infeasible. Thus, one would expect a firm that seeks to maximize profit, as opposed to hurting competitors, to charge product prices that cover capacity related fixed costs, provided there were no substantial shocks to market demand in the intermittent periods.

---

36 Similar arguments apply in the inventory model of Baldenius and Reichelstein (2005), where a firm can store parts of the output produced in inventory. Managerial incentives to adopt an optimal inventory policy can be provided by charging the manager for the historic cost of the units sold from inventory, even though those costs are sunk at the date of sale.
6 Appendix

Proof of Lemma 1: By definition:

\[ C_{T-1}(K_T|I_0, ..., I_{T-2}) \equiv \min_{\{I_t\}_{t=T-1}^{\infty}} \sum_{t=T-1}^{\infty} v \cdot I_t \cdot (1 + r)^{-(t-T+1)} \]

subject to:

(i) \[ I_t \geq 0 \quad \text{for} \quad t \geq T - 1, \]

(ii) \[ x_1 \cdot I_{T-1} + x_2 \cdot I_{T-2} + ... + x_T \cdot I_0 \geq K_T \]
\[ x_1 \cdot I_T + x_2 \cdot I_{T-1} + ... + x_T \cdot I_1 \geq K_{T+1} \]
\[ \vdots \]
\[ x_1 \cdot I_{T-1+t} + x_2 \cdot I_{T+t-2} + ... + x_T \cdot I_t \geq K_{T+t} \]
\[ \vdots \]

The NEC condition ensures that any solution, \( I^*_T \), to this minimization problem is such that the inequality constraints in (ii) are met as equalities. It will be convenient to define the following infinite sequences:

\[ X = (0, x_1, ..., x_T, 0, 0...); \]
\[ I = (I_0, ..., I_{T-2}, I_{T-1}); \]
\[ K = (K_0, ..., K_{T-1}, K_T), \]

where \( I_{T-1} \equiv (I_{T-1}, I_T, ...) \), and \( K_T \equiv (K_T, K_{T+1}, ...) \). Furthermore, we define:

\[ K_0 \equiv 0, \quad K_1 \equiv x_1 \cdot I_0, \quad K_2 \equiv x_1 \cdot I_1 + x_2 \cdot I_0, ... \]
\[ K_{T-1} \equiv x_1 \cdot I_{T-2} + x_2 \cdot I_{T-1} + ... + x_T \cdot I_0. \]

While these are definitions for \( (K_0, ..., K_{T-1}) \), we note that these capacity levels are the ones that would emerge if there had not been any investments prior to date 0. The optimal \( I^*_T \) then induces a unique vector \( K \). Moreover,

\[ K = X \circ I, \]

where "\( \circ \)" stands for convolution.\(^{37}\)

\(^{37}\)The convolution of any two infinite sequences, \( X \) and \( Y \), is denoted by \( X \circ Y \). It is defined as:

\[ (X \circ Y)_0 = x_0 \cdot y_0, \quad (X \circ Y)_1 = x_0 \cdot y_1 + y_0 \cdot x_1 \]
\[ (X \circ Y)_2 = x_0 \cdot y_2 + x_1 \cdot y_1 + x_2 \cdot y_0, \ldots \text{ and so on.} \]
We denote the one-sided Z-transform of any infinite sequence $X$ by

$$Z(X | z) \equiv \sum_{t=0}^{\infty} x_t \cdot z^{-t}.$$  

It is well known that the Z-transform of any convolution has the following multiplicative property:\(^{38}\)

$$Z(X \circ Y | z) = Z(X | z) \cdot Z(Y | z).$$

By definition,

$$C_{T-1}(K_T | I_0, ..., I_{T-2}) = v \cdot (1 + r)^{T-1} \cdot [Z(I | 1 + r) - \sum_{t=0}^{T-2} I_t \cdot (1 + r)^{-t}]. \quad (19)$$

Invoking the multiplicative property of Z-transforms, the right-hand side of (19) can be rewritten as:

$$v \cdot (1 + r)^{T-1} \cdot \left[ \frac{Z(K | 1 + r)}{Z(X | 1 + r)} - \sum_{t=0}^{T-2} I_t \cdot (1 + r)^{-t} \right]$$

$$= (1 + r)^{T-1} \cdot \left[ c \cdot Z(K | 1 + r) - v \cdot \sum_{t=0}^{T-2} I_t \cdot (1 + r)^{-t} \right]. \quad (20)$$

The last step uses the definition of the marginal cost $c$ as:

$$c = \frac{v}{Z(X | 1 + r)}.$$  

Finally, because

$$Z(K | 1 + r) = \sum_{t=0}^{T-1} K_t \cdot (1 + r)^{-t} + \sum_{t=T}^{\infty} K_t \cdot (1 + r)^{-t}$$

and $(K_0, ..., K_{T-1})$ is a linear function of $(I_0, ..., I_{T-2})$, there exist coefficients $\{u_t\}_{t=0}^{T-2}$ such that:\(^{39}\)

$$C_{T-1}(K_T | I_0, ..., I_{T-2}) = \sum_{t=0}^{T-2} u_t \cdot I_t + c \cdot \sum_{t=T}^{\infty} K_t \cdot (1 + r)^{-t+T-1}.$$  

This proves that the date $T - 1$ marginal cost of one extra unit of capacity at date $T$ is $\gamma \cdot c$. \hfill \blacksquare

\(^{38}\)See Proakis and Manolakis (2006) for a standard reference.

\(^{39}\)As one would expect, the coefficients $\{u_t\}$ depend neither on $v$ nor on $r$, but only on $x_t$. 

27
Proof of Proposition 1: 
We first establish the following claim.

Claim Let \( \eta \in \mathbb{R}^{T+1}_+ \) and \( \theta \in \mathbb{R}^{T+1} \) be such that 
\[
\theta_0 - \eta_0 \leq \theta_1 - \eta_1 \leq \ldots \leq \theta_{T-1} - \eta_{T-1} \leq \theta_T - \eta_T = 0.
\]
Then for any \( \Delta \in \mathbb{R}^{T+1}_+ \), 
\[
\Delta \cdot \theta \geq 0 \quad (21)
\]
whenever: (i) \( \Delta \cdot \eta = 0 \) and (ii) \( \sum_{i=0}^{t} \Delta_i \cdot \eta_i \leq 0 \), for \( 1 \leq t \leq T-1 \).

Proof of Claim:
Since \( \Delta \cdot \eta = 0 \) and \( \theta_T = \eta_T \), the claim is equivalent to:
\[
\sum_{t=0}^{T-1} \Delta_t (\theta_t - \eta_t) \geq 0
\]
whenever
\[
\sum_{i=0}^{t} \Delta_i \cdot \eta_i \leq 0, \quad \text{for } 1 \leq t \leq T-1. \quad (22)
\]
By Farkas Lemma (Rockafellar, 1970), \( \Delta \cdot \theta \geq 0 \) holds for any \( \Delta \) satisfying (22), if and only if there exists a vector \( w \in \mathbb{R}^{T+1}_+ \) such that
\[
w \cdot \Gamma = -(\theta_0 - \eta_0, \theta_1 - \eta_1, \ldots, \theta_{T-1} - \eta_{T-1})
\]
where
\[
\Gamma = \begin{bmatrix}
\eta_0 \\
\eta_0 \ \eta_1 \\
\eta_0 \ \eta_1 \ \eta_3 \\
\ldots \ldots \ldots \\
\eta_0 \ \eta_1 \ \eta_2 \ \ldots \ \eta_{T-1}
\end{bmatrix}
\]
By hypothesis,
\[
\eta_{T-1} \cdot w_{T-1} = \eta_{T-1} - \theta_{T-1} \geq 0,
\]
and therefore \( w_{T-1} \geq 0 \). Proceeding to \( w_{T-2} \), we find :
\[
\eta_{T-2} \cdot w_{T-2} + \eta_{T-1} \cdot w_{T-1} = \eta_{T-2} - \theta_{T-2}.
\]
Since \( \eta_{T-1} \cdot w_{T-1} = \eta_{T-1} - \theta_{T-1} \) and \( \eta_{T-2} - \theta_{T-2} \geq \eta_{T-1} - \theta_{T-2} \), it follows that \( w_{T-2} \geq 0 \). By induction, we conclude that for any \( t \),
\[
\eta_t \cdot w_t + \sum_{i=t+1}^{T-1} \eta_i \cdot w_i = \eta_t - \theta_t,
\]

and
\[ \eta_t \cdot w_t + \eta_{t+1} - \theta_{t+1} = \eta_t - \theta_t. \]

Since \( \eta_t - \theta_t \geq \eta_{t+1} - \theta_{t+1} \), it follows that \( w_t \geq 0 \).
This concludes the proof of the Claim. \( \blacksquare \)

Suppose now the depreciation schedule \( d \) is more accelerated than \( d' \). The inequality
\[ H_T(\lambda, d) \leq H_T(\lambda, d') \]
is equivalent to:
\[ \Delta_T + (1 + \lambda_1) \cdot \Delta_{T-1} + \ldots + \prod_{i=1}^{T} (1 + \lambda_i) \cdot \Delta_0 \geq 0 \]
with
\[ \Delta_t \equiv z_t(d') - z_t(d). \]

To apply the above claim, we set:
\[ \eta_t = \gamma^t. \]

Lemma 1 shows that \( d \) is more accelerated than \( d' \) if and only if:
\[ \Delta \cdot \eta = 0 \text{ and } \sum_{i=0}^{t} \eta_i \cdot \Delta_i \leq 0, \forall 0 \leq t \leq T - 1. \]

Finally, we set
\[ \theta_0 = \gamma^T \cdot \prod_{i=1}^{T} (1 + \lambda_i) \]
\[ \theta_1 = \gamma^T \cdot \prod_{i=1}^{T-1} (1 + \lambda_i) \]
\[ \vdots \]
\[ \theta_T = \gamma^T \]

If \( \lambda_t \leq r \), we note that
\[ \gamma^{-T} [\theta_t - \eta_t] = \prod_{i=1}^{T-t} (1 + \lambda_i) - (1 + r)^{T-t} \]
\[ \leq \prod_{i=1}^{T-t-1} (1 + \lambda_i) - (1 + r)^{T-t-1} \]
\[ = \gamma^{-T} [\theta_{t+1} - \eta_{t+1}]. \]
Furthermore, \( \theta_T = \eta_T \) and therefore the conditions of the above claim are met. To prove the result for aggressive growth rates, that is, for \( \lambda_t \geq r \), we note that the inequalities in the above Claim reverse once \( \theta_t - \eta_t \geq \theta_{t+1} - \eta_{t+1} \), and these inequalities hold precisely if \( \lambda_t \geq r \) for all \( t \).

**Proof of Proposition 2:** Proposition 0 established that:

\[
H_T(\lambda, d^*) = c \cdot K_T(\lambda)
\]

for the R.P.C. rule. Therefore

\[
AHC_T(\lambda, d^*) = c.
\]

The claim now follows again from Proposition 1 which established that:

\[
H_T(\lambda, d) \leq H_T(\lambda, d^*),
\]

provided \( \lambda_t \leq r \) for all \( 1 \leq t \leq T \) and \( d \) is more accelerated than the R.B.D. rule. The same inequalities then apply to the function \( AHC_T(\lambda, d) \).

**Proof of Proposition 3:**

By definition, \( d \) is uniformly more accelerated than the R.P.C., if and only if \( z_0^o \geq 0 \) and \( z_t^o \geq z_{t+1}^o \) for \( 1 \leq t \leq T - 1 \). Therefore, it suffices to show that \( AHC_T(\lambda, d) \) is everywhere increasing in each \( \lambda_t \) whenever \( z_t^o \) is decreasing in \( t \) and \( z_0^o \) is non-negative. We first establish two auxiliary claims.

**Claim 1.** For any numbers \( g_1, g_2 \) and positive numbers \( b_1, b_2 \), the ratio \( f(y_1) = \frac{g_1 y_1 + g_2 y_1}{b_2 + b_1 y_1} \) is increasing in \( y_1 \) for \( y_1 \geq 0 \), if and only if \( \frac{g_1}{b_1} \geq \frac{g_2}{b_2} \).

**Proof.** Taking the derivative with respect to \( y_1 \), one obtains:

\[
\frac{df(y_1)}{dy_1} = \frac{g_1 b_2 - g_2 b_1}{(b_2 + b_1 y_1)^2} = \left( \frac{g_1}{b_1} - \frac{g_2}{b_2} \right) \frac{b_1 b_2}{(b_2 + b_1 y_1)^2}
\]

which is of the same sign as \( \frac{g_1}{b_1} - \frac{g_2}{b_2} \).

**Claim 2.** For any numbers \( g_1, \ldots, g_T \), positive numbers \( b_1, \ldots, b_T \) and \( y_1, \ldots, y_{T-1} \), the function

\[
f_T(y_1, \ldots, y_{T-1}) = \frac{g_T + g_{T-1} y_1 + \ldots + g_1 y_1 \cdot y_{T-1}}{b_T + b_{T-1} y_1 + \ldots + b_1 y_1 \cdot y_{T-1}}
\]

is everywhere increasing in each \( y_t \), if and only if the sequence \( \frac{g_t}{b_t} \) is decreasing in \( t \).

**Proof.** We prove the “if” part of the claim by induction. The basis of induction \( (T = 2) \) coincides with Claim 1 established above. Now, assuming that the “if” part holds for all
functions \( f_{T-1} \), we show that, if \( \frac{g_T}{b_T} \) is decreasing, then \( f_T \) is everywhere increasing in each \( y_t \).

Note that

\[
  f_T (y_1, \ldots, y_{T-1}) = \frac{g_T + G_{T-1}y_1}{b_T + B_{T-1}y_1}
\]

where \( G_{T-1} = g_{T-1} + \ldots + g_1y_2 \cdot y_{T-1} \) and \( B_{T-1} = b_{T-1} + \ldots + b_1y_2 \cdot y_{T-1} \). By Claim 1, \( f_T \) is increasing in \( y_1 \), provided \( \frac{g_T}{b_T} \leq \frac{g_{T-1}}{B_{T-1}} \). The inequalities

\[
  \frac{g_T}{b_T} \leq \frac{g_{T-1}}{B_{T-1}}
\]

imply

\[
  \frac{g_T}{b_T} b_T \leq g_T
\]

for all \( t \). Hence,

\[
  (b_{T-1} + \ldots + b_1y_2 \cdot y_{T-1}) \frac{g_T}{b_T} \leq (g_{T-1} + \ldots + g_1y_2 \cdot y_{T-1})
\]

and, therefore,

\[
  \frac{g_T}{b_T} \leq \frac{G_{T-1}}{B_{T-1}}.
\]

Thus, \( f_T \) is increasing in \( y_1 \). To show that \( f_T \) is also increasing in the other arguments, we rewrite this function as

\[
  f_T (y_1, \ldots, y_{T-1}) = \frac{G'_{T-1} + g_{T-2}y_1y_2 + \ldots + g_1y_1 \cdot y_{T-1}}{B'_{T-1} + b_{T-2}y_1y_2 + \ldots + b_1y_1 \cdot y_{T-1}}
\]

where \( G'_{T-1} = g_T + g_{T-1}y_1 \) and \( B'_{T-1} = b_T + b_{T-1}y_1 \). To apply the inductive assumption, we need to check that \( \frac{g_T}{b_T} \leq \frac{g_{T-2}}{B_{T-2}} \). Denoting \( \frac{g_{T-2}}{B_{T-2}} \) as \( l \), we have:

\[
  \frac{g_{T-1}}{b_{T-1}} \leq l
\]

or, equivalently,

\[
  g_{T-1} \leq lb_{T-1}
\]

Therefore, \( \frac{g_T + g_{T-1}y_1}{b_T + b_{T-1}y_1} \leq \frac{lb_T + lb_{T-1}y_1}{b_T + b_{T-1}y_1} = l \) and, by the inductive assumption, \( f_T \) has to be increasing in \( y_2, \ldots, y_{T-1} \).

We now turn to the “only if” part of Claim 2. Again, for \( T = 2 \), the statement coincides with the “only if” part of Claim 1. Consider any arbitrary \( T > 2 \). If \( \frac{g_T}{b_T} > \frac{g_{T-1}}{b_{T-1}} \), then we can set \( y_2 = \ldots = y_{T-1} = 0 \), so that \( f_T \) takes the following form:

\[
  f_T (y_1, 0, \ldots, 0) = \frac{g_T + g_{T-1}y_1}{b_T + b_{T-1}y_1}
\]
From the case $T = 2$, we know that $f_T(y_1, 0, ..., 0)$ is not everywhere increasing in $y_1$, therefore $f_T(y_1, ..., y_{T-1})$ can not be everywhere increasing in its arguments. Now assume that $\frac{g_T}{b_T} < \frac{g_{T-1}}{b_{T-1}}$, $f_T$ is everywhere increasing, but the sequence $\frac{g_T}{b_T}$ is not monotonically decreasing. Then, there exists $t < T - 1$, such that $\frac{g_{t+1}}{b_{t+1}} > \frac{g_t}{b_t}$. The function $f_T$, then, takes the following form:

$$f_T(y_1, ..., y_{T-1}) = \frac{g_T + ... + g_{t+1} y_1 ... y_{T-t-1} + g_t y_1 ... y_{T-t} + ... + g_1 y_1 ... y_{T-1}}{b_T + ... + b_{t+1} y_1 ... y_{T-t-1} + b_t y_1 ... y_{T-t} + ... + b_1 y_1 ... y_{T-1}}$$

Note that

$$\lim_{y_{T-t-1} \to -\infty} f_T (1, ..., 1, y_{T-t-1}, 0, ..., 0) = \frac{g_{t+1}}{b_{t+1}} > \frac{g_t}{b_t}$$

Hence, there exists $y_{T-t-1}^*$ such that $f_T (1, ..., 1, y_{T-t-1}^*, 0, ..., 0) > \left( \frac{g_{t+1}}{b_{t+1}} + \frac{g_t}{b_t} \right) / 2$. On the other hand,

$$\lim_{y_{T-t} \to -\infty} f_T (1, ..., 1, y_{T-t}^*, y_{T-t}, 0, ..., 0) = \frac{g_t}{b_t}$$

and there exists $y_{T-t}^* > 0$ such that $f_T (1, ..., 1, y_{T-t-1}^*, y_{T-t}^*, 0, ..., 0) < \left( \frac{g_{t+1}}{b_{t+1}} + \frac{g_t}{b_t} \right) / 2$.

Therefore,

$$f_T (1, ..., 1, y_{T-t-1}^*, y_{T-t}^*, 0, ..., 0) < f_T (1, ..., 1, y_{T-t-1}^*, 0, ..., 0)$$

and $f_T$ can not be everywhere increasing in $y_{T-t}$. This concludes the proof of Claim 2.  

To complete the proof, we recall that $d$ is uniformly more accelerated than the R.P.C. rule if $\frac{z_i}{x_i}$ is decreasing in $t$ and $z_0$ is non-negative. To show that $AHC_T(\lambda, d)$ is everywhere increasing in each $\lambda_t$, define:

$$A(\lambda, d) = \frac{z_0^0 + z_{T-1}^0 (1 + \lambda_1) + ... + z_1^0 \prod_{i=1}^{T-1} (1 + \lambda_i)}{x_T + x_{T-1} (1 + \lambda_1) + ... + \prod_{i=1}^{T-1} (1 + \lambda_i)}$$

and

$$B(\lambda, d) = \frac{z_0^0 \prod_{i=1}^{T} (1 + \lambda_i)}{x_T + x_{T-1} (1 + \lambda_1) + ... + \prod_{i=1}^{T-1} (1 + \lambda_i)}.$$ 

Then,

$$AHC_T(\lambda, d) = A(\lambda, d) + B(\lambda, d)$$

By Claim 2 we conclude that, under the conditions imposed, $A(\lambda, d)$ is increasing in each $\lambda_j$, $1 \leq j \leq T - 1$ ($A(\lambda, d)$ does not depend on $\lambda_T$). It remains to show that $B(\lambda, d)$ is
increasing in each \( \lambda_j \). Dividing the numerator and the denominator by \((1 + \lambda_j)\), yields

\[
B(\lambda, d) = \frac{z_0^* \prod_{i=1, i \neq j}^T (1 + \lambda_i)}{(1 + \lambda_j)^{-1} \sum_{i=1}^j \left( x_{T+1-i} \prod_{l=1}^{i-1} (1 + \lambda_l) \right) + \sum_{i=j+1}^T \left( x_{T+1-i} \prod_{l=1, l \neq j}^{i-1} (1 + \lambda_l) \right)}.
\]

Therefore, \( B = \frac{B_1}{(1+\lambda_j)^{-1}B_2 + B_3} \), where \( B_1, B_2, \) and \( B_3 \) are positive and do not depend on \( \lambda_j \). Hence, \( B \) is increasing in \( \lambda_j \).

To show the “only if” part of Proposition 3, we first note that the derivative of \( AHC_T \) with respect to \( \lambda_T \) is

\[
\frac{\partial AHC_T}{\partial \lambda_T} = B(\lambda, d) = \frac{z_0^* \prod_{i=1}^{T-1} (1 + \lambda_i)}{x_T + x_{T-1} (1 + \lambda_1) + \ldots + \prod_{i=1}^{T-1} (1 + \lambda_i)}
\]

which is of the same sign as \( z_0^* \). Therefore, if \( AHC_T \) is increasing in \( \lambda_T \), \( z_0^* \) must be non-negative. Now let us set \( \lambda_T = -1 \). Then, \( B \) equals zero for any values of \( \lambda_1, \ldots, \lambda_{T-1} \) and, hence, \( A(\lambda, d) \) must be increasing in \( \lambda_1, \ldots, \lambda_{T-1} \). Applying the “only if” part of Claim 2, concludes the proof.

**Proof of Lemma 3:**

For the R.P.C. depreciation rule we have \( z_0^* = 0 \) and

\[
z_t^* = \frac{v \cdot x_t}{\sum_i x_i \cdot \gamma^i}
\]

which implies \( \frac{x_{t+1}}{x_t} = \frac{z_{t+1}^*}{z_t^*} \) for \( 1 \leq t \leq T \). From Definition 4, depreciation schedule \( d \) is uniformly more accelerated than the R.P.C. depreciation rule if \( z_0^*(d) \geq 0 \) and:

\[
\frac{x_{t+1}}{x_t} \geq \frac{z_{t+1}^*(d)}{z_t^*(d)}.
\]

To assess these inequalities, we note that:

\[
z_{t+1}^*(d) = v \cdot (1 - \alpha) \cdot \left[ \frac{1}{T} + r \left( 1 - \frac{t}{T} \right) \right] = v \cdot (1 - \alpha) \cdot \left[ \frac{1 + rT - rt}{T} \right] \quad \text{and}
\]

\[
z_t^*(d) = v \cdot (1 - \alpha) \cdot \left[ \frac{1}{T} + r \left( 1 - \frac{(t-1)}{T} \right) \right] = v \cdot (1 - \alpha) \cdot \left[ \frac{1 + rT - rt + r}{T} \right].
\]
Therefore,
\[
\frac{x_{t+1}}{x_t} = \frac{1 - \beta t}{1 - \beta t + \beta} > \frac{z^o_{t+1}(d)}{z^o_t(d)} \iff \frac{1 - \beta t}{1 - \beta t + \beta} > \frac{1 + rT - rt}{1 + rT - rt + r}
\]
\[
\iff \frac{\beta}{1 - \beta t} < \frac{r}{1 + rT - rt}
\]
\[
\iff \beta + \beta rT - \beta rt < r - r\beta t
\]
\[
\iff \beta < \frac{r}{1 + rT} = \beta^*.
\]

(23)

**Proof of Proposition 4:**

We calculate \( \frac{\partial^2 AHC_T}{\partial \alpha \partial \lambda} \). Since

\[
z^0_t = v \cdot \alpha \geq 0;
\]

and

\[
z^t_o = v \cdot (1 - \alpha) \cdot \left[ \frac{1}{T} + r \cdot \left(1 - \frac{t - 1}{T}\right) \right],
\]

\( AHC_T(\lambda, \alpha|\beta) \) can be written as (ignoring the constant multiple \( v \)):

\[
\frac{\alpha + (1 - \alpha) \sum_t \frac{y_t}{(1 + \lambda)^t}}{\sum_t \frac{x_t}{(1 + \lambda)^t}},
\]

where \( x_t = (1 - \beta t + \beta) \) and \( y_t = \frac{z_t}{(1 - \alpha)} = \frac{(1 + rT - rt + r)}{T} \). This implies:

\[
\frac{\partial AHC_T}{\partial \alpha} = \frac{1 - \sum_t \frac{y_t}{(1 + \lambda)^t}}{\sum_t \frac{x_t}{(1 + \lambda)^t}}.
\]

(26)

For \( r < \frac{6}{T - 1} \), we want to show that if \( \lambda \geq 0 \), then \( \frac{\partial^2 AHC_T}{\partial \alpha \partial \lambda} > 0 \), i.e., (26) increases in \( \lambda \). Let \( k = \frac{1}{(1 + \lambda)} \). Then we are equivalently looking to show that \( F'(k) > 0 \), where \( k \in (0, 1] \) and

\[
F(k) \equiv \frac{\sum_t y_t k^t - 1}{\sum_t x_t k^t}.
\]

Differentiating this function, we see that \( F'(k) > 0 \) if and only if

\[
(\sum_t x_t k^t)(\sum_t t y_t k^{t-1}) - (\sum_t y_t k^t - 1)(\sum_t x_t tk^{t-1}) > 0.
\]

(27)

Since \( x_t \) is linear in \( \beta \), so is the above expression. When \( \beta \) is at its maximum allowable value of \( \frac{r}{1 + rT} \), note that \( y_t = x_t \cdot \left(\frac{1 + rT}{T}\right) \). Therefore, (27) reduces to:

\[
(\sum_t x_t k^t) \cdot \left(\frac{1 + rT}{T}\right) \cdot (\sum_t tx_t k^{t-1}) - (\sum_t y_t k^t - 1)(\sum_t x_t tk^{t-1}) + \sum_t x_t tk^{t-1}
\]

34
As a result, it is necessary and sufficient to show that (27) > 0 as β → 0.
For β = 0, x_t = 1, so (27) reduces to the following sequence of expressions:

\[
\sum_t k^t \sum_t t^2 y_t k^{t-1} - \left( \sum_t k^{t-1} \right) \left( \sum_t y_t k^t - 1 \right)
\]

\[
= \left( \sum_t k^t \right) \left[ \left( \frac{1 + rT + r}{T} \right) \sum_t t^{k-1} - \frac{r}{T} \sum_t t^2 k^{t-1} \right]
\]

\[
- \left( \sum_t k^{t-1} \right) \left[ -\frac{T}{T} + \left( \frac{1 + rT + r}{T} \right) \sum_t k^t - \frac{r}{T} \sum_t t^k \right]
\]

\[
\approx \left( \sum_t k^t \right) \sum_t t k^{t-1} + T \sum_t t k^{t-1} - \left( \sum_t t k^{t-1} \right) \left( \sum_t k^t \right) + r(1 + T) \sum_t k^t \sum_t t k^{t-1}
\]

\[
- r \cdot \left[ (\sum_t k^t)(\sum_t t^2 k^{t-1}) + (1 + T) \sum_t t k^{t-1} \sum_t k^t - (\sum_t k^t)(\sum_t t k^{t-1}) \right]
\]

\[
= T \cdot \sum_t t k^{t-1} - r \cdot \left[ (\sum_t k^t)(\sum_t t^2 k^{t-1}) - (\sum_t k^t)(\sum_t t k^{t-1}) \right].
\]

This expression is linear in r. At r = 0, it equals T · \sum_t t k^{t-1} > 0. Moreover, the term multiplying −r in (28) is equivalent to:

\[
\frac{1}{k} \cdot \left[ (\sum_t k^t)(\sum_t t^2 k^{t-1}) - (\sum_t t k^{t-1})^2 \right] > 0,
\]

where the latter sign follows from the classic Cauchy-Schwarz inequality (see, for example, Marsden (1974), p. 20). Therefore, the cross-partial is positive if and only if (multiplying through (28) by k) \( r < T \cdot g(k) \), where:

\[
g(k) = \frac{\sum_t t k^t}{(\sum_t k^t)(\sum_t t^2 k^{t-1}) - (\sum_t t k^{t-1})^2}.
\]

We next show that \( g'(k) < 0 \). Direct differentiation shows that:

\[
\frac{d}{dk} \left( \frac{1}{g(k)} \right) > 0
\]

if and only if:

\[
\left( \sum_t t k^t \right) \left( \sum_t t^3 k^t \right) > \left( \sum_t t^2 k^t \right)^2,
\]

which again follows from the Cauchy-Schwarz inequality.
Finally, returning to our main claim, consider the space of positive growth rates, i.e., \( \lambda \geq 0 \Rightarrow 0 < k \leq 1 \). Then, since \( g(\cdot) \) is a decreasing function, \( r < T \cdot g(k) \forall k \in [0,1] \) if and only if \( r < T \cdot g(1) \). Expanding the term on the right using (29), this is equivalent to

\[
r < T \cdot g(1) = \frac{T \cdot (\sum_t t)}{T \cdot (\sum t^2) - (\sum t)^2} = \frac{6}{(T-1)}. \tag{30}
\]

This is of course the necessary and sufficient condition in the statement of the proposition, thereby completing our proof.

**Proof of Proposition 5:** Assuming growth in new investments has been constant and equal to \( \lambda \) over the past \( T \) years, the average historic cost per unit of capacity is given by:

\[
AHC_T(\lambda, \alpha, \beta, r, T) = \frac{z_0^0 + z_{T-1}^0 (1 + \lambda) + \ldots + z_t^0 (1 + \lambda)^{T-1} + z_1^0 (1 + \lambda)^T}{x_T + x_{T-1} \cdot (1 + \lambda) + \ldots + x_1 \cdot (1 + \lambda)^T}.
\]

Recalling equations (24-25) in the proof of Proposition 4, straightforward (but tedious) algebra shows that the ratio \( c(\beta,r,T) \) is given by:

\[
\frac{T \cdot r \cdot h(r)}{\alpha \cdot T \cdot \lambda \cdot h(\lambda) + (1 - \alpha) \cdot (\lambda - r + r \cdot T \cdot h(\lambda))} \cdot \frac{\lambda - \beta + \beta \cdot T \cdot h(\lambda) \cdot (1 + \lambda)^{-T}}{r - \beta + \beta \cdot T \cdot h(r) \cdot (1 + r)^{-T}}. \tag{32}
\]

Since \( \Delta_1(\cdot) = \left| 1 - \frac{c(\beta,r,T)}{AHC_T(\lambda, \alpha, \beta, r, T)} \right| \), the quadrant result in Proposition 2 implies that \( \Delta_1(\cdot) \) is decreasing (increasing) in \( \beta \) for \( \beta < \beta^* \) (\( \beta > \beta^* \)) if and only if the ratio \( \frac{c(\beta,r,T)}{AHC_T(\lambda, \alpha, \beta, r, T)} \) in (32) is decreasing in \( \beta \) whenever \( \lambda < r \). Conversely, we need to show that this ratio is increasing in \( \beta \) whenever \( \lambda > r \).

Suppose first that \( 0 < \lambda < r \). The first fraction in (32) then has a positive sign and only the second fraction depends on \( \beta \). Furthermore, the sign of its derivative is negative if and only if:

\[
r \cdot (H(\lambda) - H(0)) < \lambda \cdot (H(r) - H(0)), \tag{33}
\]

where \( H(u) = T \cdot h(u) \cdot (1 + u)^{-T} \). Given the properties of \( h(\cdot) \) noted in the main text, it is readily verified that \( H(\cdot) \) is decreasing and convex with \( H(0) = 1 \) and \( H(-1) = T \). Therefore the inequality in (33) holds if \( 0 < \lambda < r \). In case \( \lambda < 0 < r \), the inequality in (33) reverses and therefore the derivative of the second fraction in (32) is positive. Yet, at the same time both terms in the denominator of the first fraction become negative, yielding the desired conclusion.
Finally, if $0 < r < \lambda$, the first fraction in (32) is positive and the derivative of the second term with respect to $\beta$ is positive because the inequality in (33) is reversed and thus $\frac{\epsilon}{AHCT(\cdot)}$ is increasing in $\beta$.

**Proof of Proposition 6:** With a constant price elasticity demand curve of the form $K_t = a_1 \cdot (1 + \mu_1) \cdots (1 + \mu_{t-1}) \cdot p^{-\epsilon}$, the optimal monopoly price is independent of the growth rates $\mu_t$ and given by

$$p^* \cdot \left[ 1 - \frac{1}{\epsilon} \right] = c.$$  

The optimal capacity and production quantity at date 1 therefore becomes:

$$q_1 = a_1 \cdot \left( c \cdot \frac{\epsilon}{\epsilon - 1} \right)^{-\epsilon}.$$  

In subsequent periods the optimal quantities are:

$$q_t = q_1 \cdot (1 + \mu_1) \cdots (1 + \mu_{t-1})$$  

Since the firm is assumed to enter the product market at date 0 (or any time thereafter), the initial investment is given by $I_0 = q_1$. In order to increase capacity at date $t$ to $q_t$, the corresponding investment levels can be calculated recursively. For the investment at date 1, we have:

$$q_1 \cdot (1 + \mu_1) = I_0 \cdot x_2 + I_1.$$  

Thus,

$$I_1 = I_0 \cdot \tilde{\lambda}_1(\mu_1),$$

where $\tilde{\lambda}_1(\mu_1) \equiv 1 + \lambda_1(\mu_1) = (1 + \mu_1 - x_2)$. To streamline the notation below, it will be convenient to write:

$$\tilde{\lambda}_1(\mu_1) = (1 + \mu_1) - r_1(x_2),$$

where $r_1(x_2) \equiv x_2$. Proceeding to $I_2$, we have the capacity condition:

$$q_1 \cdot (1 + \mu_1) \cdot (1 + \mu_2) = I_0 \cdot x_3 + I_1 \cdot x_2 + I_2.$$  

Collecting terms, yields:

$$\tilde{\lambda}_2(\mu_1, \mu_2) = (1 + \mu_1) [(1 + \mu_2) - r_1(x_2)] - r_2(x_2, x_3),$$

37
with \( r_2(x_2, x_3) \equiv x_3 - x_2^2 \). Clearly, \( \tilde{\lambda}_2(\mu_1, \mu_2) \) is increasing in both \( \mu_1 \) and \( \mu_2 \) because \( 1 - r_1(x_2) \geq 0 \). We also note that \( \tilde{\lambda}_2(\mu_1, \mu_2) \) is greater than zero for all \( (\mu_1, \mu_2) \geq (0, 0) \) because:

\[
1 - r_1(x_2) - r_2(x_2, x_3) \geq 0,
\]

owing to the fact that \( x_2 \geq x_3 \). To express \( I_t \), we introduce the notation \( \mu_t \equiv (\mu_1, ..., \mu_t) \).

At date \( t \), we have:

\[
\tilde{\lambda}_t(\mu_t) = (1 + \mu_1)[(1 + \mu_{t-1})\cdots((1 + \mu_t) - r_1(x_2)] - r_2(x_2, x_3)]
\]

The remainder terms \( r_t(\cdot) \) satisfy the inequality:

\[
1 - r_1(x_1) - r_2(x_2, x_3) - ... - r_t(x_2, x_3, ..., x_t) \geq 0,
\]

because the \( 0 \leq x_t \leq 1 \) are weakly decreasing over time. As a consequence, we can express each \( I_t \) as \( I_t = I_0 \cdot \tilde{\lambda}_t(\mu_t) \), where \( \tilde{\lambda}_t(\mu_t) \equiv 1 + \lambda_t(\mu_t) \) is increasing in each \( \mu_t \). One can now adapt the arguments in the proof of Proposition 3 to show that this yields the desired conclusion.
References


