The Measurement of Business Capital, Income and Performance

by

W. Erwin Diewert

October, 1996

W.E. Diewert: Department of Economics, University of British Columbia, #997-1873 East Mall, Vancouver, BC V6T 1Z1 Canada. Telephone: (604) 822-2544, Facsimile: (604) 822-5915.

1.0 Introduction

> "In economics it is difficult to prove originality; for the germ of every new idea will surely be found over and over again in earlier writers".
>
> *Irving Fisher*[1930; ix]

> "Capital (I am not the first to discover) is a very large subject, with many aspects; wherever one starts, it is hard to bring more than a few of them into view. It is just as if one were making pictures of a building; though it is the same building, it looks quite different from different angles". *John Hicks*[1973; v]

The main topic to be addressed in this paper is how to measure the contribution of capital to periodic business income.

As many economists and accountants have observed over the past century, economic measurement cannot be divorced from the purpose at hand.[1] The main purpose of this paper is to answer the following question: how should capital input be measured in the context of evaluating business performance over a number of accounting periods?

The fundamental problem associated with measuring the contribution of a capital input to the period by period economic performance of a business unit is the *durability* of capital: a capital input is purchased in an initial accounting period but its contribution to the production of outputs persists over several subsequent periods. Thus the initial purchase cost of the capital input cannot be *entirely* allocated to the period of purchase but it is difficult to know precisely how the initial cost should be allocated over subsequent periods. This problem of determining the period by period contributions to production and the associated costs is perhaps the fundamental problem in accounting theory.[2] The difficulties associated with this fundamental allocation problem are greatly magnified if the price level is not stable. In this paper, we will *not* assume stability of prices.

Once the initial purchase cost of a durable capital input has been allocated across accounting periods, period costs can be subtracted from period revenues and accounting period income or profits can be calculated. Thus the measurement of capital goes hand in hand with the measurement of business income: different measures for the period by period cost of capital will give rise to different income measures.[3]

However, there are still some additional problems to be addressed, associated with timing problems; i.e., exactly when should a revenue or cost item be recognized by the business unit? The following quotation from the accounting literature summarizes these timing and recognition problems:

> "In the earlier, simpler days of accounting, the convention called for realization in actual cash, which was later modified by the assumption that legal transfer of title constituted realization, the resulting account receivable merely representing a deferred change to cash, while still later the mere act of shipment and billing was assumed to be equivalent to the legal transfer of title".
>
> *Stephen Gilman*[1939; 212]

The questions raised by the above quotation cannot be answered unless we have a framework or model in mind in order to classify and evaluate these problems.[4] The framework that we will use to help us answer our measurement questions is that of production theory: we view the business unit as a black box that transforms inputs used during an accounting period into outputs that were produced during the period. This is the point of view that is adopted in Social or National Income Accounting.[5] However, it must be kept in mind that most accountants think that sale of commodities is more important than production and that realization of revenue (through sale) is more fundamental than any imputed accrual of revenues that might occur through production without sale.[6] We will develop our production theory framework in more detail in section 5 below but for now, we note that the main ideas are due to the economist Hicks [1946; 191-201 and 325-326] and the accountants Edwards and Bell [1961; 71-72]. Hicks[7] realized the usefulness of his intertemporal production theory framework in being able to cast some light on a wide variety of economic phenomena; we shall make use of his intertemporal framework in sections 5, 10 and 11 below.

Once the revenues and costs which pertain to the outputs produced and the inputs utilized by a production unit during an accounting period have been determined, we can address the problem of evaluating the performance or efficiency of the unit. This can be done in at least three ways: (i) the net income or profits (the difference between revenues and costs) of the unit in one period can be compared to the profits of other units or the profits of the same unit in a previous period; (ii) the rate of return on assets employed (equal to net income divided by the value of assets employed by the unit at the beginning of the accounting period) can be compared to rates of return on similar business units or the same unit in a prior period, and (iii) the productivity change of the business unit going from one period to another can be computed.

If there were only one output produced and one input utilized by a production unit, the ex post evaluation of the performance of one period (say period 1) with another period (say period 0) would be very straightforward from the viewpoint of the productivity criterion. Let y^0 and y^1 be the physical amounts of output produced during periods 0 and 1 and let x^0 and x^1 denote the physical amounts of inputs used during the two periods by the production unit. Then the productivity change going from period 0 to 1 can be defined as the rate of growth of output y^1/y^0 divided by the rate of growth of input x^1/x^0, or alternatively, as the ratio of the period 1 output-input coefficient y^1/x^1 to the period 0 output-input coefficient y^0/x^0; i.e., we have

(1) $\quad$ Productivity $\equiv [y^1/y^0]/[x^1/x^0] = [y^1/x^1]/[y^0/x^0]$.

If the productivity number defined by (1) is greater than unity, then we say that the production unit exhibited a productivity improvement or that the production unit was more efficient in period 1 than in period 0.

In the case of many inputs and many outputs, the definition of productivity change or efficiency improvement is much more difficult and controversial. One can approach the problem of measuring productivity in the multiple input, multiple output context in at least three ways: (i) use output and input price information to weight the individual output and input growth rates (i.e., use index number techniques); (ii) use econometric techniques to estimate production or cost functions or (iii) use information on output and input quantities to construct nonparametric estimates of production unit efficiency along the lines pioneered by Farrell [1957] and Afriat [1972]. We will not utilize approaches (ii) and (iii) to efficiency measurement in the present paper[8] but the index number approach (i) will be discussed briefly in section 12 below and used in section 13.

We summarize the above discussion as follows: we want to measure the cost of capital inputs in an inflationary environment in the context of determining the period by period costs and revenues that pertain to a business unit. Once these periodic costs and revenues have been determined, periodic business income (or profits) can be calculated along with other measures of business performance.

Once we have considered the problems involved in defining periodic business income, it is natural to consider the implications of taxing alternative definitions of business income. Thus in section 11 below, we shall devote some attention to the problem of determining an appropriate base for taxing business income.

Edwards and Bell [1961; 271] noted that accounting measurement has three main functions: (i) to prevent fraud and theft; (ii) to evaluate business performance[9] and (iii) to provide a sound and equitable basis for taxation. Carsberg [1982; 64] added a few more functions: (iv) the measurement of current income may help investors estimate future cash flows from the enterprise and (v) accounting income is an indicator of the amount that can be distributed to owners without impairing the long run viability of the enterprise.[10] It can be seen that our examination of the problems involved in the measurement of capital has led us to consider some of the same measurement problems that have been studied in the accounting literature–namely items (ii) and (iii) in the above list.[11]

We turn now to a brief description of the contents of each section in the present paper. In section 2, we develop the idea of the ex post cost of a durable input. This is the fundamental concept which allows us to allocate the cost of a durable input across the accounting periods during which it is used in production. An interest rate plays a key role in the definition of user cost. Some accountants maintain that interest (in particular imputed interest on equity capital) is not a cost of production, so in section 3, we consider whether interest is a "valid" cost or not. Depreciation rates also play a role in user cost formulae. Hence in section 4, we consider the problems involved in empirically determining these depreciation rates.

In section 5 below, we examine more closely the production function model that we are using to frame our analysis of capital measurement problems. We shall distinguish the Hicksian multiperiod intertemporal production function, the single period Austrian production function, the single period reduced form "traditional" production function and the interrelationships among these functions. We will also discuss the fundamental question of what exactly are admissible productive activities and the implications of answers to this question for capital measurement.

Sections 6 and 7 deal with the core controversies in the capital measurement literature. Section 6 looks at alternative methods of asset valuation; i.e., alternative methods for determining the period by period value of a durable capital input that is held by a business enterprise over multiple accounting periods. Section 7 then looks at the implications of alternative asset valuation methods for the determination of the business unit's period by period income.

Section 8 deals with a question which is rarely discussed in the measurement literature: what is the ideal length of the accounting period? It turns out that the amount of inflation that the economy is experiencing impacts on the answer to this question. In order to calculate productivity measures for a business unit, it is necessary to decompose periodic value flows pertaining to the same commodity (i.e., sales or purchases of a homogeneous commodity made over the accounting period) into price and quantity components. Exactly how this decomposition should be accomplished is also discussed in section 8.

Section 9 focuses on the measurement problems that arise when we attempt to adapt our basic user cost framework to the treatment of inventories. The analysis here draws heavily on Diewert and Smith [1994].

Section 10 attempts to demonstrate the following proposition: with a proper treatment of interest, *any* method of depreciation (including immediately expensing the purchase of a durable capital input) will lead to the same discounted stream of period by period profits for a business unit. However, this proposition does not mean that the method of depreciation is unimportant: *different* depreciation methods will lead to *different* estimates of short run periodic profits and hence to *different* estimates of the short run efficiency of the business enterprise.

In section 11, we consider some of the problems posed by the existence of business income taxes. Are these taxes costs or distributions of financial capital to the government?

If they are costs, can they be allocated to specific commodities? We also draw on the results of section 10 to analyze the economy wide efficiency and neutrality aspects of alternative business income concepts as bases for income taxation. The distortions generated by basing business income taxes on a historical cost income concept in a period of high inflation have been noted by accountants for at least 75 years.[12]

Section 12 discusses various approaches to the measurement of business efficiency over two or more accounting periods (or comparable business units). In particular, index number methods for making productivity comparisons are discussed.

Section 13 draws on sections 5 and 12 in order to discuss alternative index number methods for aggregating capital. In particular, the merits of gross versus net capital stock concepts will be discussed.

Traditional bilateral index number theory is based on the implicit assumption that the set of commodities to be aggregated is the *same* in the two periods under consideration. The existence of new capital goods creates problems for this traditional paradigm; we address these problems in section 14.

Up to this point, we have been concerned only with the problems involved in accounting for durable capital inputs. In section 15, we present a brief discussion on how our user cost framework can be extended to deal with the problems involved in accounting for financial assets.

The problems involved in accounting for capital in a regulatory context are more severe than in the unregulated context because we cannot rely on the assumption that the "market" will be able to see through any distortions of the firm's financial position that might be generated by the conventions of historical cost accounting. In the context of traditional regulation, the regulator forces the revenues of the regulated business unit to equal its costs. If historical cost accounting is used in an inflationary environment, costs will be understated and the long run viability of the regulated firm will be impaired. Section 16 discusses some of these problems of accounting in a regulatory environment.

As the reader progresses through the various sections of this paper, it will become apparent that historical cost accounting is a very flawed instrument in the context of evaluating firm performance and of providing an equitable base for income taxation, particularly in an inflationary environment. Even though accountants have known about these problems associated with the use of historical cost for a long time, historical cost accounting persists in virtually all countries to the present day. In section 17, we attempt to answer the question: why does historical cost accounting persist?

Section 18 concludes.

I conclude this section with some personal observations. As will become apparent in reading this paper, virtually all of the ideas presented here have their roots in rather ancient academic literature. However, I hope that the present paper may be useful to accountants, economists, engineers and management scientists who may not be aware of contributions to the capital measurement literature which have been made by other disciplines. Also I hope that the organization of the paper and the framing of the issues will help the reader absorb the main points of a rather vast literature in a relatively painless manner. In particular, section 7 may help the reader grasp the essence of alternative income concepts as special cases of a general valuation framework. The reader will also note that I have attempted to trace the origins of ideas as far back as limited time resources would allow. Undoubtedly, many ideas could be traced back even further but I hope that the detailed references to the literature will be helpful to readers who want more background information or alternative treatments of the topic under consideration. Finally, this paper can be viewed as a sequel to my first survey of capital measurement problems, Diewert [1980b]. There is nothing terribly wrong about this earlier paper (as far as I can see), but I now realize (as Hicks [1973; v] did later in life) that there are many aspects to capital measurement and my earlier survey did not address many of the topics considered in the present paper.

2.0 The Ex post User Cost of a Durable Input

> "The machine should be sold if its expected quasi-rent for the coming period is less than the interest on disposal value now plus the expected decline in disposal value over the period". *Edgar O. Edwards and Philip W. Bell*[1961; 172]

> "In a perfectly competitive world the annual rent of a machine would equal the marginal product of its services. The rent itself would be determined by the interest costs on the investment, the deterioration in the future productivity of the machine due to current use, and the expected change in the price of the machine (obsolescence)". *Zvi Griliches*[1963; 120]

In order to measure the contribution of a capital stock component to the production of current period outputs, it is necessary to recognize that a capital input is not like other nondurable inputs for at least two reasons: (i) a capital input is *durable* and hence by definition, it is not completely used up in the period of its purchase and hence (ii) the purchase price of a capital input should not be entirely allocated to current period costs but instead it should be allocated over all periods from the period of purchase to the period when the capital good is scrapped or sold. How should this cost allocation over the useful life of the capital input be accomplished? We shall give a preliminary answer to this question in this section but we shall revisit this problem of intertemporal cost allocation in subsequent sections.

To provide a preliminary answer to the cost allocation problem, we specialize the dynamic production theory framework of Hicks [1946; 193-4] to work out the *ex post user cost* for a durable input. Consider a durable input that can be purchased at the beginning of period 0 at the price P^0. After using the input during period 0 the producer could sell the depreciated or "used" capital input at the beginning of period 1 price P_u^1. Assume that the producer's opportunity cost of financial capital at the beginning of period 0 is r^0. Then the present value of the net cost of buying one unit of the durable input, using it for one accounting period and selling it at the end of period 0 is the following *ex post user cost:*[13]

$$\begin{aligned} w^0 &\equiv P^0 - P_u^1/(1+r^0) \\ &= [r^0P^0 + (P^0 - P_u^1)]/(1+r^0). \end{aligned} \tag{2}$$

The first term on the right hand side of (2) is an interest cost term while the second term combines the effects of depreciation and capital gains (or losses) on the sale of the asset. These effects can be separated as follows. Let P^1 be the beginning of period 1 price for one unit of a capital input which is physically equivalent to the capital input which was purchased at the beginning of period 0. We can now use the end of period 0 "new" good price P^1 and the end of period 0 "used" good price P_u^1 to define the end of period 0 *economic depreciation rate* δ^0 as follows:[14]

$$1 - \delta^0 \equiv (P_u^1/P^1). \tag{3}$$

The logic behind definition (3) is that P^1 represents the price of a "new" capital input at the end of period 0 while P_u^1 represents the end of period 0 price of the same capital input that has been used for one additional accounting period. Thus $(1-\delta^0) = P_u^1/P^1$ is the ratio of "used" to "new" value evaluated at the end of period 0. Replacing P_u^1 in (2) by $(1-\delta^0)P^1$ leads to the following user cost formula:

$$w^0 = [r^0P^0 + \delta^0P^1 - (P^1 - P^0)]/(1+r^0). \tag{4}$$

The three terms in the numerator of (4) are the usual components that appear in user cost formulae: the first term r^0P^0 reflects the interest opportunity cost of using financial capital to purchase and hold the input over the accounting period; the second term δ^0P^1 is a depreciation term that reflects the fact that the input has been utilized during period 0 and hence its useful life has been reduced by one period and the final term, $-(P^1 - P^0)$, is the negative of the capital gains made by a new unit of the asset over period 0.

The ex post user cost formula (4) can be further simplified if we define the "new" asset inflation rate i^0 over period 0 by

$$1 + i^0 \equiv P^1/P^0 \tag{5}$$

and substitute (5) into (4) in order to obtain the following ex post user cost formula:

$$(6)\quad w^0 = [r^0 - i^0 + \delta^0(1+i^0)]P^0/(1+r^0).$$

Note that $r^0 - i^0$ is an asset specific real interest rate.

What are the informational requirements for the calculation of the ex post user cost defined by (6)? We require time series for: (i) new asset prices P^t; (ii) the opportunity cost of capital r^t and (iii) economic depreciation rates δ^t.

The user cost formulae (2), (4) and (6) discounted the end of period 0 price of the asset back to the beginning of period 0. Thus these user costs can be interpreted as beginning of the accounting period prices. For accounting purposes, it is often more convenient to work with end of the period prices.[15] To obtain end of period user costs, we need only multiply the w^0 defined by (2), (4) or (6) by $(1+r^0)$. Thus the *end of period 0 ex post user cost* counterparts to (2), (4) or (6) are:

$$
\begin{aligned}
&(7)\quad w_e^0 \equiv r^0P^0 + P^0 - P_u^1 \\
&(8)\quad\quad\; = r^0P^0 + \delta^0P^1 - (P^1 - P^0) \\
&(9)\quad\quad\; = [r^0 - i^0 + \delta^0(1+i^0)]P^0.
\end{aligned}
$$

User cost formulae of the type (9) (with i^0 set equal to zero) date back to Walras [1874; translated 1954; 269]. Other contributors to the user cost literature who derived more complex user cost formulae with income tax complications include Jorgenson [1963] and Hall and Jorgenson [1967] (in a continuous time framework) and Christensen and Jorgenson [1969] (in a discrete time framework).[16]

To see how (7) could be justified directly from an accounting perspective, suppose that the production unit purchases one unit of the asset at the beginning of period 0 at the price P^0 and at the end of period 0 (or equivalently, at the beginning of period 1) the asset is worth[17] P_u^1. Thus the net cost of using the asset over the period would appear to be $P^0 - P_u^1$. However, this net cost neglects interest cost considerations. Suppose that the production unit can issue debt at the beginning of period 0 at the interest rate r_d^0 and can raise equity capital at the beginning of period 0 cost of capital r_e^0. Suppose further that the production unit finances the fraction f^0 of its beginning of period 0 purchase of the asset by issuing debt and the fraction $1 - f^0$ by issuing equity. Then its average cost of capital (including the opportunity cost of equity capital) will be:

$$(10)\quad r^0 \equiv f^0r_d^0 + (1-f^0)r_e^0.$$

Thus the net cost of using the asset over period 0, including interest and the opportunity cost of equity capital, is $P^0 - P_u^1 + r^0P^0$, which is (7).

Should we use the beginning of the period user cost formulae (2), (4) or (6) or the end of period formulae (7) - (9) in order to determine the period 0 cost of a durable capital input which is available for use at the beginning of the period? Now period 0 capital costs are going to be determined as the product of user costs times the number of units of each capital stock component which is available at the beginning of the period. These period 0 capital costs will be added to other period 0 flow costs (such as labour, fuel and telecommunications costs) in order to determine total period 0 costs. Thus if we use the user cost formulae (2), (4) or (6), we are implicitly assuming that all period 0 flow costs are equivalent to beginning of the period values or equivalently, that all of these flow costs are "realized" at the beginning of the period. On the other hand, if we use formulae (7) - (9), we are implicitly assuming that all period 0 flow costs are "realized" at the end of period 0 (which seems more sensible). In fact, it may be more accurate to assume that flow costs are "realized" on average at the midpoint of the accounting period and to provide an interest rate adjustment to these flow costs to make them comparable to either the beginning or end of the period user costs. However, for most practical purposes, it will not be worth the trouble of carrying out this extra discounting adjustment. In view of the fact that the treatment of interest as a cost is much more straightforward when we use the user cost formula (7), we recommend that the user cost formulae (7) - (9) be used together with the usual period 0 costs for nondurable inputs.[18]

In the following section, we briefly consider the question of whether interest really is a cost of production.

3.0 Interest as a Cost of Production

> "The stock which is lent at interest is always considered as a capital by the lender. He expects that in due time it is to be restored to him, and that in the mean time the borrower is to pay him a certain annual rent for the use of it". *Adam Smith*[1776; reprinted 1963; 271]

> "And human nature being what it is, we are justified in speaking of the interest on capital as the reward of the sacrifice involved in the waiting for the enjoyment of material resources, because few people would save much without reward; just as we speak of wages as the reward of labour, because few people would work hard without reward". *Alfred Marshall*[1920; 232]

It is clear that economists from Adam Smith through Alfred Marshall regarded interest as a reward (to the lender of financial capital for abstaining from or deferring consumption and as a cost of production for producers who are recipients of the financial capital.[19]

However, the problem of explaining the factors that determined the interest rate was a much more difficult task. Böhm-Bawerk [1888; translated 1891; 24-72] summarized the literature on this topic up to his time and provided a verbal description of a modern theory of interest[20] while Fisher [1930] presented a very convincing algebraic and geometric description of the same theory.[21] Böhm-Bawerk [1891; 285-6], Fisher [1897; 522] and Hicks [1946; 141-2] explained how the present price of a good purchased now for delivery next period is equal to the (spot) price of the good next period divided by one plus the current period interest rate. Hicks [1946; 136] generalized the simple one (spot) commodity and multiple time period models of Fisher [1930] into a general model with many commodities and many time periods (his "Futures Economy" where all commodities can be bought and sold on forward markets) and Debreu [1959] provided a rigorous proof of the existence of equilibrium in such a model. But Hicks [1946; 119-127] also developed another model of intertemporal equilibrium that had a theory of the interest rate built into it– the temporary equilibrium model.[22] This second Hicksian model used the same building blocks as the futures economy model, except that instead of assuming the existence of futures markets, Hicks assumed the existence of current period (spot) markets for commodities and financial capital and the existence of definite expectations about future period spot prices (which could depend on current period prices) for all consumers and producers in the economy. In this model, these expected future period spot prices were used by producers and consumers in their intertemporal profit maximization and utility maximization plans.

The above sketch of the role of interest in modern economic theory would seem to indicate that economists generally accept interest as a valid cost of production, and indeed, interest plays a vital role in the intertemporal allocation of resources. However, many accountants and some economists object to interest as a cost of production; in particular, they object to interest that is imputed to the equity capital employed by a business unit.

A few accountants objected to associating an interest cost with the use of a durable input over an accounting period (in addition to a depreciation cost) on the grounds that such interest rate adjustments are likely to be minor in view of the errors involved in estimating depreciation.[23] However, if interest rates are high and the durable input is long lived, other accountants have pointed out that the neglect of interest can lead to substantial underestimation of costs.[24] On the other hand, the main objection of accountants[25] and some economists[26] to the inclusion of interest on equity capital as a cost is that it is an imputation or estimated value and accountants should stick to recording values rather than creating them. These objectors have a valid point: it is not a trivial matter to determine precisely what are the relevant debt and equity interest rates, r_d^0 and r_e^0, which appeared in (10) above.

Let us consider the problem of estimating the period 0 debt interest rate r_d^0 first. Suppose that we can determine the beginning of period 0 market values for all of the debt obligations of the enterprise under consideration, say B_j^0 for $j = 1, \ldots, J$ where B_j^0 is the beginning of period 0 market value of the jth bond. Suppose also that we can determine the interest payments b_j^0 that can be associated with the jth bond for period 0 for $j = 1, \ldots, J$. The effective period 0 interest rate r_j^0 for the jth debt instrument is defined as:

(11) $\quad r_j^0 \equiv b_j^0/B_j^0, j = 1, \ldots J.$

Define the total market value of the beginning of period 0 debt as[27]

(12) $\quad B^0 \equiv \Sigma_{j=1}^J B_j^0.$

The overall ex ante period 0 debt interest rate could now be defined as

(13) $\quad r_d^0 \equiv \Sigma_{j=1}^J B_j^0 r_j^0/B^0 = \Sigma_{j=1}^J b_j^0/B^0.$

The first equation in (13) indicates that r_d^0 can be regarded as a weighted average of the individual bond effective interest rates r_j^0 while the second equation indicates that the overall debt interest rate is equal to total period 0 interest payments $\Sigma_{j=1}^J b_j^0$ divided by the initial market value of the debt B^0.

If all of the debt instruments were one period bonds, (13) would be the end of the story. However, if some of the bonds had a period to maturity greater than the length of period 0 (the usual situation!), then the above analysis is unsatisfactory since it ignores possible changes in the market value of the bonds as we go from the beginning of period 0 to its end. To deal with this complication, define B_j^1 as the market value of the jth bond at the end of period 0.[28] Any appreciation in value of the jth bond over period 0, $B_j^1 - B_j^0$, should be added to the end of period 0 interest payments b_j^0 to get the total cost that can be imputed to the jth debt instrument over period 0. Thus the end of period 0 effective interest rate for the jth bond (including accrued capital costs) is

(14) $\quad \bar{r}_j^0 \equiv [b_j^0 + (B_j^1 - B_j^0)]/B_j^0; j = 1, \ldots, J.$

If market interest rates increase over period 0 (due to say unanticipated general inflation), then B_j^1 will be less than B_j^0 and the terms $B_j^1 - B_j^0$ in (14) will be negative and hence will act to reduce effective interest rates. Now define the overall ex post[29] debt interest rate as the following weighted average of the individual $\bar{r}_j^0$ defined by (14):

(15) $\quad \bar{r}_d^0 \equiv \Sigma_{j=1}^J B_j^0 \bar{r}_j^0/B^0$

(16) $\quad \qquad = \Sigma_{j=1}^J [b_j^0 + (B_j^1 - B_j^0)]/B^0$

where (16) follows from (15) and (14). The $\bar{r}_d^0$ defined by (15) could be used as the debt interest rate r_d^0 which appears in (10) above.[30] Thus if beginning and end of period market values for the firm's debt instruments are available, the problem of defining an overall effective ex post debt interest rate seems straightforward.

However, the problems involved in estimating the firm's period 0 opportunity cost of equity capital, the r_e^0 which appears in (10), are far from straightforward. To conclude this section, we shall outline six possible approaches to the determination of r_e^0.

Approach 1: Discounted Cash Flow

Suppose that a company's current period 0 dividends D^0 are expected to grow at the constant real rate $\tilde{g}$ for the indefinite future and that the expected inflation rate for the indefinite future is $\tilde{i}$. The company's current share price S^0 should equal the discounted future expected dividends. The discount rate should be the long run cost of equity capital r_e^0 minus the anticipated inflation rate $\tilde{i}$. Under these assumptions, we should have the following relationship between the company's current share price S^0 and current dividend rate D^0:

$$(17) \quad S^0 = D^0/[\tilde{r}_e^0 - \tilde{i} - \tilde{g}].$$

Formula (17) can be rearranged to give the following formula for the cost of capital:

$$(18) \quad r_e^0 = [D^0/S^0] + \tilde{i} + \tilde{g}.$$

This method for determining the opportunity cost of equity capital is due to Williams [1938] and Gordon and Shapiro [1956]. According to Myers [1992; 489], this method is widely used to determine allowed equity rates of return for regulated utilities in the United States.

There are many problems with this method. The determination of the anticipated future inflation rate $\tilde{i}$ will be problematical given that past inflation rates have been very variable during the past two decades. Dividend growth rates are also variable over the business cycle. Finally, dividend price ratios of the form D^0/S^0 are also tremendously variable and moreover, this method is not suitable for the determination of an economy wide equity cost of capital since many businesses are not incorporated and many incorporated businesses do not have publicly traded shares.

Approach 2: The Capital Asset Pricing Model

Under certain assumptions, the capital asset pricing model of Sharpe [1964], Lintner [1965] and Mossin [1966] yields the following relationship between the expected cost of capital for a company $\tilde{r}_e^0$, a safe or riskless interest rate r_s^0 and the expected return on a market portfolio of assets $\tilde{r}_m^0$:

$$(19) \quad \tilde{r}_e^0 = r_s^0 + \beta[\tilde{r}_m^0 - r_s^0]$$

where β is the covariance between the company's equity rate of return and the market portfolio rate of return divided by the variance of the market portfolio rate of return. Given a time series of ex post company rates of return r_e^t, market rates of return r_m^t and the safe rate of return r_s^t, ex post returns can be substituted into (19) in place of anticipated returns and β can be estimated in a regression model.[31] Alternatively, an estimate for β can be constructed by taking an average of past covariances $Cov(r_e^t, r_m^t)$ divided by $Var(r_m^t)$. Given this estimator for β, an ex ante $\tilde{r}_e^0$ can be calculated as the right hand side of (19) where $\tilde{r}_m^0$ is a forecast for the period 0 market rate of return.

Some of the assumptions that are required to derive (19) are: (i) each investor is a von Neuman and Morgenstern expected utility maximizer with the same preferences over current period consumption and end of the period wealth; (ii) a riskless one period asset actually exists; (iii) all investors have preferences over the same set of risky assets and the common riskless asset; (iv) all investors have the same expectations about the returns, variances and covariances of the risky assets and (v) there are no transactions costs. All of these assumptions are somewhat suspect from the empirical point of view. Machina [1992; 860-862] documents some of the empirical evidence which contradicts the expected utility model. In particular, Mehra and Prescott [1985] show that the equity premium over the safe asset seems to be too large for generally agreed upon values of relative risk aversion. Epstein and Zinn [1990] explain this equity premium by a generalization of the usual expected utility model that allows for first order risk aversion.[32] Assumption (ii), the assumption that a perfectly safe one period asset exists, is also problematical: nominal government bonds are not risk free due to inflation risk.[33] Assumption (iii) is also dubious: what is the relevant set of risky assets facing any investor? Should we include housing or foreign stock markets? Our β estimates will generally change as we change our definition of the "market" for risky assets. Assumption (iv) is also problematical: what will happen to our estimate for β as we include or exclude data for 1987–the year of the great worldwide stock market crash? Finally, assumption (v) is also far from being satisfied.

Although the capital asset pricing model could be used to estimate the cost of equity capital for some companies whose shares are traded in a stock market, it cannot be used to estimate the cost of equity capital for many companies and for the economy as a whole since a large proportion of private sector companies are not listed on any stock exchange.

Approach 3: The Ex Post Return Method

In this approach, the company's ex post operating income for the period under consideration is equated to the sum of ex post user costs of the type defined by (9) where each user cost is weighted by the quantity of capital used. The resulting equation is solved for r_e^0, the ex post return to equity capital. Suppose that there are J types of capital in use during period 0. Denote the beginning of period 0 price of the jth type of capital by P^{0j}, denote the jth depreciation rate by δ^{0j} and denote the jth asset appreciation rate in period 0 by i^{0j} for $j = 1, \ldots, J$. Finally denote the quantity of the jth type capital used during period 0 by k^{0j} for $j = 1, \ldots, J$. The ex post rate of return on equity capital r_e^0 can then be determined as the solution to the following linear equation where OI^0 represents period 0 operating income:

$$(20) \quad OI^0 = \Sigma_{j=1}^J [f^0 r_d^0 + (1 - f^0) r_e^0 - i^{0j} + \delta^{0j}(1 + i^{0j})] P^{0j} k^{0j}.$$

Formula (20) can be simplified if we replace the weighted sum of debt and the equity cost of capital, $f^0 r_d^0 + (1 - f^0) r_e^0$, by a single overall ex post cost of capital r^0. Thus r^0 can be determined by solving the following equation:

$$(21) \quad OI^0 = \Sigma_{j=1}^J [r^0 - i^{0j} + \delta^{0j}(1 + i^{0j})] P^{0j} k^{0j}.$$

The ex post cost of capital method for determining the opportunity cost of capital that is based on solving equation (21) for r^0 is due to Christensen and Jorgenson [1969] and has recently been advocated in the regulatory context by Christensen, Schoech and Meitzen [1995]. In addition to the simplicity of this method, Christensen, Schoech and Meitzen [1995; 10] note that this method can be applied in a symmetric manner to both a single enterprise as well as to the economy as a whole.

The problem with the ex post return method for estimating r^0 is that it does *not* correspond to a true opportunity cost of capital for the business unit; instead, it corresponds to an ex post measure of period 0 *performance* for the business unit.

Approach 4: The Weighted Average of Past Ex Post Returns

Instead of using the ex post returns to capital method outlined in Approach 3 above, we could switch to a forecasting framework, using a weighted average of past ex post returns to capital to forecast a current period opportunity cost of capital.

The problem with methods of this type is their arbitrariness: which ex post approach to the determination of the opportunity cost of capital should be used? Which forecasting method should be used? How far back in time should we go? It will be difficult to reach agreement on what is the most reasonable specific method in this general class of methods. Moreover, as we argued at the end of Approach 3 above, ex post returns incorporate pure profits (or losses) and hence are not true opportunity costs for equity capital.

Approach 5: The Use of An Exogenous Market Interest Rate

In this method, a relevant market interest rate is used as a proxy for the equity opportunity cost of capital. This market interest rate could be: (i) the prime lending rate that banks or other financial institutions charge borrowers in "similar" lines of business; (ii) the business unit's period 0 ex ante average interest rate defined by (13) or (iii) an index of ex ante interest rates of the form (13) where the B_j^0 now refer to market capitalizations of beginning of period 0 debt instruments for business units j that are "similar" to the business unit being studied.[34] As an example of (iii), Christensen, Schoech and Meitzen [1994] used the Moody's public utility bond as a proxy for the cost of capital for a regulated utility. These authors noted that this method has the advantages that the Moody bond yield is publicly available and is updated annually.[35]

Approach 6: The Use of An Official Rate of Return

In this approach, a government or regulatory agency would set an "official" interest rate that could be used to approximate a business unit's cost of equity capital. For example, the "official" rate might be: (i) the interest rate that is used by the taxation authorities to assess late payment of income taxes; (ii) an equity interest rate that is recommended by the country's accounting standards board or (iii) the midpoint of a regulator's range of acceptable returns to equity capital for a regulated firm.

A problem with this method is that there is no guarantee that the "official" rate set by a taxation authority, accounting standards board or regulator will be "reasonable"; i.e., this method gives no guidance on how the authority will in fact determine the "official" rate. In practice, official rates determined by the tax authorities are probably based on Approach 5(iii) outlined above.

Of the 6 broad approaches to the empirical determination of an equity opportunity cost of capital, I feel that Approaches 5 and 6 are fairly satisfactory.

Why is it important to recognize imputed interest on equity capital as a cost of production? The following quotations answer this question:

> "Once again, the basic reason why interest on the use of total capital should be recorded as a cost is that interest *is* a cost .... A company has not performed satisfactorily, either for its shareholders or for society, if it has not generated enough revenue to cover all its costs, including the cost of using capital. The current income statement does not show whether or not the company has met this fundamental test. Its implication is that any earnings above the cost of debt interest are a 'plus' ". *Robert N. Anthony*[1973; 96]

> "The argument in favor of including interest as an element of cost is twofold. From the viewpoint of the business as a whole, it helps to point out an important fact to the managers of any enterprise which persistently fails to return a normal current rate of interest on the investment. From the more detailed cost accounting viewpoint, it is said to make an important cost distinction between those manufacturing departments using costly machinery and those using inexpensive machinery or none at all". *Stephen Gilman*[1939; 322]

We turn now to the problems involved in determining the depreciation rates which appear in user cost formulae.

4.0 The Determination of Depreciation Rates

> "Depreciation is defined simply as rate of decrease of value".
> *Harold Hotelling*[1925; 341]

> "The net stock concept is motivated by the observed fact that the value of a capital good declines with age (and/or use). This decline is due to several factors, the main ones being the decline in the life expectancy of the asset (it has fewer work years left), the declines in the physical productivity of the asset

(it has poorer work years left), and the decline in the relative market return for the productivity of this asset due to the availability of better machines and other relative price changes (its remaining work years are worth less). One may label these three major forces as exhaustion, deterioration, and obsolescence".

Zvi Griliches[1963; 119]

Definition (3) in section 2 above defined the depreciation rate of a durable input in terms of the decline in value of a "new" machine or other durable input compared to a "used" machine that had been used one additional accounting period. The two values that were compared were market values that pertained to the end of the accounting period. If the price of "new" machines were the same at the beginning and the end of the accounting period (so that $P^0 = P^1$), then the decline in the value at the end of the period, P_u^1/P^1, would be the same as the decline in value over the period, P_u^1/P^0. Many of the early treatments of depreciation implicitly assumed price stability (i.e., $P^0 = P^1$), and hence depreciation was identified with the decline in value of the durable input over the accounting period.[36] As the above quotation by Griliches [1963; 119] indicates, economists tried to analyze the factors that determine depreciation rates.[37]

Accountants, engineers, statisticians and economists have all made contributions to the literature on depreciation. We shall review many of their approaches in sections 4.1 to 4.6 below. Sections 4.1 and 4.2 summarize some of the early accounting approaches to the treatment of depreciation. These early approaches do not lead to empirical estimates of depreciation rates as we have defined them (recall (3) above); however, the approaches outlined in sections 4.3 to 4.6 do lead to estimates of depreciation rates. Section 4.3 reviews accounting approaches while sections 4.4 – 4.6 review some economic approaches.

4.1 Early Accounting Approaches to the Measurement of Depreciation

"[There are] various methods of estimating the Depreciation of a Factory, and of recording alteration in value, but it may be said in regard to any of them that the object in view is, so to treat the nominal capital in the books of account that it shall always represent as nearly as possible the real value. Theoretically, the most effectual method of securing this would be, if it were feasible, to Revalue everything at stated intervals, and to write off whatever loss such valuations might reveal without regard to any prescribed rate .... The plan of valuing every year instead of adopting a depreciation rate, though it might appear the more perfect, is too tedious and expensive to be adopted .... the next best plan, which is that generally followed ... is to establish average rates which can without much trouble be written off every year, to check the result by complete or partial valuation at longer intervals, and to adjust the depreciation rate if required". *Ewing Matheson*[1884; 35]

"One of the first clear references to depreciation accounting was in the annual report of the Baltimore and Ohio Railroad for the year ended September 30, 1835. That report explained that income for the year was determined 'after carrying $75,000 to the debit of profit and loss to make good deterioration of the railway and machinery . . .'. During the years following 1835, there was no consistent policy followed by any group of companies or even by any one company. Apparently, some companies made a separate provision for depreciation as did the Baltimore and Ohio Railroad, while other companies charged replacement costs to expense in lieu of depreciation". *P.D. Woodward*[1956; 71]

The very earliest treatments of the depreciation problem seem to have been on the basis of periodic appraisals of the value of fixed assets. Thus our first early approach to the determination of depreciation is (i) the appraisal approach: changes in appraised values, if negative, were regarded as costs to be charged to the accounting period between appraisals.[38] However, as the quotation by Woodward above indicates, there were two additional early treatments of depreciation: (ii) engineers made estimates of the value of the physical deterioration and loss of productive life that equipment and machinery might have experienced during an accounting period and (iii) new purchases of durable inputs were simply expensed in the period of purchase.

Obviously, the third approach (which is consistent with cash flow accounting) is not helpful in the determination of periodic income, which is the focus of the present paper. The other two approaches are reasonable but not helpful in the context of the explicit determination of depreciation rates: approach (i) mixes up capital gains with the determination of depreciation rates while approach (ii) gives no indication as to how depreciation rates would be determined.

4.2 The Provision for Future Replacement or Sinking Fund Approach

> "The depreciation problem, in general terms, is the problem of writing off from fiscal period to fiscal period sums sufficient to return the capital invested in a property when that property has outlived its usefulness".
>
> *J.S. Taylor*[1923; 1010]

> "After the straight line formula the one perhaps most widely used is the sinking fund formula and the modifications of this method".
>
> *John B. Canning*[1929; 273]

As the above quotations indicate, some early statisticians and accountants viewed the depreciation problem as a method for funding the future replacement of a durable input. In this sinking fund method, the focus shifted from changes in asset values to the problem of setting aside period by period accounting charges into a fund which will cumulate over the useful life of a fixed asset into an amount which will be sufficient to replace the asset on its retirement date. This treatment of depreciation has some elements in common with the maintenance of physical capital approach to income determination to be discussed in section 7.4 below, but the two methods are distinct.

In the 1930's, the sinking fund approach to depreciation was successfully attacked by a number of accountants on the grounds that depreciation accounting should be viewed as a method of spreading the initial cost of a durable input over its useful life rather than as a solution to the logically distinct problem of deciding whether the asset should be replaced at the end of its life:

> "There is a much larger number who in some way try to identify the deposits to a sinking fund or the deposits and earnings thereon with 'depreciation expense'. The only source of an 'expense' of depreciation is the outlay or outlays made or agreed to be made for the asset in order to have the enjoyment of the service". *John B. Canning*[1929; 274]

> "Depreciation exists whether the property being used is to be replaced or not. In no sense does the depreciation allowance ('reserve') account represent the accumulation for the purpose of acquiring future assets".
>
> *M.B. Daniels*[1933; 306]

> "Generally it is conceded that it is the purpose of recording depreciation to recover the original expenditure, the purchase of a new truck being a separate and distinct transaction having no possible connection with, or relation to, the

recovery of the original investment. It is simple to point out that the truck at the end of six years may be replaced at a higher price or a lower price, or at the same price, or by teams and wagons, or that no replacement will be made if the need for a truck has disappeared. The above arguments seem sufficiently convincing to discredit this replacement theory in so far as it is related to fixed assets ". *Stephen Gilman*[1939; 349]

"The availability of money for replacement may offer serious financial problems. The problem of financing replacements may be sufficiently difficult to tax the resourcefulness and foresight of business men but it is in no sense whatever an accounting problem. The originally acquired asset was a deferred charge and its cost is recovered by the depreciation program. The replacement whether it be an identical item or not, is a fresh transaction resulting in the creation of a new deferred charge the cost of which in turn must, from the accounting viewpoint, be recovered over the years which follow its acquisition".

Stephen Gilman[1939; 494]

We turn now to accounting treatments of depreciation that take the intertemporal cost allocation viewpoint.

4.3 Straight line Depreciation and Other Accounting Allocation Methods

"A plough, for instance, which lasts twenty years, will contribute a twentieth part of its life-work and use to the ingathering of twenty different harvests".

Eugen von Böhm-Bawerk[1891; 305]

"*Straight Line Formula* ... In general, only two primary estimates require to be made, viz., scrap value at the end of n periods and the numerical value of n .... Obviously the number of periods of contemplated use of an asset can seldom be intelligently estimated without reference to the anticipated conditions of use. If the formula is to be respectable at all, the value of n must be the most probable number of periods that will yield the most economical use".

John B. Canning[1929; 265-6]

The first method that comes to mind in attempting to determine a sequence of depreciation rates of the form (3) for a durable capital input as it ages is the one suggested by Böhm-Bawerk above: estimate the expected number of accounting periods n that the input is likely to be used in production and assume that the single period depreciation rate is $\delta = 1/n$. This straight line method of depreciation can be used to allocate the initial purchase cost of the asset, say P^0, across the n periods of its life; these historical cost allocations under straight line depreciation would be $(1/n)P^0, (1/n)P^0, \ldots, (1/n)P^0$, a sequence of n equal allocations. The straight line depreciation method can also be used in conjunction with current values of new units of the asset, yielding the following sequence of current value depreciation charges: $(1/n)P^1, (1/n)P^2, \ldots, (1/n)P^n$, where P^t is the price of a new unit of the asset at the beginning of period t, for $t = 1, 2, \ldots, n$.

Another commonly used method for the determination of depreciation rates rests on the assumption that depreciation occurs on the undepreciated value of the asset at a constant geometric rate δ where $0 < \delta < 1$. The sequence of historical cost allocations of original cost P^0 that this method generates is $\delta P^0, \delta(1-\delta)P^0, \delta(1-\delta)^2 P^0, \delta(1-\delta)^3 P^0, \ldots$ while the corresponding stream of periodic current cost accounting charges is $\delta P^1, \delta(1-\delta)P^2, \delta(1-\delta)^2 P^3, \delta(1-\delta)^3 P^4, \ldots$. This method of depreciation is sometimes called the reducing balance method[39] or the declining balance method.[40] As we shall see in section 13 below, this method of accounting for depreciation (applied to current values) is very convenient when it is necessary to construct capital aggregates for productivity measurement purposes.[41] Empirical estimates for the declining balance depreciation parameter δ generally come from: (i) "official" estimates by broad asset class made by the national tax or regulatory authorities[42]; (ii) estimates made by the engineers or managers of the business unit[43], or (iii) statistical studies such as those to be discussed in sections 4.4 and 4.5 below.

Saliers [1922] and Canning [1929; 260-309] list many other rather arbitrary methods that accountants have used to estimate depreciation rates. The arbitrariness of these accounting depreciation methods and the fact that the estimates are generally based on a prior reasoning rather than on empirically observable declines in value[44] has of course attracted comment from many accountants and economists over the years:

> "Accountants immediately discard their own figures and demand an appraisal of the plant and other fixed assets, whenever they are called upon to compute capital value for the purpose of sale, reorganization etc. Apart from such occasions they adhere to their depreciation methods with the proviso that the method itself matters less than consistent adherence to it, once it has been adopted. These methods generally limit guessing to a minimum considered unavoidable in the circumstances". *Gabriel A.D. Preinreich*[1938; 240]

> "For the past hundred years accountants have been searching for the 'true' depreciation method which would allocate the cost of the machine over its lifetime in accordance with the rate at which it is actually being 'used' up. They have reluctantly concluded that there is no 'true' depreciation method, and that all

> the methods used or proposed are mere conventions, the choice between which is a matter of convenience". *F. Lutz and V. Lutz*[1951; 7]

However, historical cost accountants such as Daniels and Ijiri have defended the arbitrariness of accounting cost allocations as follows:

> "The function of depreciation is recognized by most accountants as the provision of a means for spreading equitably the cost of comparatively long lived assets. Thus, if a building will be of use during twenty years of operations, its cost should be recognized as operating expense, not of the first year, nor the last, but of all twenty years. Various methods may be proper in so allocating cost. The method used, however, is unimportant in this connection. The important matter is that at the time of abandonment the cost of the asset shall as nearly as possible have been charged off as expense, under some systematic method".
> *M.B. Daniels*[1933; 303]

> "However, there is a diametrically opposite problem in historical cost accounting ... the problem is one of disaggregation or allocation. Suppose that resources A and B are purchased together for \$20, but at the end of the year the firm had only Resource A. How much of the \$20 should be assigned to Resource A? Depreciation is a typical problem of this kind. However, accountants have devised many methods, however arbitrary they may be, by which such allocations are carried out objectively". *Yuji Ijiri*[1979; 67]

Both of the above authors recognize the arbitrariness of historical cost accounting allocations of asset cost; the best that can be said of these methods is that they are "systematic". If the tax authorities specify that one or more depreciation formulae must be used for tax purposes, then the use of the resulting historical cost allocations might also be characterized as "objective".

Since historical cost accountants regularly criticize current value accountants for their use of imputed or estimated values, it is important to recognize that historical cost accounting is subject to precisely the same criticism: historical cost accounting, by accepting an arbitrary a priori pattern of depreciation rates *imputes period by period depreciation costs*. If we attempt to estimate the period by period durable input costs accurring to a business unit, then *any* method of accounting will have to resort to imputed or estimated values.[45]

We conclude this section with two additional criticisms of historical cost depreciation allocations.

The first criticism is due to Canning[46] who asked that criteria be developed to choose among the many depreciation methods that were used by historical cost accountants.

This request for a rational criterion for choosing a depreciation formula has not been answered because the answer cannot be given on the basis of a priori reasoning: period by period empirical evaluation of the physical condition and market value of the assets to be depreciated is required.

The second criticism is due to Edwards and Bell[47] who noted that the historical cost accountant would need to be clairvoyant in order to determine the useful life of an asset; i.e., identical new assets are not all retired at the same time.[48] On the other hand, current value accounting techniques, by estimating period by period used asset values (recall formula (7) above), avoid in principle the difficult problems that arise when durable inputs are used at different intensities. We can paraphrase this second criticism of historical cost accounting techniques as follows: different historical cost accountants will estimate different lengths of life (and scrap values) for the same asset, leading to variable or "nonobjective" period by period depreciation estimates. Parker[49] in an ingenious empirical study showed that current value estimation of depreciation for a certain calculating machine gave rise to less dispersion in the estimates for depreciation expense than historical cost estimation. Thus current value accounting may be more "objective" than historical cost accounting.

We now turn to an examination of economic methods for estimating depreciation rates. These methods are based on the observation of current market values.

4.4 Economic Approaches Based on Used Asset Prices

> "If there is a perfect second hand market for the goods in question, so that a market value can be assessed for them with precision, corresponding to each particular degree of wear, then the value-loss due to consumption can be exactly measured ...". *John R.Hicks*[1939; 2nd edition 1946; 176]

> "Some depreciation patterns have very little economic justification (except accounting convenience), but most of them at least purport to approximate the decline in the economic value of the remaining services (i.e., market value). Of the various possible depreciation schemes (net stock measures), two measures seem to be of the most interest: (a) a net stock concept based on a purely physical deterioration depreciation scheme, and (b) the market value of the existing stock of capital. The latter figure can be approximated by the use of depreciation rates derived from studies of used machinery prices". *Zvi Griliches*[1963; 120]

Economists (like Hicks and Griliches) and accountants (like Bell and Edwards[50]) have long realized that a possible method of estimating the decline in value of a durable input due to its use over an accounting period is to use information on the market prices of used assets at a point in time and to compare differences in price as a function of the age of the input.

Let t index the beginning of a time period and suppose that we know the purchase price of a new durable input at the beginning of period t and we have collected information on the average market price at time t for the same input that has been used for n periods, P^{tn}, for $n = 1, 2, \ldots, N$.[51] Then the one period depreciation rate for a new asset could be defined by δ_0 in the following equation (provided that depreciation is independent of the time period t):[52]

$$(22) \quad P^{t1}/P^{t0} \equiv (1 - \delta_0).$$

Given that δ_0 has been determined, then the one period depreciation rate for a one period old asset can be defined by δ_1 in the following equation:

$$(23) \quad P^{t2}/P^{t0} = (1 - \delta_0)(1 - \delta_1).$$

In general, given that the one period depreciation rates $\delta_0, \delta_1, \ldots, \delta_{n-1}$ have been determined, then the one period depreciation rate δ_n for an asset that is n periods old at the start of an accounting period can be defined by δ_n in the following equation:

$$(24) \quad P^{tn+1}/P^{t0} = (1 - \delta_0)(1 - \delta_1) \ldots (1 - \delta_n), \quad n = 0, 1, \ldots, N - 1.$$

Thus given market data on the prices of the used asset, period by period depreciation rates δ_n can be obtained by solving equations (24).[53] If we have information on used asset prices for many different time periods t and we are willing to make the assumption that depreciation rates are stable over time, then a stochastic specification of a variant of (23) can be made and econometric techniques can be used to estimate the sequence of one period depreciation rates. If the market data on used asset prices is sparse, then instead of estimating a completely general pattern of period to period depreciation rates $\delta_0, \delta_1, \ldots, \delta_{N-1}$, various restrictions on these parameters can be imposed.[54] The simplest such restriction is that δ_n be constant from period to period; i.e.,

$$(25) \quad \delta_n = \delta_0 \text{ for } n = 1, 2, \ldots, N - 1$$

which is the geometric or declining balance depreciation model mentioned in section 4.3 above.

Empirical studies of depreciation rates using second hand asset prices have been made by the accountant Beidleman [1973] [1976][55] and the economists Hall [1971], Hulten and Wykoff [1981a] [1981b] and Oliner [1996]. The literature on this used asset approach is ably reviewed by Hulten and Wykoff [1996][56] and Jorgenson [1996c].

Many economists and accountants have objected to the use of second hand data to estimate depreciation rates for a variety of reasons:

> "We readily agree that where a market is sufficiently large, generally accessible, and continuous over time, it serves to coordinate a large number of subjective estimates and thus may impart a moment of (social) objectivity to value relations based on prices formed on it. But it can hardly be said that the second-hand market for industrial equipment, which would be the proper place for the determination of the value of capital goods which have been in use, satisfies these requirements, and that its valuations are superior to intra-enterprise valuation". *L.M. Lachmann*[1941; 376-377]

> "But why, if market values are the key to asset values, does not the accountant find depreciation by direct reference to market quotations for assets of different ages, and abandon his formulae? Various answers suggest themselves .... Second-hand markets tend to be small and scrubby, so that quotations may be hard to find and harder to trust. Many assets are built specifically for the one firm, and therefore worn replicas do not exist. An owner usually regards his own worn assets as different from, and better than, replicas in the market, because he knows their history, condition and foibles". *William T. Baxter*[1971; 31]

> "One argument, drawing on the Akerlof Lemons Model, is that assets resold in second hand markets are not representative of the underlying population of assets, because only poorer quality units are sold when used. Others express concerns about the thinness of resale markets, believing that it is sporadic in nature and is dominated by dealers who under-bid".
> *Charles R. Hulten and Frank C. Wykoff*[1996; 17-18]

Inspite of the above objections to the use of the second hand market method for estimating depreciation rates, this method seems more "objective" than simply guessing at the appropriate rates.

A more serious objection to the above model of depreciation rate determination is that the method includes only length of asset service or time in use as an explanatory variable and thus the method neglects variations in the intensity of use of the durable input.[57] There are at least two ways of meeting this criticism: (i) we can follow the advice of Edwards and Bell[58] and estimate separate sequences of depreciation rates that pertain to assets that are used with approximately the same intensity and have similar maintenance policies or (ii) we can incorporate utilization and maintenance variables as explanatory variables in stochastic versions of (24).[59] Jorgenson [1996c, 27-29] reviews the literature on these extensions of the basic used asset model of depreciation rate determination.

We turn now to another economic method for the determination of depreciation rates.

4.5 Economic Approaches Based on Production Function Estimation

Suppose a durable input is used by a business unit and it purchases I^t units of it at the beginning of period t in order to produce y^t units of output using the vector x^t of nondurable inputs[60] during period t as well as the services of past purchases of the durable input. Suppose further that the durable input lasts N periods and that after adjusting for physical loss of efficiency, all unretired units of the durable input are perfect substitutes in production. The production function which relates output flow to inputs used during a period is F and if there is no technological progress, we have the following relationship between output produced and inputs used during period t:

$$(26) \quad y^t = F(x^t, I^t + (1-\delta_0)I^{t-1} + (1-\delta_0)(1-\delta_1)I^{t-2} + ... + (1-\delta_0)...(1-\delta_{N-1})I^{t-N}).$$

The production function method for determining depreciation rates works as follows: (i) collect data on output produced during period t, nondurable inputs used x^t and durable input purchases I^t for a number of periods t; (ii) assume a functional form for the production function F; (iii) add a stochastic specification to equations (26) for $t = 0, 1, ..., T$ and (iv) use econometric techniques to simultaneously estimate the unknown parameters which appear in the production function F as well as the depreciation rates $\delta_0, \delta_1, ..., \delta_{N-1}$. Variants of this basic method include: (i) restricting the depreciation parameters $\delta_0, \delta_1, ..., \delta_{N-1}$ in some a priori fashion (e.g., recall (25) above); (ii) using the assumption of short run profit maximizing or cost minimizing behavior on the part of the business unit in order to add extra estimating equations involving period t prices to the single estimating equation (26) or (iii) instead of estimating the production function F, estimating the unknown parameters in the dual cost or profit function.[61]

Empirical studies using this approach to the estimation of depreciation rates include Epstein and Denny [1980], Pakes and Griliches [1984], Nadiri and Prucha [1996] and Doms [1996].

It should be noted that the depreciation rates which are estimated using this production function approach will, in general, be different from the estimates that result from the used asset approach studied in section 4.4 above. The latter approach incorporates the effects of both exhaustion and deterioration (to use Griliches' terminology), while the production function approach incorporates only the effects of physical deterioration.[62]

There are two major problems with the production function approach: (i) the approach will only work in a highly aggregated model with a small number of outputs, nondurable inputs and durable inputs due to the difficulties involved in estimating the parameters pertaining to a general production function when there are numerous inputs and outputs[63] and (ii) the assumption that the different vintages of capital can be combined together in the additive capital aggregate that appears as the last term on the right hand side of (26) is restrictive; i.e., it is assumed that the deterioration adjusted different vintages of capital are perfect substitutes in production, an assumption which may or may not be true.

Thus as a good general method for the empirical determination of depreciation rates, the production function method is unsatisfactory.

We turn now to our final class of economic methods for the determination of depreciation rates.

4.6 Economic Approaches Based on Rental Prices

Suppose that a rental market for a durable input exists so that we can observe in period 0 the rental price w^0 which appears in the user cost formulae (2) or (4). Then given information on the firm's cost of capital r^0 and the prices P^0 and P^1 of a "new" unit of the durable input being used by the firm at the beginning and end of period 0, we can use equation (4) to solve for the period 0 depreciation rate δ^0. Thus given that rental markets exist for durable inputs being used by a business unit, these rental prices can be equated to the corresponding user costs and depreciation rates can be derived from the resulting equations. Of course, if rental markets do not exist, then this method will not work, an obvious point made by Hulten and Wykoff [1996;15]. Jorgenson [1996c;32] reviewed the few studies that have used this method.

Even when rental markets exist, this method is unlikely to generate very accurate depreciation rates, particularly when prices are changing rapidly over the course of the accounting period. The problem is that rental prices are generally determined at the beginning of the period *before* the price of the used durable at the end of the period is known. Hence an observed rental price is actually an ex ante price of the form

$$\begin{aligned}\tilde{w}^0 \equiv\ & P^0 - \tilde{P}_u^1/(1+r^0) \\ & = [r^0 - \tilde{i}^0 + \delta^0(1+\tilde{i}^0)]P^0/(1+r^0)\end{aligned}$$

where $\tilde{P}_u^1$ is the (beginning of period 0) *anticipated* price of the used durable input (at the end of period 0) and $\tilde{i}^0$ is the *anticipated* period 0 inflation rate for "new" units of the durable input. Thus equating the right hand side of (28) to an observed rental price w^0 and solving for δ^0 is impossible unless we *know* the anticipated inflation rate $\tilde{i}^0$ at the beginning of period 0, an unlikely possibility.[64]

Even though the rental price method is unlikely to be a useful method for the empirical determination of depreciation rates, rental prices are useful when they exist–they can be used as period 0 (opportunity) costs for the use of the corresponding durable inputs during period 0.[65]

Our overall conclusion in evaluating different methods for the empirical determination of depreciation rates is that the used asset price method seems best when the relevant second hand markets actually exist. However, for firm specific assets that are not traded in second hand markets, it appears that depreciation will have to be determined on the basis of engineering estimates or appraisals.

It is evident from our discussion of depreciation that different authors have different concepts of depreciation. Goldberg summed up some of the alternative meanings as follows:

> "If, by 'charging depreciation' we mean an allocation of historical cost, I suggest that we use words (such as 'cost allocation' and 'proportion' of cost allocated against past revenue') which will convey this meaning. If we mean attempting to provide resources for future replacement of assets, why not use words (such as 'provision for future replacement') which will bring this meaning out? If we mean adjustment to present market costs, why not use words which say so; if we mean estimate of wearing out, let us indicate this clearly and unequivocally. To use a word like 'depreciation' or a phrase like 'provision for depreciation' which is now so confused is not quite fair to ourselves or to the readers of our statements and reports". *Louis Goldberg*[1955; 484]

We interpret "depreciation" to encompass exhaustion and deterioration over an accounting period but to exclude capital gains or losses due to changes in the prices of assets over the accounting period; i.e., depreciation is the hypothetical loss of value of an asset that would take place over an accounting period under "normal" use, holding the structure of new and used asset prices fixed.[66]

We turn now to a closer examination of the production function concept and the meaning of "production."

5.0 The Production Function Framework

> "Thus far, however, we have left out of consideration the fact that commodities are *products* which result from the combination of productive factors such as land, men and capital goods". *Léon Walras*[1954; 211]

> "Almost all of our theorizing about investment and the desired stock of capital rests implicitly on some technological considerations and is derived from some kind of general production function. As long as we stick to the production function framework, it is clear that *quantity* rather than *value* is the relevant dimension, since the production function is defined as a relationship between the quantity of output and the quantity of various inputs".
>
> *Zvi Griliches*[1963; 118]

In order to measure the contribution of capital to the production of outputs, it is useful to have an idealized model of how capital inputs interact with other flow inputs to produce outputs. The idealized models that economists utilize are based on *production functions*, or more specifically, on *production possibilities sets* which are technologically feasible sets of inputs and outputs that can be produced by a specified business unit in a specified time period. There are a number of different production function concepts that can be distinguished. Thus in section 5.1, we discuss the short run production function which distinguishes capital as an input at the beginning of the accounting period and (depreciated) capital as an output at the end of an accounting period. In section 5.2, we consider an intertemporal production function which relates inputs to outputs over many accounting periods. In this production function concept, the capital stocks the firm has available at the start of the first accounting period are distinguished as inputs and the (depreciated) capital stocks at the end of the last accounting period (when the assets of the firm are sold) are distinguished as outputs, but there is no apparent necessity to keep track of used capital inputs in intermediate accounting periods in this framework (unless they are sold before the final period). Purchases of new capital inputs over intermediate periods are distinguished in this framework. In section 5.2, we also attempt to reconcile this intertemporal production function concept with the one period "Austrian" production function concept in section 5.1. In section 5.3, we indicate how the usual one period production function that treats capital just as an input in each accounting period can be extracted from the Austrian production function framework.

In section 5.4, we address a very fundamental question in accounting theory; namely, is speculative activity productive? The answer to this question will determine the validity of including asset inflation rates in our user cost formulae.

Finally, in section 5.5, we discuss briefly what are the factors determining the production function for a firm or business unit.

5.1 The Austrian Production Function

> "We must look at the production process during a *period* of time, with a beginning and an end. It starts, at the commencement of the Period, with an Initial Capital Stock; to this there is applied a Flow Input of labour, and from it there emerges a Flow Output called Consumption; then there is a Closing Stock of Capital left over at the end. If Inputs are the things that are put in, the Outputs are the things that are got out, and the production of the Period is considered in isolation, then the Initial Capital Stock is an Input. A Stock Input to the the Flow Input of labour; and further (what is less well recognized in the tradition, but is equally clear when we are strict with translation), the Closing Capital Stock is an Output, a Stock Output to match the Flow Output of Consumption Goods. Both input and output have stock and flow components; capital appears both as input and as output". *John R. Hicks*[1961; 23]

> "The business firm can be viewed as a receptacle into which factors of production, or inputs, flow and out of which outputs flow...The total of the inputs with which the firm can work within the time period specified includes those inherited from the previous period and those acquired during the current period. The total of the outputs of the business firm in the same period includes the amounts of outputs currently sold and the amounts of inputs which are bequeathed to the firm in its succeeding period of activity".
> *Edgar O. Edwards and Philip W. Bell*[1961; 71-72]

Hicks [1961; 23] and Edwards and Bell [1961; 71-72] obviously had the same model of production in mind: in each accounting period, the business unit combines the capital stocks and goods in process that it has inherited from the previous period with "flow" inputs purchased in the current period (such as labour, materials, services and additional durable inputs) to produce current period "flow" outputs as well as end of the period depreciated capital stock components which are regarded as outputs from the perspective of the current period (but will be regarded as inputs from the perspective of the next period). We call this an Austrian model of production in honour of the Austrian economist Böhm-Bawerk [1891] who viewed production as an activity which used raw materials and labour to further process partly finished goods into finally demanded goods.[67]

It should be noted that the neo-Austrian model of Hicks [1973] is different from the model that we are describing in this section. Hicks [1973; 7] interpreted Böhm-Bawerk's production model as follows:

> "Like Böhm-Bawerk (or Hayek) I think of the general productive process as being composed of a number (presumably a large number) of *separable elementary processes* .... The elementary process, of the older Austrian theory, was of a simple, too simple, type. There was associated with a unit of output, forthcoming at a particular date, a sequence of units of input at particular previous dates. The sequence of inputs, and the single output, constituted the process".

Hicks [1973; 8] then asserted that Böhm-Bawerk's production model was inconsistent with the existence of fixed capital inputs:

> "For the only kind of capital-using production which will fit into the old Austrian scheme is production without fixed capital, production that uses working capital (or circulating capital) only. Fixed capital (plant and machinery) will not fit in. For fixed capital goods are 'durable-use goods'; their essential characteristic is that they contribute, not just to one unit of output, at one date, but to a sequence of units of output, at a sequence of dates".

However, Böhm-Bawerk [1891; 299-300] certainly mentioned various durable capital inputs such as "tools", "machines" and "agricultural implements" as inputs in his model of production; what he did not explain explicitly is how time and use (i.e., depreciation) would transform these inputs into less valuable outputs at the end of a production period. What Böhm-Bawerk emphasized was the transformation of partly finished goods into more valuable partly finished goods and final products.[68]

It will be useful in subsequent sections to develop some notation to describe the one period Austrian model of production of this section. We suppose that there are M durable inputs that the business unit is using at the beginning of period 0. These durable inputs include machines, transportation equipment, other equipment, computers, plant structures, office buildings, tools, office furnishings and furniture, etc. These fixed capital stock components are classified into discrete categories according to their age and other relevant physical characteristics. The list of durable inputs also includes circulating capital stock components: inventories of raw materials, finished goods and partly finished goods (goods in process). Finally, we include in the business unit's list of initial capital stock components any patents or other marketable knowledge products as well as any holdings of land or other natural resources that it might possess. We denote the business unit's beginning of period 0 holdings of durable capital inputs by the nonnegative vector $\bar{k}^0 \equiv [\bar{k}_1^0, \bar{k}_2^0, \ldots, \bar{k}_M^0]$ where $\bar{k}_m^0 \geq 0$ denotes the initial stock of durable input m for $m =$

$1, \ldots, M$. We also suppose that $P_m^0 \geq 0$ is the beginning of period 0 market opportunity cost[69] for a unit of durable input m for $m = 1, \ldots, M$ and the vector of these initial market values is $P^0 \equiv [P_1^0, P_2^0, \ldots, P_M^0]$.

Next, we suppose that there are N outputs or inputs that the business unit can sell or purchase in the marketplace during period 0. The vector of average market prices that the business unit faces for the N commodities in period 0 is $p^0 \equiv [p_1^0, p_2^0, \ldots, p_N^0]$ where $p_n^0 \geq 0$ is the average market price for commodity n in period $0, n = 1, \ldots, N$. The vector of *net outputs* that the business unit produces during period 0 is denoted by $y^0 \equiv [y_1^0, y_2^0, \ldots, y_N^0]$; if $y_n^0 > 0$, then y_n^0 units of commodity n are being produced by the business unit during period 0 while if $y_n^0 < 0$, then $-y_n^0 > 0$ units of commodity n are being used as inputs during period 0.[70] The list of period 0 "flow" inputs includes various types of labour services, including both establishment employees and contracted professional services, purchases of electricity, heating fuels and telecommunications services. In principle, the entire list of durable capital stock components could be included in the list of flow inputs, since the business unit could purchase additional units of capital during period 0 to add to its initial stocks. The list of "flow" outputs will include the usual outputs that the business unit produces, classified as finely as seems necessary for the purpose at hand.[70] In principle, all of the initial capital stock components held by the firm at the start of period 0 should appear in the list of flow outputs, since these stocks could be sold in the marketplace during period 0. Note that we are distinguishing as separate flow commodities *sales* of initial capital stock components from additional *purchases* of capital stock components during period 0 for two reasons: (i) the selling price of an asset will usually differ from the purchase price of a similar unit of the asset due to transactions costs and (ii) the technological impact of the sale of a fixed capital stock component is often quite different from the purchase of an additional unit due to internal transactions costs (such as installation and training costs for purchases and dismantling and renovation costs for sales). Thus the dimensionality of the "flow" commodity space N will generally be much greater than the dimensionality of the initial durable input "stock" space M.

Finally, at the end of period 0, the business unit will have at its disposal a vector $k^1 \equiv [k_1^1, k_2^1, \ldots, \ldots, k_M^1]$ of durable inputs which can be valued at the end of period 0 (or beginning of period 1) nonnegative market opportunity costs vector $P^1 \equiv [P_1^1, P_2^1, \ldots, P_M^1]$. The components of k^1 consist of depreciated units of the business unit's beginning of period 0 vector of capital stocks k^0 that were not sold during period 0 plus any additional units of capital that might have been purchased during period 0 .[71]

We now turn our attention to the definition of the period 0 Austrian production function or more generally, the period 0 Austrian production possibilities set for the business unit under consideration. Several definitions are possible.

The broadest definition for the period 0 production possibilities set S^0 is to define S^0 as the set of all technologically feasible vectors of the form (k^0, y^0, k^1) where $k^0 \geq 0_M$[72] is a nonnegative beginning of period 0 vector of capital inputs, y^0 is an N dimensional vector of period 0 net outputs that can be produced given k^0 and $k^1 \geq 0_M$ is a nonnegative vector of capital stocks that are left over at the end of period 0. In this broadest definition of the period 0 production possibilities set, we allow the business unit to choose its initial vector of capital stocks k^0. In our next definition of the period 0 production possibilities set, we restrict the business unit's choices for the initial capital stock vector k^0 to be an observed vector of capital stocks $\bar{k}^0$ for the business unit under consideration. In this case, the period 0 production possibilities set for the business unit can be described as a feasible set of net outputs and end of the period capital stocks $\{y^0, k^1\}$ that could be produced using the observed initial vector of capital stocks $\bar{k}^0$; i.e., the technology set now has the form[73]

$$(29) \quad \{(y^0, k^1) : (\bar{k}^0, y^0, k^1) \in S^0\}.$$

The production possibilities set defined by (29) is our most narrow definition for the period 0 feasible set of inputs and outputs for the business unit under consideration. A third possible definition for the period 0 production possibilities set might be the following one:

$$(30) \quad \{(k^0, y^0, k^1) : (k^0, y^0, k^1) \in S^0, k^0 \leq \bar{k}^0\};$$

i.e., in definition (30), we allow the business unit to dispose of units of its initial capital vector $\bar{k}^0$, so that the capital stock vector k^0 actually used in the period 0 production process is equal to or less than $\bar{k}^0$.

Note that the "narrower" or more restricted production possibilities sets defined by (29) and (30) are defined in terms of the "broader" or least restricted production possibilities set S^0 and the business unit's initial observed capital stock vector $\bar{k}^0$.

The business unit's competitive profit maximization problem that corresponds to the broadest definition of the period 0 production possibilities set S^0 can be formalized as:

$$(31) \quad max_{k^0,y^0,k^1}\{-P^0 \cdot k^0 + (1+r^0)^{-1} p^0 \cdot y^0 + (1+r^0)^{-1} P^1 \cdot k^1; (k^0, y^0, k^1) \in S^0\}$$

where $P^t \cdot k^t \equiv \Sigma_{m=1}^M P_m^t k_m^t$ for $t = 0, 1, p^0 \cdot y^0 \equiv \Sigma_{n=1}^N p_n^0 y_n^0, r^0$ is the period 0 interest rate or opportunity cost of capital and S^0 is the period 0 production possibilities set for the business unit. Note that we have divided the net "flow" revenues for period 0, $p^0 \cdot y^0$, and the market value of the business unit's end of the period holdings of capital stocks, $P^1 \cdot k^1$, by one plus the interest rate, $(1 + r^0)$. Thus we are assuming that period 0 "flow" revenues and costs $p^0 \cdot y^0$ are "realized" at the end of period 0 along with the end of period

0 value of the business unit's capital stocks, $P^1 \cdot k^1$. These end of period 0 capital stocks are discounted to make them equivalent to beginning of period 0 values, as is traditional in economics. However, from the perspective of accounting theory, it is more natural to express all values in terms of end of the period values and thus from this perspective, the business unit's period 0 profit maximization problem becomes:

$$(32) \quad max_{k^0,y^0,k^1}\{-(1+r^0)P^0 \cdot k^0 + p^0 \cdot y^0 + P^1 \cdot k^1 : (k^0, y^0, k^1) \in S^0\}.$$

Now let us shift our focus to the business unit's profit maximization problem in period 1. Let $S^1 \equiv \{(k^1, y^1, k^2)\}$ denote the business unit's unrestricted period 1 production possibilities set, which consists of feasible vectors of starting capital stocks k^1, period 1 "flow" inputs and outputs y^1 and end of period 1 finishing capital stock vectors k^2. If there is no technological progress or managerial improvement in the organization of production, S^1 will equal S^0; i.e., the period 1 and period 0 production possibilities sets will be the same.

The period 1 counterpart to the period 0 unrestricted profit maximization problem (32) is:

$$(33) \quad max_{k^1,y^1,k^2}\{-(1+r^1)P^1 \cdot k^1 + p^1 \cdot y^1 + P^2 \cdot k^2 : (k^1, y^1, k^2) \in S^1\}$$

where r^1 is the beginning of period 1 opportunity cost of capital; $P^1 \equiv (P_1^1, \ldots, P_M^1) \geq 0_M$ is the vector of beginning of period 1 opportunity costs for capital stock components; $P^2 \equiv (P_1^2, \ldots, P_M^2) \geq 0_M$ is the vector of end of period 1 opportunity costs for capital stock components; $p^1 \equiv (p_1^1, \ldots, p_N^1) \geq 0_N$ is the vector of period 1 average prices for units of outputs and inputs and $y^1 \equiv (y_1^1, \ldots, y_N^1)$ is a period 1 net output vector (positive components of y^1 denote outputs, negative components denote inputs).

Obviously, one period profit maximization problems that are analogous to (32) and (33) can be defined for each accounting period t that the business unit is in operation.

We can also define a one period profit maximization problem that has the same structure as (32) except that the restricted production possibilities set defined by (29) is used in place of S^0. This restricted period 0 profit maximization problem (with k^0 restricted to equal the fixed initial capital stock vector $\bar{k}^0$) is:

$$(34) \quad max_{y^0,k^1}\{-(1+r^0)P^0 \cdot \bar{k}^0 + p^0 \cdot y^0 + P^1 \cdot k^1 : (\bar{k}^0, y^0, k^1) \in S^0\}.$$

Suppose $\bar{y}^0$ and $\bar{k}^1$ solves (34). Then the end of period 0 capital stock vector $\bar{k}^1$ can serve as a vector of fixed starting capital stocks for the business unit's period 1 restricted profit maximization problem which is analogous to (33) except that k^1 is fixed at $\bar{k}^1$:

$$(35) \quad max_{y^1,k^2}\{-(1+r^1)P^1 \cdot \bar{k}^1 + p^1 \cdot y^1 + P^2 \cdot k^2 : (\bar{k}^1, y^1, k^2) \in S^1\}.$$

Finally, we can define a one period profit maximization problem that has the same structure as (32) except that the production possibilities set defined by (30) is used in place of S^0:

$$(36) \quad max_{k^0,y^0,k^1}\{-(1+r^0)P^0 \cdot k^0 + p^0 \cdot y^0 + P^1 \cdot k^1 : (k^0, y^0, k^1) \in S^0; k^0 \leq \bar{k}^0\}.$$

The differences between the three period 0 profit maximization problems (32), (34) and (36) can be explained as follows: in (32), the business unit is allowed to sell its initial holdings of capital (the components of the vector $\bar{k}^0$) or buy additional units of capital at the beginning of period 0 at the prices P^0; in (36), the business unit can only sell its initial holdings of capital $\bar{k}^0$ at the prices P^0; in (34), the business unit is stuck with its initial holdings of capital $\bar{k}^0$ at the beginning of period 0 and values these initial holdings at the prices P^0. Thus the different period 0 profit maximization problems reflect different assumptions about what options are open to the business unit at the beginning of period 0. However, for each of the three problems, the Austrian production possibilities set S^0 plays a crucial role.

In the following section, we no longer assume that the business unit's decision horizon is limited to a sequence of single periods; we will allow the business unit to make production plans that extend over a number of periods.

5.2 The Intertemporal Production Function

> "An option is any possible income stream open to an individual by utilizing his resources, capital, labor, land, money, to produce or secure said income stream. An investment opportunity is the opportunity to shift from one such option, or optional income stream, to another .... Some of the optional income streams, however, would never be chosen, because none of their respective present values could possibly be the maximum." *Irving Fisher*[1930; 151]

> "The problem of the firm, dynamically considered, is to find that stream of outputs, capable of being produced from the initial equipment, which shall have the maximum capital value .... If we write $x_{r0}, x_{r1}, x_{r2}, \ldots, x_{rv}$ for the [net] outputs of x_r planned to be sold in successive 'weeks' from the present, then the production function takes the form $f(x_{10}, x_{20}, \ldots, x_{n0}; x_{11}, x_{21}, \ldots, x_{n1}; x_{12}, x_{22}, \ldots, x_{n2}; \ldots; x_{1v}, x_{2v}, \ldots, x_{nv}) = 0$ assuming that the plan extends forward for v weeks. The capitalized value of the plan is $C = \Sigma_{r=1}^{n}\Sigma_{t=0}^{v}(\beta_t p_{rt} x_{rt})$ where $\beta_t = 1/(1+i_t)$, and i_t is the rate of interest per week for loans of t weeks;

p_{r0} is the current price of x_r and p_{rt} is the price the entrepreneur expects to rule in the week beginning t weeks hence". *John R. Hicks*[1946; 326]

In this section, we will utilize the intertemporal production function concepts developed by Fisher [1930; 151] and Hicks [1946; 136]. As in section 5.2, we assume that there are M types of durable capital equipment and that the business unit's initial holdings of capital stock components at the beginning of period 0 is $\bar{k}^0 \equiv [\bar{k}_1^0, \ldots, \bar{k}_M^0] \geq 0_M$. We now assume that the business unit's time horizon extends over T periods. Denote a vector of planned net outputs for period t by $y^t \equiv [y_1^t, \ldots, y_N^t]$ for $t = 0, 1, 2, \ldots, T-1$ and denote the corresponding vector of anticipated average prices for period t by $p^t \equiv [p_1^t, \ldots, p_N^t] \geq 0$ for $t = 0, 1, 2, \ldots, T-1$. At the end of period $T-1$ (or equivalently, at the beginning of period T), we assume that the business unit is sold[74]; i.e., the components of its beginning of period T capital stock vector $k^t \equiv [k_1^T, \ldots, k_M^T] \geq 0$ are sold at the prices $P^T \equiv [P_1^T, \ldots, P_M^T] \geq O_M$. The *intertemporal production possibilities set* $S \equiv \{(k^0, y^0, y^1, \ldots, y^{T-1}, k^T)\}$ is the feasible set of net output vectors $y^0, y^1, \ldots, y^{T-1}$ for periods $0, 1, \ldots, T-1$ and beginning of period T (depreciated) capital stock components k^T that can be produced by existing technology (and technology that can be anticipated to exist over the time horizon of the business unit) and an initial vector of capital stock inputs, k^0.

Let r^t be the interest rate or opportunity cost of financial capital that is relevant to the business unit at the beginning of period t for $t = 0, 1, \ldots, T-1$ and let $P^0 \equiv [P_1^0, \ldots, P_M^0] \geq 0_M$ be the vector of opportunity costs for capital stock components at the start of period 0. Then assuming that period t "flow" net revenues, $p^t \cdot y^t \equiv \Sigma_{n=1}^N p_n^t y_n^t$, are "realized" at the end of period t, the business unit's intertemporal planned profit maximization problem can be written as follows:

$$
\begin{aligned}
(37) \quad & max_{k^0, y^0, y^1, \ldots, y^{T-1}, k^T} \{-P^0 \cdot k^0 + (1+r^0)^{-1} p^0 \cdot y^0 + (1+r^0)^{-1}(1+r^1)^{-1} p^1 \cdot y^1 \\
& \qquad + \ldots + (1+r^0)^{-1} \ldots (1+r^{T-1})^{-1}(p^{T-1} \cdot y^{T-1} + P^T \cdot k^T) : \\
& \qquad (k^0, y^0, y^1, \ldots, y^{T-1}, k^T) \in S\}.
\end{aligned}
$$

Note that all values in the objective function of (37) that are realized after the beginning of period 0 are discounted by interest rate terms $(1 + r^t)$. Thus all values are expressed in beginning of period 0 equivalent values. Note also that the intertemporal profit maximization problem (37) reduces to the single period Austrian profit maximization problem (31) if the business unit's time horizon is only one period; i.e., if $T = 1$. Note also that in both (31) and (37), we allowed the initial vector of beginning of period 0 capital stocks k^0 to be variable. A counterpart to (37) which freezes k^0 to equal the business unit's historically determined capital stocks $\bar{k}^0$ is[75]:

$$
\begin{aligned}
(38)\qquad & max_{y^0,y^1,\ldots,y^{T-1},k^T}\{-P^0\cdot\bar{k}^0+(1+r^0)^{-1}p^0\cdot y^0+(1+r^0)^{-1}(1+r^1)^{-1}p^1\cdot y^1 \\
& \quad +\ldots+(1+r^0)^{-1}\ldots(1+r^{T-1})^{-1}(p^{T-1}\cdot y^{T-1}+P^T\cdot k^T): \\
& \quad (\bar{k}^0,y^0,y^1,\ldots,y^{T-1},k^T)\in S\}.
\end{aligned}
$$

If we divide the objective function in (34) by $(1+r^0)$, it can be seen that the resulting version of (34) is the same problem as (38) if $T=1$; i.e., the intertemporal profit maximization problem (38) is equivalent to our restricted one period Austrian profit maximization problem (34) when the business unit's time horizon is only a single period.

In the case where the business unit has a multiperiod planning horizon (i.e., the case where $T>1$), it is possible to relate the one period technology sets of section 5.2 to the intertemporal technology set S of the present section. Suppose that the one period Austrian technology sets $S^0,S^1,\ldots,S^{T-1}$ are given for periods $0,1,\ldots,T-1$. Then these Austrian technology sets can be used to define a Hicksian intertemporal technology set S as follows:

$$
\begin{aligned}
(39)\qquad & S\equiv\{(k^0,y^0,y^1,\ldots,y^{T-1},k^T):(k^0,y^0,k^1)\in S^0,(k^1,y^1,k^2)\in S^1,\ldots, \\
& \qquad (k^{T-1},y^{T-1},k^T)\in S^{T-1}\};
\end{aligned}
$$

i.e., in defining (39), we simply force the end of period t capital stocks k^{t+1} to be equal to the starting capital stocks for period $t+1$ for $t=0,1,\ldots,T-1$. Thus we do not allow the business unit to sell or purchase any units of capital at the very end of each period t in definition (39) for $t=0,1,\ldots,T-2$.[76]

We conclude section 5.3 by indicating that under certain conditions, solutions to the Hicksian intertemporal profit maximization problem (38) are also solutions to a sequence of Austrian single period profit maximization problems, provided that the period by period capital stock valuation vectors $P^1,P^2,\ldots,P^{T-1}$ that appear in the Austrian problems (but do not appear in (38)) are chosen appropriately. In order to minimize notational complexity, we will demonstrate the above assertion for the case of a two period intertemporal technology; i.e., we will assume $T=2$ in (38).[77]

Suppose that y^{0*},y^{1*},k^{2*} solves (38) when $T=2$. Under certain conditions, we can define end of period 0 or beginning of period 1 capital stock price and quantity vectors P^{1*} and k^{1*} such that: (i) y^{0*} and k^{1*} solve the period 0 profit maximization problem (34) provided that $P^1=P^{1*}$ and (ii) (k^{1*},y^{1*},k^{2*}) solves the period 1 profit maximization problem (33) provided that the P^1 which appears in (33) is equal to P^{1*}. The translation of the last rather technical sentence is this: period by period "Austrian" profit maximization can be consistent with the intertemporal Hicksian profit maximization model (38) provided that the correct "economic" capital stock prices P^{t*} are used in the single period profit maximization problems.

In order to establish the above assertion, it is necessary to introduce the period t variable profit function π^t that is dual to the Austrian technology set S^t for $t = 0, 1$:[78]

$$(40) \quad \pi^t(p^t, k^t, k^{t+1}) \equiv max_{y^t}\{p^t \cdot y^t : (k^t, y^t, k^{t+1}) \in S^t\}, t = 0, 1.$$

Using definition (40), the single period constrained profit maximization problem (34) can be rewritten as the following unconstrained profit maximization problem involving only the components of k^1:

$$(41) \quad max_{k^1}\{-(1 + r^0)P^0 \cdot \bar{k}^0 + \pi^0(p^0, \bar{k}^0, k^1) + P^1 \cdot k^1\}.$$

Suppose that $\hat{k}^1$ is a solution to (41) and each component of $\hat{k}^1$ is positive; i.e., $\hat{k}^1 >> 0_M$. If $\pi^0(p^0, \bar{k}^0, k^1)$ is differentiable with respect to the components of k^1 at $k^1 = \hat{k}^1$, then the vector of first order partial derivatives of π^0 with respect to the components of $k^1, \nabla_{k^1}\pi^0(p^0, \bar{k}^0, \hat{k}^1) \equiv [\partial\pi^0(p^0, \bar{k}^0, \hat{k}^1)/\partial k_1^1, \ldots, \partial\pi^0(p^0, \bar{k}^0, \hat{k}^1)/\partial k_M^1]$, will satisfy the following first order necessary conditions to solve (41):[79]

$$(42) \quad \nabla_{k^1}\pi^0(p^0, \bar{k}^0, \hat{k}^1) + P^1 = 0_M.$$

Now use definition (40) for $t = 1$ and rewrite the period 1 constrained profit maximization problem (33) as the following unconstrained profit maximization problem involving the vector of beginning of period 1 capital stocks k^1 and the vector of end of period 1 capital stocks k^2:

$$(43) \quad max_{k^1,k^2}\{-(1 + r^1)P^1 \cdot k^1 + \pi^1(p^1, k^1, k^2) + P^2 \cdot k^2\}.$$

Suppose that $\tilde{k}^1 >> 0_M$ and $\tilde{k}^2 >> 0_M$ are solution vectors for (43) and that $\pi^1(p^1, k^1, k^2)$ is differentiable with respect to the components of k^1 and k^2 at $(k^1, k^2) = (\tilde{k}^1, \tilde{k}^2)$. Then the vector of first order partial derivatives of π^1 with respect to the components of $k^1, \nabla_{k^1}\pi^1(p^1, \tilde{k}^1, \tilde{k}^2)$, and the vector of first order partial derivatives of π^1 with respect to the components of $k^2, \nabla_{k^2}\pi^1(p^1, \tilde{k}^1, \tilde{k}^2)$, will satisfy the following first order necessary conditions to solve (43):[80]

$$(44) \quad \nabla_{k^2}\pi^1(p^1, \tilde{k}^1, \tilde{k}^2) + P^2 = 0_M;$$

$$(45) \quad \nabla_{k^1}\pi^1(p^1, \tilde{k}^1, \tilde{k}^2) - (1 + r^1)P^1 = 0_M.$$

Now assume that the intertemporal production possibilities set S is constructed using the one period technology sets S^0 and S^1 and definition (39) when $T = 2$. Using definitions (40), we can rewrite the constrained intertemporal profit maximization problem (38) (when $T = 2$) as the following unconstrained profit maximization problem involving the beginning and end of period 1 capital stock vectors k^1 and k^2 as decision variables:[81]

$$(46)\quad max_{k^1,k^2}\{-P^0 \cdot \bar{k}^0 + (1+r^0)^{-1}\pi^0(p^0,\bar{k}^0,k^1) + (1+r^0)^{-1}(1+r^1)^{-1}\pi^1(p^1,k^1,k^2) + (1+r^0)^{-1}(1+r^1)^{-1}P^2 \cdot k^2\}.$$

Assume that $k^{1*} >> 0_M$ and $k^{2*} >> 0_M$ solves (46) and that π^0 and π^1 are differentiable with respect to the components of k^1 and k^2 when $(k^1, k^2) = (k^{1*}, k^{2*})$. Then k^{1*} and k^{2*} will satisfy the following first order necessary conditions for solving (46):

$$(47)\quad (1+r^0)^{-1} \bigtriangledown_{k^1} \pi^0(p^0,\bar{k}^0,k^{1*}) + (1+r^0)^{-1}(1+r^1)^{-1} \bigtriangledown_{k^1} \pi^1(p^1,k^{1*},k^{2*}) = 0_M;$$

$$(48)\quad (1+r^0)^{-1}(1+r^1)^{-1} \bigtriangledown_{k^2} \pi^1(p^1,k^{1*},k^{2*}) + (1+r^0)^{-1}(1+r^1)^{-1}P^2 = 0_M.$$

We use the vector of partial derivatives $\bigtriangledown_{k^1}\pi^0(p^0,\bar{k}^0,k^{1*})$ in order to define a vector of end of period 0 "economic values" or shadow prices of capital P^{1*} :[82]

$$(49)\quad P^{1*} \equiv - \bigtriangledown_{k^1} \pi^0(p^0,\bar{k}^0,k^{1*}).$$

Now suppose that the P^1 in (41) and (42) is equal to P^{1*}. Since k^{1*} satisfies (49), then k^{1*} will also satisfy (42) if the P^1 in (42) is replaced by P^{1*} (and $\hat{k}^1$ is replaced by k^{1*}). Thus k^{1*} will solve the period 0 Austrian profit maximization problem (41) if the vector of shadow prices P^{1*} defined by (49) is used as the end of period 0 capital stock price vector P^1 in (41). In this case, the one period maximization of profits coincides with the business unit's long run intertemporal maximization of profits.

Now substitute the definition (49) of the shadow prices P^{1*} into the first order conditions (47). The resulting equation is (after simplification):

$$(50)\quad -(1+r^1)P^{1*} + \bigtriangledown_{k^1}\pi^1(p^1,k^{1*},k^{2*}) = 0_M,$$

which is (45) if we replace P^1 by P^{1*} (and $\tilde{k}^1, \tilde{k}^2$ are replaced by k^{1*}, k^{2*}). Finally, after simplification, the first order conditions (48) are equivalent to the first order conditions (44) (where $\tilde{k}^1, \tilde{k}^2$ are replaced by k^{1*}, k^{2*}). Thus k^{1*}, k^{2*} will solve the period 1 Austrian profit maximization problem (43), provided that the beginning of period 1 price vector for capital stock components P^1 is replaced by the vector of shadow prices P^{1*} defined by (49).

The thrust of the above algebra is this: under some regularity conditions,[83] single period Austrian profit maximization is perfectly consistent with the long run intertemporal maximization of profits, provided that the business unit uses "economic" prices to value its end of period capital stock components. The problem with this result is that it is usually extremely difficult or impossible to determine these economic prices as outside observers of the business unit (or even as insiders); i.e., at the end of period 0, how can we determine P^{1*} defined by (49)? The difficulties involved in the practical determination of "economic" prices explain why most accountants dismiss the use of "economic values" as a practical alternative for the valuation of a business unit's end of the period capital stocks. We will discuss alternative methods of valuation in more detail in section 6 below.

Traditional production function models do not distinguish capital as an input at the beginning of a period and capital as an output at the end of the same period as was done in the Austrian production function. In the following section, we indicate how a traditional production function can be derived from an Austrian production function.

5.3 The Traditional Production Function

> "I belong to the party which is still looking to find, at the end of its journey, a rehabilitation of the so-called 'Production Function' $P = f(L, C)$ [where P is product, L is labour input and C is capital input] in some form or other; what I am looking for is a concept of capital which will ultimately allow us to think, more or less, in those terms". *John R. Hicks*[1961; 18]

> "In the context of the Hicksian model, it is clear that we can construct several capital aggregates that must be carefully distinguished: (a) a current period capital stock aggregate (an input from the viewpoint of the current period) using current period capital stock prices as weights in the aggregation procedure; (b) a (depreciated) following period capital stock aggregate (an output from the viewpoint of the current period) using discounted expected following period capital stock prices as weights; (c) a current period investment aggregate (an output) using current period investment goods prices as weights in the aggregation procedure; and (d) a capital aggregate that is an aggregate of (a) and (b) where capital as an input and capital as an output are oppositely signed in the index number formula that is used". *W. Erwin Diewert*[1980b; 474-475]

We return to the Austrian model of section 5.1 and note that there is an easy way of simplifying the model so that we do not have to distinguish each durable commodity as both an input and an output: we need only use Leontief's [1936; 54-57] Aggregation Theorem. This result says that if two commodities are always used or produced in fixed proportions by a production unit in each period t, then the two commodities can be aggregated into a single composite commodity. More specifically, let x_1^t and x_2^t denote the quantities of say two inputs used during period t and let $p_1^t \geq 0$ and $p_2^t \geq 0$ denote the period t average price for each commodity. If $x_1^t = \alpha x_2^t$ for all periods t under consideration, then the two commodities can be aggregated into a composite commodity with period t

aggregate input X^t equal to the quantity of input 1 during period t, x_1^t and with period t composite price P^t equal to the period t value of the two commodities divided by x_1^t; i.e.,

(51) $\quad X^t \equiv x_1^t;\ P^t \equiv [p_1^t x_1^t + p_2^t x_2^t]/x_1^t = p_1^t + \alpha p_2^t.$

Definitions (51) can still be used to aggregate the two commodities even if say commodity 1 is an input and commodity 2 is an output; in this case, x_1^t and x_2^t have opposite signs and α, the factor of proportionality, is negative.

Now consider the case of a single durable input that lasts 2 or more periods and whose productivity declines only with age (and not use). Suppose that $k^0 > 0$ units of the (new) durable input were purchased at the start of period 0 at price $P^0 > 0$, and suppose that the end of period 0 price for depreciated units is $P_u^1 >> 0$. Then from the perspective of the end of period 0, the net cost of using k^0 units of the durable input during period 0 is

(52) $\quad (1+r^0)P^0k^0 - P_u^1k^0 = w_e^0k^0$

where w_e^0 is the end of period 0 user cost defined by (7) above. Now define $x_1^t \equiv k^t, x_2^t \equiv -k^t, p_1^t \equiv (1+r^t)P^t$ and $p_2^t \equiv P_u^t$ for $t = 0, 1, \ldots$ and apply Leontief's Aggregation Theorem. It can be seen that $X^t = k^t$ and $P^t = w_e^t$ for $t = 0, 1, \ldots$; i.e., the ex post end of period 0 user cost w_e^0 can be viewed as the period 0 price for the use of one unit of an aggregate of capital where the two capitals are capital input at the beginning of period 0 and capital output at the end of period 0. The resulting aggregate capital can be viewed as the capital input which appears in a "traditional" production function and a user cost is the price which is associated with the capital aggregate.

Obviously, the above aggregation technique can be applied to all vintages of a capital input provided that declines in value over the period are independent of use. If declines in value are not independent of use, then we need to distinguish different end of period prices that depend on the intensity of use of the durable input over the accounting period. For example, suppose that k_L^0 units of our new durable good in the previous paragraph were used "lightly" during period 0 and that k_H^0 units were used "heavily" (or more intensively), where $k^0 = k_L^0 + k_H^0$. Suppose that the end of period 0 price for a unit that is used lightly is P_{uL}^1 and for a unit that is used heavily is P_{uH}^1. We can define period 0 Leontief aggregate prices for new units of capital that are lightly and heavily used, w_{eL}^0 and w_{eH}^0 respectively, as follows:

(53) $\quad w_{eL}^0 \equiv (1+r^0)P^0 - P_{uL}^1;\ w_{eH}^0 \equiv (1+r^0)P^0 - P_{uH}^1.$

The Leontief aggregate quantities that are matched up with the above user costs are k_L^0 and k_H^0 respectively. Of course, the disaggregation of each type of beginning of the period capital input into separate categories depending on period 0 use can be carried out as finely as seems empirically necessary.[84] Thus Leontief's Aggregation Theorem can be applied to aggregate capital inputs in an Austrian production function even if the value of the assets declines with use as well as with age.

The above aggregation technique will not work for assets that lose their identity during the period 0 production process; e.g., a computer chip on hand at the beginning of the period emerges as part of a computer at the end of the period[85] or a concrete foundation at the beginning of the period becomes part of a building at the end of the period, etc.

We turn now to a discussion of what seems to be a rather philosophic question: what exactly is productive activity? It turns out that answers to this question are crucial in determining how to define business income.

5.4 The Three Basic Forms of Productive Activity

> "Production, in the narrow sense, changes the form and nature of products. Trade and transport change their external relations".
>
> *Alfred Marshall*[1920; 64]

> "If the difference of the *place* at which goods are available is a sound economic reason for exchanging fungible goods that are in other respects entirely similar, and if the advantage and convenience of the present *place* may justify the claim and allowance of a premium, just as much may the difference of the *time* at which similar goods are available be a sound reason for their exchange, and a guarantee that there will be premium on the – more valuable – present goods. This premium, and nothing else, is Interest".
>
> *Eugen von Böhm-Bawerk*[1891; 295]

Over the years, three general forms of productive activity have been identified: (i) the *transformation* of less valuable commodities into more valuable commodities; (ii) the *transportation* of commodities from one location to a more valuable location, and (iii) the *storage* or *holding* of goods from one time period to a future time period when they will be more valuable.

The first two types of productive activity are not controversial and were identified long ago by Marshall [1920; 64] in 1890; the transformation activity is regarded as being the most fundamental type of productive activity by economists and accountants.[86] Böhm-Bawerk [1891; 295] noticed the analogy between transportation and storage activities and introduced the third type of productive activity into the literature. Over the years, many accountants (such as Schmidt,[87] Edwards and Bell[88] and Chambers[89]) and economists (such as Lerner[90] and Debreu[91]) have argued that the third type of productive activity is just as valuable as the other two types. However, when the third type of activity is labelled as "speculative activity" instead of "storage activity", many economists and accountants have objected to treating the third type of productive activity in exactly the same manner as the first two types. Specifically, these economists and accountants argue against the inclusion of capital gains on assets held by a business unit over an accounting period. These objections will show up in section 6 below when we consider various approaches to the valuation of end of the period capital stocks. However, in the present section, we shall present some of the arguments that have been advanced by economists and accountants to *deny* treating speculative gains or capital gains on assets held through an accounting period as net revenues or components of business income in a manner that is symmetric to the treatment of transformation and transportation activities. We shall consider four types of objection.

The first objection states that capital gains on assets held through the accounting period are capital losses to someone else and hence there is no net gain to the community. Consider the following quotations by the accountants Schmidt and Crandell:

> "Only in one case can appreciation be real profit to the business man, viz., when he uses money credit to buy goods for speculation outside of his regular business needs. If his selling prices thereafter are higher than the money lent plus interest and costs after selling the goods, the difference will be his realized speculative gain. This kind of profit is especially high in times of rising general price levels. But this kind of private profit is no profit to the community, because the lender of money loses the same buying power on his money that the borrower gains". *Fritz Schmidt*[1931; 291]

> "What treatment should be accorded the speculative gains and losses realized from trading among individuals in securities? It is obvious that these sorts of transactions cannot increase the national wealth, hence the national income cannot be affected thereby. Whatever one gains the other loses".
> *William T. Crandell*[1935; 399]

The argument that the capital gains made by one business unit must be offset by capital losses made by some other consumer or business unit does not seem to be correct. Consider the case of a one person economy that controls a single business unit. Any capital gains made by the business unit that result from an optimal intertemporal allocation of resources are not offset by capital losses.[92]

The second objection to the inclusion of capital gains in income is more subtle: speculative holding activities do not enhance the productive powers of the economy and hence any increase in revenues resulting from these activities should not be recognized as a benefit to the economy; in fact, focusing on speculative gains may be bad for the economy because it will cause managers to not focus on the first two types of productive activities[93] (transformation and transportation). Consider the following quotations which are representative of this point of view:

> "Some theoretical explanation of the reasons why appreciation cannot be profit is needed at this point. For this purpose we must consider the enterprise as a part of the national production machine. It will then be clear that a maintenance of total productivity as of a certain moment will only be possible, if the productive instrumentalities of all individual enterprises concerned are preserved intact. The maintenance of productive power as a whole is not possible if accounting is based on an original value basis. The reason is that pure appreciation would then appear as profit whenever a change of value has taken place between the purchase and selling dates for the materials and wages that compose a product". *Fritz Schmidt*[1931; 289]

> "The appreciation in value of capital assets and land must not be treated as an element in national income. Depreciation due to physical wear and tear and obsolescence must be treated as a charge against current income, but not the depreciation of the money value of an asset which has remained physically unchanged. Appreciation and depreciation of capital were included in the American statistics of national income prior to 1929, but now virtually the same convention has been adopted in all countries". *Colin Clark*[1940; 31]

> "Enhancement of asset values as a result of increased market prices does not, without realization of such appreciation through sale, constitute a basis for recognition of revenue to the business enterprise. However, the realization of gain on the sale of a capital asset does not necessarily imply any contribution by the seller to the social product during the period of realization. Because such gains are irrelevant to production of the period, capital gains (and losses) are excluded from calculations of national income and product. It is seen then that, whereas standards of accounting for revenue provide for recognition of capital gains once they have been realized, such gains find no place at all in the accounting for the economy". *Gilbert P. Maynard*[1952; 190]

> "The essence of the difference between financial capital maintenance and all concepts of physical capital maintenance is in the treatment of the effects of price changes while assets are held. Under financial capital maintenance, all such effects are included in income .... Under physical capital maintenance,

the effects of price changes are excluded from income on the grounds that, if positive, they do not enable an enterprise to increase its operating capability or, if negative, they do not force a reduction". *Bryan Carsberg*[1982; 62]

The above authors have implicitly ruled out storage and holding activities as being productive like transformation and transportation activities. However, we can follow Böhm-Bawerk and argue that holding activities are completely analogous to transportation activities. Since transportation activities are regarded as being productive, so should holding activities.

The third objection to the inclusion of capital gains in the period by period income statements of a business unit has been made by accountants and it is an objection only to the inclusion of unrealized capital gains (i.e., no *sale* of the asset which has experienced a capital gain over the accounting period has been made) in income, not to the inclusion of realized capital gains (i.e., the appreciating asset has been sold during the accounting period). The objection is that unrealized capital gains should not be included in the period's income due to their *hypothetical* and *unverifiable* nature. Consider the following quotations:

"*Appreciation, Capital Gains and Losses.* A part of the ultimate net income of an enterprise can be assigned in some cases to natural growth and other increases in value. In the case of timber tracts, orchards and similar properties, natural increase, commonly called *accretion*, is an important factor in financial history. In other cases enhancement of property values due to changing business and general economic conditions, a general rise in the price level, or other factors which result in an increase in effective value over actual cost, usually referred to as *appreciation*, are of marked significance. At what point, in the succession of events that lead to final fruition of these gains in cash, should the accountant recognize the change? Eventually, if no cognizance is taken of it before, the gain will be realized in cash when the property itself, or the product resulting from its use, is sold. Until such time as the gain is validated by sale, the increased value is commonly characterized as unrealized and the gain as 'unearned' or 'unrealized income"'. *William T. Crandell*[1935; 389]

"The various codifications of accounting doctrine during the past two decades have been in general agreement that revenue should be recognized in the accounts only when certain tests of realization have been met. In the vast majority of cases, revenue realization is marked by a discrete event, that of sale and delivery of goods or services. Thus accountants draw a distinction between the earning or accrual of revenue throughout the productive processes and the realization of revenue, giving recognition in the accounts only to the latter….

The national income accountant is concerned with the *creation* of product, not alone with its subsequent sale". *Gilbert P. Maynard*[1952;189][94]

"There is another important respect in which business and social accounting differ which is worthy of comment here. Although business accountants are fully aware of the tentative nature of their measurements of income, they place great emphasis upon the objectivity and verifiability of the business data to which they grant recognition in the accounts". *Gilbert P. Maynard*[1952; 193]

The accountant's objection to the inclusion of *unrealized* capital gains as a contribution to the income of an accounting period due to their hypothetical nature is a valid one. However, the traditional accounting solution to the unrealized capital gains problem is to assume that no capital gains occur in any accounting period unless a realization occurs in some period (i.e., the asset is sold) in which case, all of the capital gains that accrued over the many accounting periods that the asset was held are imputed to the period of sale. This historical cost treatment of capital gains can create tremendous distortions (particularly in inflationary environments) to both the periodic income statements and balance sheets of the business unit.[95] Thus the accountant's treatment of unrealized capital gains (i.e., to exclude them from the income statement) is just as hypothetical (and more misleading in an inflationary environment) as including them in periodic income.[96] However, the historical cost accountant's objection to the hypothetical nature of period by period valuations of the capital stock components held by the business unit could be used to justify a separate treatment of unrealized capital gains on income statements rather than simply lumping them in with the more objective (transformation and transportation) sources of income.[97]

A fourth argument against the inclusion of capital gains in income statements runs as follows: for most businesses, capital gains or losses are an unintended consequence of their normal productive activities and moreover, in the long run, these gains and losses will tend to cancel. Hence it is not worth the bother of including these gains and losses as income, particularly when income may be taxed and hence a large unrealized capital gain may lead to a large tax bill which in turn may lead to a curtailment of the firm's normal productive activities. However, this line of thought led to a difficulty: what if the normal activity of a business unit was speculative (e.g., a commodities trader or a land speculator)? These business units would seem to be excluded from paying any income taxes on their earnings from speculative activities. To get around this difficulty, Plehn and other economists introduced the concept of *recurrence* of income:

"Income is essentially wealth available for recurrent consumption recurrently (or periodically) received. Its three essential characteristics are: *receipt, recurrence, and expendability*". *Carl C. Plehn*[1924; 5]

"It will, I think be readily admitted that those particular gains and profits which are recurrent, expendable receipts are the ones about whose income character there is seldom any doubt. Thus the gains and profits of a merchant are his income. The possible or even probable irregularity or uncertainty which distinguishes them from some other incomes does not seem to militate against their inclusion in income, provided they are *expected* to be recurrent. The same is true of the gains and profits of dealers in capital assets, for the lands, stocks and bonds, houses and the like are their stock in trade .... But it is when gains and profits lack one or two of the three characteristics of income, or have them in less than complete form, that a question arises. The one that is most often lacking is recurrence. Thus gains and profits from transactions outside of one's regular vocation or line of business, like the profit from the sale of a home, are of doubtful income character". *Carl C. Plehn*[1924; 10]

"The British income tax places very heavy stress upon the annual character of income. For an explanation of this conception, which results in the exclusion from taxable income of gains of an irregular nature, one must go back as far as the fifteenth century, when, with an agricultural society where few fortuitous gains developed, the idea of receipts as being annual in character became deeply impressed upon the minds of the people. It became the habit to think of one's regular receipts as his income, and to consider irregular receipts as additions to capital". *Robert Murray Haig*[1921; reprinted 1959; 69]

Thus if a business unit *regularly* makes profits on its speculative activities, the resulting profits are regarded as income but any capital gains on *occasional* speculative activities are not regarded as income according to the recurrence criterion for income. Of course, the problem with this concept is that it is difficult to draw the *boundaries* of recurrence:

"When is income recurrent? Professor Plehn expressly says it need not be perfectly regular. But how irregular can it be and still be 'recurrent'? The big profit on the sale of an old homestead may well occur twice in a lifetime. Does it not then 'recur'? If we extend the picture through two or more lifetimes 'recurrence' becomes altogether likely. In the case of corporations whose life goes on indefinitely every windfall, or extraordinary profit, may some day be duplicated. Evidently the 'recurrency' concept turns out to be too elusive to pass muster as a basis for analysis". *Irving Fisher*[1924; 666]

More fundamentally, if capital gains are regarded as being valid additions to income in some contexts, then why should they be excluded in other contexts?[98] When we view speculative activity as being analogous to transportation activities, it is obvious that a strong case can be made for including it as a valid form of productive activity.

We turn now to our final topic in the general area of production theory.

5.5 Where does the Production Function Come From?

> "The theory of the firm subsumed in general equilibrium theory is incredibly lean. It is summarized in the representation of a firm as a subset of the commodity space (preferably convex, closed, etc.). The simplicity of this theory derives from the activity analysis model of production which describes a firm as a given technology–a book of proprietary information about feasible transformations of inputs into outputs. A technology can evolve over time but it is not a matter of choice; indeed, a basic premise of general equilibrium theory is that information is not a decision variable". *Robert Wilson*[1975; 184]

> "An invention is viewed as a new process of production, or as a new vector of input-output coefficients". *William D. Nordhaus*[1969; 19]

Up to now, we have assumed that the single period production functions of section 5.1 or the intertemporal production function of section 5.2 are exogenously determined and known to the business unit. We think of the corresponding production possibilities sets as a given set of plans or operating procedures that are known to the management of the production unit. But where does this knowledge of the production possibilities set come from? How does this knowledge expand over time; i.e., how does innovation occur? These are the questions that we wish to address in this section.[99]

Knowledge of the set of feasible input and output combinations that a business unit in a specific geographic location[100] could use and produce during an accounting period comes from at least three sources: (i) operating manuals or other written (or computer accessible) materials that are available in the establishment; (ii) knowledge of production techniques that is embodied in employees and managers who work in the establishment and (iii) knowledge that is embedded in establishment machines. This provides a brief answer to the first question above. Our answer to the second question will not be so brief.

How does knowledge of new techniques of production (process innovations) and of new products (product innovations) get created? Traditional production theory as is embedded in general equilibrium theory[101] is silent on this point (even though many economists have noted that knowledge creation cannot be regarded as exogenous[102] and critics[103] have noted this deficiency of traditional production theory).

Obviously, specialized schools, universities and publicly supported research labs are a primary source of the creation of new knowledge but a considerable amount of innovative activity is undertaken by individual inventors and the research departments of private firms.

Arrow[104] and others[105] have attributed increases in productivity (more output for the same amount of input) to experience or the incidental effect of new investments. Arrow [1962; 155-157] explains his theory of innovation as follows:

> "I would like to suggest here an endogenous theory of the changes in knowledge which underlie intertemporal and international shifts in production functions. The acquisition of knowledge is what is usually termed 'learning' and we might perhaps pick up some clues from the many psychologists who have studied this phenomenon .... I advance the hypothesis here that technical change in general can be ascribed to experience, that it is the very activity of production which gives rise to problems for which favorable responses are selected over time .... The first question is that of choosing the economic variable which represents 'experience' ... I therefore take instead cumulative gross investment (cumulative production of capital goods) as a index of experience".

A somewhat similar theory of innovation was advanced by Allen [1983] which he called collective invention.[106] Allen explained his theory as follows:

> "Thus, if a firm constructed a new plant of novel design and that plant proved to have lower costs than other plants, these facts were made available to other firms in the industry and to potential entrants. The next firm constructing a new plant could build on the experience of the first by introducing and extending the design change that had proved profitable .... Collective invention was thus like modern research and development in that firms (and not individual inventors) generated the new technical knowledge. However, collective invention differs from R & D since the firms did not allocate resources to invention–the new technical knowledge was a by-product of normal business operation–and the technical information produced was exploited by agents other than the firms that discovered it". *Robert C. Allen*[1983; 2]

> "As long as the rate of investment was high, the rate of experimentation and the discovery of new technical knowledge was also high. On the other hand, if the rate of investment fell for any reason, the rates of experimentation and invention fell with it". *Robert C. Allen*[1983; 3]

Allen illustrated his theory using data on changes in the height and operating temperatures of blast furnaces in England between 1850 and 1875 and he summarized his results as follows:

> "Increasing furnace height and blast temperature led to lower fuel consumption and costs. The first firms to build tall furnaces might have treated this knowledge as a trade secret, but they did not. This information was made available to other parties through two channels–informal disclosure and publication in the engineering literature". *Robert C. Allen*[1983; 6-7]

Thus Allen modelled innovation as follows: as firms invested in new facilities, bolder firms undertook marginal changes in the design of their facilities or machines; successful design changes were then communicated to the industry as a whole through trade associations or formal publication in journals or magazines. It is interesting to note that Marshall advanced similar ideas many years ago.[107]

The next batch of theories of innovation date back to the origins of economics.

Adam Smith [1963; 8] observed that many inventions or innovations are made by workers who simply figure out better ways of accomplishing a task that they are presently engaged in:

> "I shall only observe, therefore, that the invention of all those machines by which labour is so much facilitated and abridged, seems to have been originally owing to the division of labour. Men are much more likely to discover easier and readier methods of attaining any object when the whole attention of their minds is directed towards that single object, than when it is dissipated among a great variety of things. But in consequence of the division of labour, the whole of every man's attention comes naturally to be directed towards some one very simple object. It is naturally to be expected, therefore, that some one or other of those who are employed in each particular branch of labour should soon find out easier and readier methods of performing their own particular work, whenever the nature of it admits of such improvement".[108]

Smith also observed that many improvements in productivity result from the specialization of labour: a worker who is able to concentrate or specialize on one task will become more proficient at that single task due to: (i) improvements in dexterity or physical skill and (ii) the elimination of the fixed costs in going from one type of task to another:

> "This great increase of the quantity of work, which, in consequence of the division of labour, the same number of people are capable of performing, is owing to three different circumstances; first, to the increase of dexterity in every

> particular workman; secondly, to the saving of the time which is commonly lost in passing from one species of work to another; and lastly, to the invention of a great number of machines which facilitate and abridge labour, and enable one man to do the work of many". *Adam Smith*[1963; 7]

Note that Smith suggested a third productivity benefit due to the increased specialization of labour: specialized routine operations by workers lend themselves to replacement by more efficient machines. Marshall[109] and Young[110] made similar observations. These observations are still valid today; e.g., many clerical and lower level managerial jobs are being replaced by computers.[111]

Smith [1963; 14] also pointed out that the division of labour was limited by the extent of the market; i.e., as the scale of the establishment grows due to the growth of markets for its outputs, the possibility of using specialized labour (and capital!) inputs also grows. As a corollary to his general principle, Smith pointed out that cities had larger markets than small towns and hence would support a higher degree of specialization in labour markets:

> "There are some sorts of industry, even of the lowest kind, which can be carried on no where but in a great town. A porter, for example, can find employment and subsistence in no other place. A village is by much too narrow a sphere for him; even an ordinary market town is scarce large enough to afford him constant occupation". *Adam Smith*[1963; 14]

Alfred Marshall further refined Adam Smith's idea that increases in the scale of an enterprise would generally lead to more efficient production by introducing the ideas of internal and external economies of scale:

> "We may divide the economies arising from an increase in the scale of production of any kind of goods, into two classes–firstly, those dependent on the general development of the industry; and, secondly, those dependent on the resources of the individual houses of business engaged in it, on their organization and the efficiency of their management. We may call the former *external economies*, and the latter *internal economies*". *Alfred Marshall*[1920; 266]

Internal economies of scale occur if output expansion leads to a less than proportional increase in the use of inputs; i.e., internal economies are equivalent to increasing returns to scale in more modern language. The increasing returns to scale phenomenon could be regarded as meaning that the production possibilities set of an establishment has a particular shape and hence it might appear that the increasing returns to scale phenomenon can be accommodated by traditional production theory. This is true once a business unit has actually run an establishment at a higher scale and has demonstrated that the technology works at the higher output levels, but the *first* successful demonstration of operating a technology at a higher scale has much the same character as establishing the feasibility of an innovation.[112]

There appear to be four main sources of internal economies of scale:

(i) *Indivisibilities*; i.e., most labour and capital inputs cannot be purchased in fractional amounts and all capital inputs have upper and lower limits on their capacities.[113]

(ii) *The Laws of Physics*; i.e., Kaldor[114] (and Marshall[115]) noted that the three dimensional nature of space leads to certain economies of scale.

(iii) *The Existence of Fixed Costs*; i.e., these are the efficiencies which result from averaging or amortising fixed costs over higher output levels. Before a machine yields a benefit from its operation, it may require the services of an operator who may have to be transported from one location to another[116] and the machine may require a warming up period before production can begin. These are examples of fixed costs whose effect becomes relatively smaller the greater the length of time that the machine is continuously operated.

(iv) *The Law of Large Numbers*; i.e., these are efficiencies that result from the laws of probability theory. For example, consider a power plant that uses a number of identical engines. If the probabilities of engine failure are independently distributed, then having one set of spare parts on hand will generally be sufficient whether the plant has one engine or ten engines. Similarly, a large bank will not require as high a proportion of cash reserves to meet random demands as a small bank.[117] In a similar vein, a large property insurance company whose risks are geographically diversified faces a smaller probability of bankruptcy than a small insurance company, etc.

We note that operations research and management science have developed mathematical techniques which enable the business unit to achieve internal economies with respect to the factors listed in (i), (iii) and (iv) above.

Two examples of Marshall's external economies of scale are: (i) reduced prices for inputs due to bulk purchasing[118] and (ii) the large scale of a business unit translates into a large demand for inputs and this in turn can encourage specialized suppliers to come into existence.[119] Thus external economies of scale reflect favorable changes in the environment facing the expanding business unit (lower input prices and new intermediate input suppliers).

What is the underlying cause of both internal and external economies? It seems that Adam Smith [1963; 14] had the answer to this question: growth of the market.

Some of the obvious factors that facilitate growth of the market are: (i) transportation improvements[120]; (ii) population growth[121]; (iii) reduction in trade barriers[122]; (iv) improvements in advertising[123]; (v) improvements in communications[124]; and (vi) growth of physical and human capital which leads to income growth which in turn leads to a growth in effective demand.

Perhaps the most interesting aspect of the creation of new knowledge about new production techniques is not the initial creation of the new knowledge but its diffusion to the local establishment level. The fact that a new product or production process has been developed somewhere in the world is of little significance to a local establishment that could use the innovation if the original knowledge is not transmitted or diffused to the establishment. Some of the factors that facilitate the rapid diffusion of new (and old) knowledge into a local market area are: (i) access to public libraries and university libraries[125]; (ii) access to newspapers, periodicals, journals, magazines, how to do it books, etc.[126]; (iii) memberships in trade associations, industry associations, professional societies, etc.;[127] (iv) access to international meetings and trade fairs where knowledge can be transmitted on a face to face basis[128] (adequate local transportation infrastructure will facilitate this access[129]); (v) access to good schooling and specialized training programs[130]; (vi) access to specialized consulting services and product information and (vii) access to telecommunications services[131] (i.e., having good local telecommunications infrastructure). The point that we are trying to make here is that a local market area does not necessarily have to devote a high percentage of its resources to primary research and development (i.e., to the creation of new products and processes): it need only have easy access to the sources of new knowledge.

We can summarize the results of this section as follows: (i) Traditional production theory that regards the feasible set of inputs and outputs for a business unit as an exogenous set is a serious oversimplification of reality. (ii) Current operating techniques can be regarded as exogenous, but the search for new products and new processes is an endogenous activity of the business unit that leads to an expansion of its single period production possibilities set. (iii) Exactly how this endogenous expansion takes place has not yet been adequately modeled, but it seems that the process of creating an innovation involves fixed costs and uncertain benefits.[132] (iv) It is difficult to separate a shift in the establishment production function (due to innovative activity) from a movement along a production function (due to a change in scale or to a change in input prices[133]); any new choice of inputs and outputs by a firm during an accounting period may involve innovative activity or it may simply involve a routine response to a change in the market environment.

We turn now to the discussion of a much narrower topic that will help us to evaluate the period by period performance of a business unit.

6.0 Alternative Approaches to Asset Valuation

"There are four possible bases which might be adopted: (1) liquidation value, or that value which is likely to be realized if the assets were thrown onto the market in the process of an orderly or forced liquidation; (2) original cost with proper allowance for decline in value of current assets and allowances for depreciation and depletion of fixed assets; (3) capitalized income producing value; (4) present market price of replacing or reproducing a similar asset in its present state of condition". *H.C. Daines*[1929; 98]

"Regarded from the more technical viewpoint of accounting, the problems faced and solved by German accountants during the period of absurd price fluctuations are quite worthy of study by accountants of all other countries. For the problems associated with keeping and interpreting financial records that must be expressed in a monetary unit oscillating even hourly are problems that the rest of the world must face to a less extreme degree".

Henry W. Sweeney[1927; 180-181]

In this section, we return to the fundamental problem of accounting: how to determine end of the accounting period prices for durable assets that are held by the business unit for multiple accounting periods. As can be seen from the above quotation by the accountant Daines, there are many possible methods for asset valuation that could be used. We shall consider six methods: (i) historical cost valuations; (ii) general purchasing power adjusted historical costs; (iii) net realizable values or appraisal values; (iv) replacement costs; (v) future discounted cash flows and (vi) asset specific index number adjusted historical cost. The method of valuation that is in general use today by accountants is the first method: historical cost accounting. However, the quotation by Sweeney should alert us to the problems associated with this method when there is general inflation in the economy. The other five methods attempt to deal with the valuation problem when there is general price instability.

6.1 Historical Cost Valuation

"Its greatest advantage is the fact that an original cost method is most easily subject to objective verification; it is the easiest to use in practice".

H.C. Daines[1929; 98]

"Today's dollar is, then, a totally different unit from the dollar of 1897. As the general price level fluctuates, the dollar is bound to become a unit of different magnitude. To mix these units is like mixing inches and centimeters or measuring a field with a rubber tape-line".

Livingston Middleditch[1918; 114-115]

We have already discussed how historical cost depreciation (i.e., decline in asset value over an accounting period) is determined in section 4.3 above: once a useful life for an asset has been estimated and a corresponding depreciation schedule has been determined, the initial purchase cost of the asset is allocated across accounting periods as a sum of periodic depreciation allowances. The corresponding historical cost value of the asset at the end of an intermediate accounting period is simply the initial purchase cost less the accumulated depreciation allowances over prior periods.

The main problem with historical cost valuation of assets shows up if there is a large change in the price of the asset (due to general inflation for example) from the time of its purchase to the end of the current accounting period: the historical cost valuation may bear no resemblance at all to a current market valuation for the asset. Thus in an inflationary situation, historical cost depreciation allowances will be understated, income will be overstated and income taxes may become capital taxes.[134] The problem is that historical cost accounting implicitly assumes that monetary values at the end of an accounting period are comparable to monetary values at the beginning of the accounting period; i.e., there is an implicit assumption of price level stability. The accountant Middleditch [1918] challenged this implicit assumption, having observed the tremendous inflation that occurred during World War I.[135]

There are two main virtues that are claimed for historical cost accounting: (i) it is *objective* and (ii) it is *conservative*. Both of these virtues are subject to criticism. We have already seen in section 4.3 that the determination of the length of life of an asset and the determination of the associated depreciation rates are far from being objective. Moreover, even if "economic" depreciation rates were used in order to allocate the initial purchase cost of an asset over its useful life, the resulting historical cost end of period values will be completely meaningless in a high inflation environment; i.e., they will not reflect current opportunity costs or market values. Thus historical cost accounting values might be *objective* but *irrelevant*.[136] *Conservatism*, on the other hand, conflicts with *accuracy*; i.e., if we wanted to be super conservative, why not assume all intermediate asset values are zero? The absurdity of this statement should make us realize that accuracy is a much more important virtue than conservatism.[137]

It is perhaps useful to elaborate a bit more on the meaning of "accuracy" in the context of determining period by period values for the assets of a business unit. It seems clear that there cannot be an answer to the problem of constructing intermediate values that is as unambiguous as the actual selling price of an asset; i.e., we can only make estimates of these intermediate values. Thus it seems reasonable to follow the example of Morgenstern [1963; 77] and regard these estimated intermediate values as probability distributions. "Accuracy" in this context could be defined as providing a suitable measure of central tendency (e.g., a mean valuation) along with a measure of dispersion (e.g., a variance). Unfortunately, accounting theory (and practice) has not proceeded along these lines[138], although occasionally, accountants recognize that introducing statistical concepts into accounting would be useful.[139]

We turn now to a discussion of other methods for valuing assets on a periodic basis– methods that will more closely approximate current market values or opportunity costs.

6.2 Purchasing Power Adjusted Historical Cost

> "It is obvious, therefore, that if quantities, whether measured in pounds or bushels or dollars, are to be correctly combined or compared, the unit of measurement must be homogeneous .... Yet many men who are not measuring their heights with fluctuating rulers, and who would throw verbal stones at such a silly doing, are complacently living in a similar kind of glass house, a business structure where in the substance of value continues to be measured by a dollar of seriously fluctuating size". *Henry W. Sweeney*[1936; reissued 1964; 11]

> "Professor Baxter [1976] has characterized the development of Latin American inflation accounting systems as having two stages: firstly, fixed assets and depreciation are adjusted by reference to a general index, and, secondly, at a later stage, the 'time-log' error on stocks [inventories] and monetary working capital is corrected by the application of an index".
> *David Tweedie and Geoffrey Whittington*[1984; 243]

This method of constructing a current value at the end of an accounting period originates with Middleditch [1918][140] and works as follows. Suppose an asset was purchased at the beginning of accounting period 0 at the price P^0, the period 0 depreciation rate is δ^0 and a general rate of price inflation over period 0 is ρ^0 (i.e., the general price level at the end of the period divided by the general price level at the beginning of the period is $1+\rho^0$). Then the historical cost accounting value of the asset at the end of the period is $(1-\delta^0)P^0$ but the *GPLA (General Price Level Adjusted)* value is

$$(54) \quad (1-\delta^0)(1+\rho^0)P^0.$$

The advantage of this method for constructing current asset values on a period by period basis is its relative *simplicity* (adjusted historical cost values at the beginning of the period need only be inflated by the common indexation factor $1 + \rho^0$) and its *objectivity* (once the appropriate indexation factor $1 + \rho^0$ has been chosen).[141]

In response to rapid inflation or a hyperinflation, GPLA accounting is the main form of current value accounting that has been used historically.[142]

Note the difference between ρ^0, a general inflation rate, and the asset specific inflation rate i^0 defined earlier by (5). In general, ρ^0 will not equal i^0 and hence the GPLA value for the asset will not equal its end of period selling price replacement value. This is the weakness of General Price Level Adjusted accounting. However, its strength is that it will adjust for the effects of general inflation.

The remaining topic to be discussed is how to choose the general inflation rate ρ^0.[143]

One of the simplest choices is to use the inflation rate for a widely traded commodity (such as gold[144]) as the index of general inflation. Another alternative is to use the rate of increase in the exchange rate of the country where the business unit is (primarily) located against a stable currency.[145] Instead of using the price of gold or any single commodity as the indicator of inflation, the general inflation between the beginning and the end of the accounting period might be better captured by looking at the price change of a "representative" basket of goods. As a further refinement, we could replace a fixed basket price index by a more general price index such as the Fisher [1922] ideal price index, which allows for substitution in response to price changes.[146]

Accountants and economists have struggled with the problem of choosing an appropriate price index to represent inflation for approximately a century. Many of the problems have still not been resolved[147]: (i) Which commodities should be included in the index? (ii) How should the individual price ratios be weighted[148]; i.e., what is the theoretically correct functional form for the price index? (iii) If the accounting period is shorter than a year, how can we deal with seasonal commodities that might be present in the index?[149] (iv) If we decide to use a consumer price index as a measure of general inflation, what is the appropriate class of consumers whose preferences should be reflected in the price index; i.e., should we construct a consumer price index that pertains only to investors in the business unit (some of whom may reside in foreign countries[150]) or should we broaden the class of consumers to be included in the index to include employees, creditors and potential investors?[151]

Even though the above questions are difficult to answer, we agree with Staubus that adjusting historical costs for general inflation by an imperfect index will generally be an improvement over historical cost accounting:

> "The argument that the corporate accountant cannot use the different purchasing power indexes of each individual shareholder must be read as either a weak excuse for inaction or an insistence on a degree of perfection that accountants have not reached in the past and are not likely to reach in the future.

Surely a broadly based price index provides a better measure of the change in the measuring unit than the assumption that there is no change at all, as the millions of people who base contracts on such indexes recognize".

George J. Staubus[1975; 44-45]

It is sometimes asserted that General Price Level Accounting adds no additional information over that which is available from reading historical cost accounting balance sheets[152]; i.e., if investors know historical cost values and they can look up the relevant general inflation index, then they can readily calculate the adjusted asset values defined by (54). This would be true if the business unit made the following information available to investors in each accounting period: (i) the value of new investments made in each period and (ii) the historical cost residual value of all assets that are sold or retired during the accounting period. In general, this information is not provided in balance sheets; hence providing investors with an aggregate GPLA asset value will provide new information that could not be calculated by individual investors.

6.3 Net Realizable Values (Exit Values)

"Some economists, notably Professor Jacob Viner of the University of Chicago, hold the belief that the value which the assets would bring in the market is the only proper basis of value for use in accounting". *H.C. Daines*[1929; 98]

"These markets [for assets] can be divided into two kinds, the markets in which the firm could buy the asset in its specified form and at the specified time and the markets in which the firm could sell the asset in its specified form and at the specified time. The prices obtained in markets of the first group we shall call *entry prices*; the prices obtained in markets in the second group we shall call *exit prices*". *Edgar O. Edwards and Philip W. Bell*[1961; 75]

In our earlier discussion of depreciation, we saw that a century ago, it was not unusual for accountants to value the fixed assets of a business unit at the end of an accounting period by appraised values; i.e., estimates of the net realizable values that the assets would bring in the market at the moment in time. However, during the first 35 years of the current century, many business firms arbitrarily revalued their fixed assets to suit their immediate purposes.[153] By the 1930's, the accounting profession reacted against these abuses by adopting the historical cost accounting methodology for valuing assets, and the accounting profession as a whole has stuck to this position since that time (except when an economy experienced very rapid inflation in which case General Price Level Adjusted accounting has been temporarily adopted). However, most economists and some accountants, such as Sweeney [1936; 44-53], Staubus[154], Edwards and Bell [1961], Chambers[155] and Sterling[156], have advocated the use of current values to value assets at the end of each accounting period.

The basic problem with the use of current values is that it is difficult to determine exactly what is the "correct" concept for a current value. Edwards and Bell [1961; 75] distinguish between an *entry value* (the minimum cost of purchasing a replacement for a currently held asset) and an *exit value* (the maximum price a currently held asset could be sold for in the market less the transactions costs of the sale; i.e., the net realizable value for the asset).[157] In this section, we will focus on the problems associated with the use of exit values and we will deal with entry values in the next section.

Historical cost accountants have two principle objections to the use of (imputed) net realizable values to value assets held by a business unit at the end of an accounting period: (i) they are not *objective* and (ii) they are not *additive*.

On the lack of objectivity of net realizable values, consider the following quotations:

> "Which alternative should be used as a basis? The highest, or the lowest, or an average? How should the search area, to get offers or find prices, be determined?". *Yuji Ijiri*[1979; 66]

> " 'Forced liquidation value' is also ill defined, but it sometimes seems to mean the price that could be obtained by selling to the first man on the street that one happened to meet. If this is the meaning, then we agree that it would be absurd to report such values. A less radical notion of immediate exit price is obviously called for". *Robert R. Sterling*[1970; 328]

Thus to find an estimated net realizable value for an asset, it is necessary to determine what is the appropriate set of potential buyers and how their price bids could be elicited. If instead of seeking prices from potential buyers of the asset, we resort to appraisal values for the asset, we again encounter a certain lack of determinancy: how many appraisals should be made; what are the credentials of the appraisers; what criteria do the appraisers use[158]; etc.

Rather than saying that hypothetical net realizable values or appraised values are not objective[159], it might be more accurate to say that they do not pass the *reproducibility test*; i.e., two accountants attempting to construct net realizable values for a firm's assets would not generally come up with the same values. This is the major advantage of historical cost accounting and general price level adjusted accounting; aside from the major problems involved in defining asset lives and depreciation rates[160], these two methods of accounting can claim that they pass the reproducibility test.

Turning now to the lack of additivity of net realizable values, consider the following quotation:

> "The second factor which makes current cost income more disputable than historic cost income is the non-additivity of current costs. The historical cost of Resource A and Resource B is by definition the sum of the historical cost of Resource A and Resource B .... This addivity does not exist in current cost valuation, insofar as the price of a resource is not necessarily equal to the sum of the prices of its components. If the current cost of Resource A is \$20 and Resource B is \$30 but that of A and B together is \$60, should we use \$50 or \$60 as the current cost of Resource A and Resource B?". *Yuji Ijiri*[1979; 67]

Thus if we have two assets that can be combined together to produce an extraordinary revenue stream (e.g., a machine and a building to house the machine that together produce a new product with a high profit margin), then the joint asset may have a net realizable value that is much greater than the sum of the separated net realizable values; i.e., net realizable values for assets are not necessarily additive.[161]

In order to overcome the lack of additivity of net realizable values, it will be necessary to make some rather arbitrary judgments. For example, current values could be obtained for each asset that was purchased separately (or for each group of assets that was purchased jointly) on a stand alone basis; e.g., if a tractor were purchased with several supplementary attachments, then we could attempt to find a net realizable value for the entire asset package. Thus the additivity problem is "solved" by restricting the collection of net realizable values to the asset combinations that were actually purchased by the business unit.[162]

To overcome the lack of reproducibility objection to the use of net realizable values is a bigger task and might involve considerable costs.[163] Accounting standards organizations or the government (in its role as a collector of business income taxes) would have to specify acceptable methods for constructing net realizable values. One possible (partial) solution might be to utilize appraised values for property insurance purposes. Insurance companies have an incentive to insure property up to its maximum value to the business unit (if premium revenue is proportional to insured value) but they also have an interest in not allowing overinsurance (in order to minimize carelessness and fraud on the part of the insured business unit). Another possible solution to the lack of reproducibility problem would be for the Accounting Standards Board or the Government to develop appraisal criteria and to train and license appraisers.

We leave the final words on possible methods for the objective or reproducible determination of net realizable values to Chambers:

> "We will take a more or less common sense view–namely that a statement of financial position as at a date will include singular statements, in respect of plant assets, which are indicative of one or more of the following: the cost at that date of acquiring plant in the condition in which it then stands, the valuation which a lender might place on it as a security for a loan, the valuation which the owner might place on it for insurance purposes, or the price which might be obtained for it if it were decided to change the character of the company's investments. Anyone is at liberty to contend that these would all be different; but they have one thing in common, they are all estimates made in the context of conditions operating about the time at which the financial statements are prepared. They are approximations to contemporary value in the market".
>
> *Raymond J. Chambers*[1964; 270]

We turn now to a discussion of entry values.

6.4 Replacement Costs (Entry Values)

> "The replacement cost is the sum of money which would have to be expended at the present time to reproduce a physical property identical with that in existence at the present time and used for the benefit of the public".
>
> *Hammond V. Hayes*[1913; 618]

> "The values which the accountant uses in closing the books and preparing statements ideally should be based upon economic conditions at the moment of closing. If plant and equipment assets were valued at the close of each period on the basis of costs of replacement–effective current costs–depreciation changes would be increased in a period of rising prices and the other concomitant effects would be registered in the accounts in a rational manner".
>
> *William A. Paton*[1920; 6-7]

The description of an entry price or replacement value of an asset has already been provided in section 6.3 above: it is the current market cost of purchasing a physically identical replacement for an asset currently being held by a business unit. As can be seen from the above quotations, the concept of a replacement value dates back at least 80 years.

Replacement cost as a basis for asset valuation grew in popularity during the 1920's due to the inflationary upheavals that took place at that time and in the prior decade:

> "In Germany, during the severe inflation period, the orthodox practice of calculating depreciation on the basis of original book costs was eventually swept aside because accountants and business men came to perceive that, in maintaining the substance of capital, it was no longer useful. At first various supplementary measures were adopted, such as charging all new fixed asset costs to expense and creating a special reserve to provide for maintenance of plant value and business efficiency (e.g., the prevalent Werkerhaltungskonto). Later, computation of depreciation on the basis of reproductive cost grew in popularity, which, indeed, is still evident from a survey of contemporary German depreciation theory". *Henry W. Sweeney*[1931; 166]

> "Prices go up and prices go down, and with each change in the price level the discussion of replacement cost usage recurs. It appears that businessmen and accountants were willing to experiment with the use of replacement cost in the 1920's and early 1930's. But this receptivity to its use has declined steadily since then: in the 1940's practicing accountants were opposed to its use; .... Thus if past experience holds true for the future, replacement cost will still receive its share of attention from theoreticians while practicing accountants largely ignore it". *Germain Boer*[1966; 97]

Even though replacement cost accounting is no longer used by business accountants, it should be noted that it is still used today by national income accountants (or social accountants as they are sometimes called) as the basis for computing depreciation on a current cost basis.

The net realizable value and replacement cost of an asset can be regarded as the selling and buying prices for the asset in the relevant second hand market. Replacement cost will generally exceed the corresponding net realizable value due to the existence of transactions costs.

There is a variant of replacement cost accounting that at first sight seems to eliminate the need to consider second hand markets: find a current purchase price for a new asset that corresponds to the used asset on hand, apply the same method of depreciation to this new asset price (instead of the original historical cost price for the asset) and the resulting depreciated current price is an estimate for replacement value. However, this method of constructing replacement values implicitly assumes that the business unit is using the "correct" depreciation rates (the "correct" rates are reflected in current used asset markets).

Replacment cost can exceed the corresponding net realizable value for reasons other than transactions costs. Consider the following example due to Paton:

> "One example will be sufficient to show the ruinous error which may flow from a slavish adherence to the cost-of-replacement theory in appriasals. In 1924, a valuation was made of the properties of the Kansas City Railways by two independent engineers. One of the items to be appraised was three old engines in the power house. These were of the massive type, with enormous flywheels, and were standard equipment twenty or twenty-five years ago, or more. This equipment was in excellent physical shape, but was utterly obsolete, and a couple of the engines were no longer even connected. The company's power at the time of appraisal was entirely supplied by other and more modern equipment, although the old units were capable of giving service if required. One of the engineers went to the Westinghouse Company, with complete specifications, and secured an estimate of what it would actually cost, as of the date of the appraisal, to construct these engineers, on special order. He then made an estimate of the cost of shipment, installation, etc. The result was a cost of replacement figure considerably over a million dollars. The other engineer treated the units as scrap and gave them a net value of $20,000". *William A. Paton*[1931; 95]

What happend in the above example is that technical progress occurred which caused the net realiable value for the used asset to plummet, but the replacement value for the asset was high, since the old asset was no longer being produced.

There is a logical difficulty associated with the use of replacement cost values for unique assets such as a specially constructed machine or an engineering structure that is specific to the business unit: no replacement cost values are available in the marketplace for unique assets.[164] A solution to this difficulty is inherent in the approach of the first engineer in the above example: simply calculate the estimated cost of building the specific asset using the technology and input prices that pertain to the end of the accounting period.[165]

Replacement cost values are subject to the same two difficulties that were associated with the use of net realizable values: replacement costs are not generally *reproducible* (different accountants will generally obtain different estimates of replacement cost) and replacement costs are not generally *additive* (if a group of assets is replaced, the aggregate replacement cost may be less than the sum of the individual replacement costs). The lack of additivity is not a serious problem: we can again impose additivity by seeking replacement costs for assets according to how they were originally purchased; i.e., if a group of assets were jointly purchased, then we attempt to find a joint replacment cost for the same group of assets. However, the lack of reproducibility is a serious limitation on the use of replacement values.[166]

In this section, we considered the use of replacement costs and in section 6.3, we considered the use of net realizable values as a basis for valuing the assets held by a firm at the end of an accounting period. Is there a rational basis for choosing between these alternative valuation methods? One way of answering this question is to consider whether the business unit is likely to *buy* additional units of the asset in the near future (in which case an appropriate opportunity cost would appear to be *replacement cost*) or whether the business unit is likely to *sell* the asset in question (in which case the relevant opportunity cost would appear to be *net realizable value*). Thus several accountants[167] have argued for the use of replacement values for raw material inventories and for net realizable values for inventories of finished products. Following this same logic, an expanding firm might value its fixed capital stock components at replacement values while a contracting firm might use net realizable values. While this line of reasoning does not provide a complete answer to the question of which valuation base to use, it does seem helpful.

We turn now to a brief discussion of yet another basis for interim asset valuations.

6.5 Future Discounted Cash Flows

> "The flow of services issuing from an article of capital may have any duration and any distribution of rate. In every case the capital value of the article is the discounted value of its anticipated services". *Irving Fisher*[1897; 527]

> "If one could approximate the whole future series of money outgoes and of money receipts of an enterprise, one could find, given a rate of discount, a direct capital value of that enterprise". *John B. Canning*[1929; 207]

The view that the appropriate value for an asset is the discounted stream of the future net revenues that can be attributed to it was actively advocated by Irving Fisher [1897] [1930]. In the accounting literature, estimating a current asset value as the discounted stream of its future expected returns is known as the *economic approach* to asset valuation. Of course, a current purchase price for an asset can be thought of as representing a lower bound to the asset's economic value to the purchaser, but in this section, we will define an asset's economic value as an estimated discounted stream of net returns that can be attributed to the asset.

Accountants pointed out that the economic approach to asset valuation suffers from two fatal flows: (i) future discounted net returns are generally not known with any degree of certainty and hence the resulting estimates will not be *reliable*[168] and (ii) even if we did know future revenue flows with certainty, revenue flows are produced by the joint efforts of all assets and it is generally impossible to allocate the resulting joint net revenue flows to individual assets.[169] Another way of phrasing the first objection is to say that economic values will not generally pass the *reproducibility test*; i.e., different accountants will generally obtain different estimates for economic values. In principle, the second objection to the economic approach can be overcome–the economic values or shadow prices defined by (49) in section 5.3 above provide the "correct" allocations of future discounted expected profits to the end of period 0 capital stock components held by the business unit. However, it is not a straight-forward matter to compute these shadow prices–an econometric estimate for the firm's Austrian variable profit function would have to be obtained and the resulting shadow prices would still depend on somewhat arbitrary assumptions about future technical progress and about future expected input and output prices that the firm will face.

Inspite of the above rather negative evaluation of the Fisherian economic approach to asset valuation, accountants[170] have recognized that for certain *unique* assets held by a business unit, the economic approach may be the only relevant approach for obtaining current asset values. For example, a reasonable estimate for the value of a unqiue oil field held by an exploration company might be the estimated discounted net revenues generated by the crude oil pumped out of the field over the life of the field. In order to obtain these estimates, it will be necessary to: (i) estimate how much crude will be extracted in each future period; (ii) estimate future spot prices per barrel of crude (less applicable taxes); (iii) estimate future extraction costs and (iv) provide an appropriate discount rate. In fact, there are engineering firms that will provide such estimates and accountants accept their valuations in order to put an estimated value on oil reserves. As another example, suppose a business unit holds the rights to a movie or a patent; (both are unique assets). Then a reasonable current asset value for the movie might be the discounted value of future expected rental income and for the patent might be the discounted value of future anticipated royalty payments.

In order for economic valuations to pass the objectivity or reproducibility test, it seems necessary that these valuations be done by specialized valuation firms, which could be accredited by the relevant accounting standards board or by the relevant governmental authority.

We turn now to our final class of methods for valuing assets.

6.6 Specific Price Level Adjusted Historical Cost

> "On account of the expense involved, to argue for yearly appraisals of fixed assets, would sound impractical. When price levels remain fairly constant they would prove to be unnecessary. During periods of price fluctuation an adjustment could be made in previous appraisals to reveal this condition or an entirely new appraisal resorted to. In this connection, price indexes may prove very helpful in the future to both the accountant and the appraisal engineer".
>
> *H.C. Daines*[1929; 101]

> "Knowing the exact composition of the client's property as at the date for which the new appraisal is to be made, the appraisal company then values such property at the prices prevailing on that date .... A method that may very conveniently and profitably be used as a quick and cheap substitute under certain conditions is the index-number method. This method is a phase of 'stabilized accounting', which is concerned with the use of index numbers to restate accounting figures in a uniform price level before combining or comparing them".
>
> *Henry W. Sweeney*[1934; 110]

The specific price level method for constructing current values for an asset held by a business unit through successive accounting periods was suggested by Daines [1929; 101], Sweeney [1934; 110] and many other accountants.[171] The method works as follows. First, assets held by the business unit at the beginning of period 0 are classified into a finite number of distinct asset classes. Secondly, it is supposed that index numbers that pertain to each asset class are available at the beginning and end of each accounting period. Finally, suppose that an asset was purchased at the beginning of accounting period 0 at the price P^0 [172], the period 0 depreciation rate for the asset is δ^0 and the asset inflation rate for the relevant asset class over period 0 is i^0 (i.e., the specific asset index number at the end of the period divided by the specific asset index number at the beginning of the period is $1 + i^0$). Then the *Specific Price Level Adjusted* (SPLA) historical cost of the asset at the end of period 0 is defined as

$$(55) \quad (1 - \delta^0)(1 + i^0)P^0.$$

Comparing (55) with (54), we see that the present specific price index number method for constructing an end of period estimated asset value is very similar to the General Price Level Adjusted asset value defined earlier by (54); the only difference is that now a presumably more relevant specific price index is used for revaluation purposes rather than an index of general inflation.

If the same set of asset specific price indexes is given to all accountants, then Specific Price Level Adjusted values will satisfy the reproducibility test. The SPLA asset value should also be closer to its end of period market value (i.e., an end of period purchase cost or net realizable value) since presumably, the index numbers reflect a sample of market transaction prices for new units of the asset (or similar assets) during a time period that includes the end of period 0. Thus SPLA values will tend to be *reproducible* and *relevant*. We also note that Specific Price Level Adjusted accounting is not completely impractical since it has occasionally been used historically.[173]

There are some problems associated with the use of Specific Price Level Adjusted values: (i) None of the available specific price indexes may be relevant for the particular asset on hand.[174] A related problem is that different accountants may classify the same asset into different asset classes thus destroying the reproducibility property for the method.[175] (ii) The asset specific index numbers will generally pertain to a discrete interval of time instead of the precise date at which the accounting period ends. Under these conditions, the exact adjustments (if any) that the accountant should make to the specific indexes is ambigous.[176] (iii) The related issue of the timeliness of the specific indexes should also be raised: annual specific price indexes for capital stock components that appear with a half year time lag will be useless in the context of quarterly accounting.[177] (iv) The construction of SPLA values is mainly suitable for the valuation of fixed capital stock components and not circulating capital stock components.[178] How then should end of period prices for inventory stocks be constructed? We shall address this problem in section 9 below. The problems involved in constructing current prices for inventory items are generally not as severe because relevant market prices for inventory components held by the business unit are often available in the records of the business unit: market prices for used fixed assets are more difficult to obtain. (v) The SPLA values for assets at the end of the accounting period are still dependent on the rather arbitrary depreciation rates (recall δ^0 in (55)) that are associated with historical cost accounting.[179] To cure this lack of reproducibility in the method, the Agency that provides the asset specific index numbers should also provide "standard" depreciation rates for assets in each class (or alternatively, provide index numbers for not only new assets but also used assets[180]). The adoption of this last suggestion will not only lead to reproducible SPLA values, but it will also lead to reproducible estimates of depreciation.[181]

Which Agency should provide the relevant index numbers and depreciation rates? Three possible choices are: (i) the relevant National Statistical Agency; (ii) the relevant Accounting Standards Board or (iii) an Agency or Department of the relevant National Government (e.g., the income taxation authority).

We note that historical cost valuations for fixed assets have proved to be very resilient from a historical perspective, being temporarily abandoned only in the face of dramatic inflationary shocks when the method clearly became absurd.[182] It seems likely that the alternative valuation methods described above in sections 6.2-6.6 have failed to be adopted permanently for a number of reasons: (i) the alternative method was thought to be too *inaccurate* (General Price Level Adjusted valuations); (ii) the alternative method was thought to be too nonobjective or not *reproducible* (all other methods) or (iii) the alternative method was thought to be too *expensive* or too *complex*. However, it seems possible that all of these objections could now be overcome with the use of Specific Price Level Adjusted values, provided that a National Authority could provide the accounting profession with the relevant asset specific index numbers and standard depreciation rates.

We leave our last words on the subject of asset valuation to one of the pioneers of current value accounting:

> "Even crude attempts should result in an improvement over present depreciation practices. During periods of rapidly changing prices crude measurements of a relevant item are likely to be much more meaningful than accurate measurements of an irrelevant one (in this case, historic cost)".
>
> *Edgar O. Edwards*[1954; 268]