THE “ARCHITECTURE” OF CAPITAL ACCOUNTING:
BASIC DESIGN PRINCIPLES

Charles R. Hulten
University of Maryland and NBER

July, 2005
(revised)

“... it is not reasonable for us to expect the government to produce statistics in areas where concepts are mushy and where there is little professional agreement on what is to be measured and how (Griliches (1994), page 14).”

National income accounting would be a relatively simple matter were it not for “capital.” All flows of output would then be for immediate consumption, and labor would be the sole factor of production (and relatively undifferentiated labor at that, since there would be no investments in health and education to complicate matters). The question of how the boundary between market and non-market activity should be defined would be one of the main issues of contention, how to measure the real output associated with intangible products like services would be another. However, both problems are essentially issues of implementation rather than of basic theory, since there is no conceptual reason to exclude the non-market use of economic resources from a complete set of national accounts, nor is there a controversy about the need to express inputs and outputs in both current and constant prices.

When it comes to capital, however, it is more a question of “what” to do rather than “how” to do it. No issue has given economic theory more trouble, from Karl Marx and the

---

Austrian capital theorists, to Keynes and the Cambridge Controversies, and the ambiguity has only gotten worse with the increased theoretical focus on Schumpeterian uncertainty, partial information, imperfect competition, and with the emerging literature on the importance of intangible capital assets. This unsettled state of affairs is obviously a problem for the design of national income accounts, since, as Griliches (1994) observed, it is hard to measure something when there is a fundamental disagreement about what exactly “it” is. This ambiguity is reflected in the current design of national accounting systems, as well as in the structure of financial accounting systems. No system currently in place achieves a complete account of capital in its many facets and dimensions.

These observations are the starting point for this paper on the architecture of the capital accounts. To use the architectural analogy, the paper is about abstract design principles and is not a blueprint for a particular building; it is about the logically prior question of “what” should be done, rather than a discussion of how do “it.” It is inspired by Koopmans’ (1947) famous injunction about the need to avoid “measurement without theory.” This injunction argues that theory should guide measurement practice in order to guide the selection and definition of the variables included in the accounts and to define the boundaries, insure internal consistency among these variables, and facilitate their interpretation and subsequent use. However, while the Koopmans’ injunction is especially important for defining the role of capital in the national accounts, it does not specify any particular theory and, of course, there are many candidates for this role: capital accounts can be built along Keynesian lines (as with the traditional structure of the U.S. National Income and Product Accounts (NIPA); or, they can be broadly defined to include environmental, social, and quality-of-life indicator variables; even within the corpus of ‘standard economics,’ there are at least three ways of describing capital within the neoclassical
growth model alone. No single approach can claim to be unambiguously superior for all purposes (or views of the world), since the objective of any set of accounts is to inform a particular issue, and it is the user that defines both the relevant questions and the desired method of informing the answer. However, a choice of architecture does have to be made, if for no other reason than to insure internal consistency, and this paper describes an architecture design based on neoclassical economics. This architecture has as its foundation the familiar circular flow model of payments and commodities derived from standard supply and demand analysis, and it has the neoclassical theory of production and consumption as its superstructure. This architecture is implicit in contemporary accounts like the NIPA and the United Nation’s System of National Accounts, though not fully realized there, and is similar to the structure outlined in the work of Christensen and Jorgenson (1969, 1970).

The circular flow model (or “CFM”) is organized along lines of functional activity (consumption and production), rather than structural lines (non-financial business, financial business, government, foreign sector, households).² Agents have dual roles in the accounting structure: acting as producers, they supply products for current consumption or for future consumption via investment goods, and they demand the factor inputs that are necessary for the production of goods; acting as consumers, they supply these factor inputs and demand the producers’ output. The sectors are linked by markets in which the inputs and outputs are exchanged and valued according to the ‘laws’ of supply and demand. Some of the exchanges may exist only as shadow prices outside the formal market context and therefore require

---

² Patinkin (1973) traces the circular flow model, in its modern form, to the work of Frank Knight in the 1920s and 1930s. Earlier forms of the model can be found, according to Patinkin, but were apparently not intended as a representation of the allocation and distribution of goods and services in a complete economic system (the use to which the model is put in this paper).
imputation by the economic statistician, but this difficulty is, again, a problem of implementation and not of basic design.

Because of the complex nature of capital, the paper starts with a minimalist description of the CFM in which there is no capital of any kind. Using this as the baseline case, the paper then introduces various aspects of “capital” in order of increasing complexity, starting with a variant in which capital arises only from a temporal mismatch between the production and use of consumption goods, without any actual capital goods. The CFM is then expanded to allow for capital, starting with the stock of inventories and proceeding thereafter to productive capital inputs in both tangible and intangible forms. In following this sequence, it may appear as though there are many separate and distinct entities called “capital.” However, a comparison of each case reveals the following unity: all aspects of capital ultimately are derived from the decision to defer current consumption in order to enhance or maintain expected future consumption.

The functional structure of the CFM also reveals the dual nature of this unified conception of “capital.” From the standpoint of consumers, deferred consumption involves the diversion of current income from consumption to saving, which adds to the stock of consumer wealth and which leads, in turn, to higher income in the future and thus to higher future consumption. From the standpoint of the production sector, deferred consumption involves the diversion of resources away from consumer-good producing industries to investment-good industries. This diversion adds to the stock of productive capital and leads, in turn, to a larger output of consumption goods in the future. This dual structure helps clarify the various linkages between producers’ capital and consumers’ wealth, between producers’ investment and consumers’ saving, and between the cost of the capital to the producer and the income from
The distinction between capital cost and capital income is particularly important for any discussion of the architecture of capital accounting, since it is largely ignored in contemporary national accounting practice. By insisting on its importance, indeed necessity, the CFM establishes its utility as an architectural model for a consistent system of national income and wealth accounts.

II. The Basic Circular Flow Diagram

The structure of the circular flow model is shown in Figure 1, which describes the flow of payments and quantities in a four-part diagram in which the production activities are located on one side of a diagram and consumption is on the other side, and in which inputs are consigned to the lower half of the diagram and outputs are put in the upper part. Input and output flows are linked by “markets” in the upper and lower parts of the diagram, where the goods are transferred from one sector to the other. These bilateral exchanges are the essential feature of the CFM, and when validated by market valuations, establish the equivalence between revenue, cost, income, and expenditure. These exchanges are portrayed as a counter clock-wise flow around the outer edge of the Figure 1, and they give rise to the fundamental accounting identity relating the value of output to the value of input. When consumption (C) and labor (L) are the only goods in the account, the equivalence of flows in the CFM reduces to \( P^C C = P^L L \).

The ability to track the quantities of input and output over time is one of the main reasons that national accounts are constructed, since it is the flow of goods at any point in time (and not their nominal value) that is the determinant of economic well-being. The flows of consumption goods and labor input are portrayed as a clock-wise inner flow around the inner edge of Figure 1. This flow can easily be derived from the opposing flow of nominal values in any one period by

---

3 The distinction between capital cost and capital income is particularly important for any discussion of the architecture of capital accounting, since it is largely ignored in contemporary national accounting practice. By insisting on its importance, indeed necessity, the CFM establishes its utility as an architectural model for a consistent system of national income and wealth accounts.
simply normalizing all prices to equal one (i.e., $P^C = P^L = 1$). If these prices remain constant over time, there is no problem of intertemporal comparability and, in fact, no real need for considering prices at all. However, there is no reason to expect them to remain constant, since both productivity change and monetary inflation will cause nominal prices to change over time, both relatively and absolutely. In this case, the value flows $P^C C$ and $P^L L$ must be separated into price and quantity components in each period, using either independent estimates of price deflators or quantity indexes. The result is a time series of income and product accounts in both nominal prices,

$$P^C(t)C(t) = P^L(t)L(t) \quad (1a)$$

and constant prices

$$P^C(0)C(t) = A(t)P^L(0)L(t). \quad (1b)$$

The factor $A(t)$ must be included in the constant price account in order to allow for autonomous changes in productivity over time.

There has thus far been little reference to the economic structure that gives rise to the flows of the CFM. There is, however, an implicit structure embedded in the architecture of Figure 1 simply by virtue of its organization into sectors and markets. The flow of $L$ into the producers’ sector and the flow of $C$ out of the sector imply a transformation of input into output, which is formalized in standard theory by the production function $C(t) = F(L(t),t)$. The “$t$”, here, allows for costless advances in the efficiency of production, and is the source of the term $A(t)$ in the constant-price identity (1b). Similarly, the flow of $C$ into the consumers’ sector and the outflow of $L$ implies consumer choice among competing alternatives, which is modeled in
standard theory by the utility function, $U(C(t), L(t))$.\textsuperscript{4} These production and utility functions can be linked to the accounting identities (1a) and (1b) using Euler’s Theorem (see, for example, Hulten (2001)) and also have the helpful feature that they establish natural boundaries for the flow accounts in the CFM of Figure 1, or, indeed, for any set of accounts whose purpose is to provide a complete description of how available resources are used to satisfy economic wants. This theoretical structure suggests that any produced good that yields utility, and any input that is necessary for production, should be located within the boundaries of a complete set of economic accounts, regardless of whether they are distributed outside of formal markets.\textsuperscript{5}

The formulation of the accounting model (1) makes no reference to the goods that are produced for immediate use in other industries (steel to make autos, for example). These intermediate goods are important when the aggregate economy is broken into sectors, and introduce additional complexity into the CFM architecture. However, since these goods are produced and used up within each accounting period, they disappear in an aggregate account of the economy. Unfortunately, this can be deceptive and lead to the erroneous conclusion that the

\textsuperscript{4} The utility function is more commonly expressed in terms of consumption and leisure, rather than hours worked, i.e., as $U(C(t), H-L(t))$, where $H$ denotes the hours available for work and leisure is thus $(H-L)$. In a multi-product version of the CFM, each of the $N$ types of consumer goods has its own production function and, in principle, use some of each of the $M$ types of labor input, $C_i(t) = F(L_{i1}(t), \ldots, L_{iM}(t), t)$; the associated utility is then $U(C_1(t), \ldots, C_N(t); H - \sum \sum L_{ij}(t))$. For clarity, we will assume the fewest number of goods necessary for the exposition. We will also suppress the leisure term by ignoring the constant $H$, in order to simplify the exposition, since it does not play a central role in the capital accounts described in this paper.

\textsuperscript{5} The empirical problems associated with the development of the price deflators are often a practical constraint on the choice of accounting boundary. The problem is particularly difficult when the goods in question are intangible, where the units of measurement may be hard even to define in principle, much less to estimate accurately. Determining the appropriate deflators for non-market goods is also notoriously difficult. The result has been the tendency to exclude investments in human capital, research and development, and other intangibles from the accounts, along with non-market uses of time, despite the theoretical rationale for their inclusion.
concept of intermediate goods is largely irrelevant in the aggregate (except, possibly, for the
distinction between value added, now $P^L$, and deliveries to final demand, $P^C C$). The deception
arises because the distinction between what is an intermediate good and what is capital depends
on the length of the accounting period selected. If the period is one month, a pencil is likely to
be a capital good, whereas if the accounting period is one year, is it likely to be an intermediate
good. A machine tool with an average economic life of 20 years is capital when the accounting
period is one year but is an intermediate input when the period is a century. Since the length of
the accounting period is arbitrary, capital as a productive good is itself an arbitrary accounting
concept that can, in fact, be dispensed with in certain accounting models (see, for example,
Hulten (1979)).

III. Capital as Deferred Consumption

The all-consumption-labor model of the preceding section envisions a world without
capital goods. Since all goods are consumed when they are produced and all input is
contemporaneously generated, the magnitudes of these aggregate economic variables refer to the
current accounting period, and there is no connection between periods (i.e., between $C(t-1)$ and
$C(t)$, or between $L(t-1)$ and $L(t)$). However, a connection may exist at the level of the individual
agent’s utility function. Since the life of most people spans multiple accounting periods, the
individual utility function is more plausibly written as a function of consumption and leisure
over an entire lifetime $T$, i.e., as $U(C_1(1), \ldots, C_T(T); L_0(0), \ldots, L(T))$, rather than a single period of
that life. Maximization of this intertemporal utility function subject to the amount of income
that can be earned in each year, $P^L(t)L(t)$, results in an optimal consumption plan in which
desired consumption may exceed or fall short of income in any year. For example, individuals may want to shift consumption from periods of high income to others where income is lower (e.g., to years of retirement).

The opportunity for individuals to shift consumption arises if there are financial instruments that accommodate intertemporal transfers. The existence of such instruments allows individual consumers to lend or borrow part of their current income, \( P^l(t)L_j(t) \), which is to say, it allows individuals to save or dissave. Since total consumption is fixed, the saving of lenders must just balance the dissaving of borrowers in every year. The income and product accounts that accommodate this consumption-shifting mechanism are an elaboration of the aggregate account (1a,1b). When consumption shifting occurs at the individual level, an individual either saves or dissaves the amount \( S_j(t) \), leading to the individual income identity

\[
P^c(t)C_j(t) + S_j(t) = P^l(t)L_j(t).
\]

For both (1a,1b) and (2) to hold simultaneously, total savings across all individuals, \( \sum_j S_j(t) \), must be zero in each year, because all goods produced within a given year must be consumed within that year.

Moreover, individual saving or dissaving must balance over the lifetime of each person since all loans must be repaid with interest. The basic intertemporal constraint on each individual’s borrowing and lending is the discounted present value of lifetime labor income, which must equal the discounted present value of lifetime consumption. Assuming that there are no bequests to future generations nor inherited wealth from the past, the lifetime budget constraint at the start of economic life thus takes the form:
The time-discount factor \((1 + r)\) is assumed to be constant over time for simplicity of exposition.

In light of (2), individual savings must also have a zero balance over the lifetime of each person:

\[
W_j(0) = \sum_{t=1}^{T} \frac{P^C(t)C_j(t)}{(1+r)^t} = \sum_{t=1}^{T} \frac{P^I(t)L_j(t)}{(1+r)^t}.
\]  

(3)

This reflects the fact that all loans must be repaid out of lifetime income. Moreover, aggregate net worth at each point in time is zero, since the contemporaneous sum of individual net worth in each, \(\Sigma_j NW_j(t)\), must reflect the condition that \(\Sigma_j S_j(t)\) is zero.

Because no net wealth is created at the economy-wide level of aggregation, there is no aggregate sheet balancing assets and liabilities. However, a balance sheet based on these present value equations does exist for each individual agent, which records in each year the net consequences of all past saving and dissaving. The existence of these individual net worth positions implies that wealth, and the corresponding balance sheets, can exist even though there are no explicit capital goods and no consumption goods are actually shifted between years.

This conclusion must be modified when an economy is open to international flows. In this case, borrowing and lending can occur across national boundaries, and there can be a non-zero net balance of claims or debits against future consumption for the residents of any one country. While the net position of all countries combined is still zero, the aggregate wealth constraint of each country is now
The aggregate net worth (and implied balance sheet) of this open economy takes the form

$$W(0) = \sum_{t=1}^{T} \frac{P^C(t)C(t)}{(1+r)^t} = P^C(0)K(0) + \sum_{t=1}^{T} \frac{P^L(t)L(t)}{(1+r)^t}.$$  

(3')

$P^C(0)K(0)$ is the cumulative balance of past external loans or debt (i.e., past saving or dissaving) up to the beginning of the decision interval (the “present”), measured in terms of current consumption. The net worth can be either positive or negative at any point in time, leading to the conclusion that a form of “capital” is implicit in economic activity even when there are no explicit capital goods and, indeed, even when all the consumption goods produced within a time period are also consumed (by someone) during that period.

IV. Capital as an Inventory of Goods

A small “tweak” to the analysis of the preceding section gives further insights to the capital problem. The discussion of Section III examined the situation in which the consumption good had to be consumed in the period it was produced, but could be effectively shifted between time periods through the issuance of “paper” debt agreements among people with different preferences about the timing of their consumption. An important variant on this theme arises when the consumption good can be stored, and therefore shifted directly from one period to the
next. While this tweak is small, the implications are not. There is now a transfer of real goods over time, not just debt obligations, and it is thus possible to speak of a “stock of capital goods,” albeit a stock composed entirely of consumption goods.

In order for an inventory of goods to be carried over from one period to the next, there must be some provision for storing the goods until they are consumed. It is natural to locate the storage activity in the production sector, given the functional classification of activity into either consumption or production and the observation that the act of storing the good can be thought of as production for future consumption. In this formulation, the production function for the consumption good must be modified to reflect the possibility that part of the current output is diverted to future use. The production function is now \( C(t) + I(t) = F(L(t),t) \), where \( I(t) \) is the amount of the goods sent forward to the next period. The aggregate income identity is \( P^C(t)C(t) + P^C(t)I(t) = P^L(t)L(t) \).

Consumption can exceed the total quantity of the good produced during any period in this framework if there is a stock of stored consumption goods carried over from past production. This stock is equal to the addition to inventory, \( I(t) \), plus any balance of unused goods left over from previous periods, \( K(t-1) \), adjusted for wastage at a rate \( \delta \), into the inventory goods, \( K(t) \), available for consumption in the next period:

\[
K(t) = I(t) + (1-\delta)K(t-1). \quad (5)
\]

The value of this stock in any year is determined by its replacement value, \( P^C(t)K(t) \), at the current commodity price. It is also the value that the producers could capture if they were to sell their entire inventory of goods, and can thus be regarded as an asset of the producer sector.
There is, however, an offsetting liability arising from the fact that producers have revenue of $P^C(t)C(t)$ from the sale of the consumption good, but have a wage obligation of $P^L(t)L(t)$. The difference, $P^L(t)L(t) - P^C(t)C(t)$, is a deferral of wages that can be thought of as the consumers’ claim against the inventory stock held by the producers’ sector.

One way to model this deferral is to suppose that producers issue a paper claim, e.g., a debt or equity instrument, that promises to pay an amount equal to the wage deferral: $P^C(t)S(t) = P^L(t)L(t) - P^C(t)C(t)$. The $S(t)$ in this formulation is the amount of deferred consumption in units of the good and can therefore be thought of as the quantity of goods saved (an amount equal to inventory investment, $I(t)$), even though the saving takes the form of paper claims. $P^C(t)S(t)$ is the nominal value of these claims, and can be thought of as an increment to consumers’ net worth. This formulation of saving in terms of claims to future consumption is evocative of the pure consumption-loan model of the preceding section: the net worth equation (4') applies equally to the analysis of this section, since net worth is the difference in the present values of the future streams of consumption and labor in both cases; and, in both cases, net worth is equal to $P^C(t)K(t)$. The main difference between the pure consumption-loan model and the inventory model lies in the fact that the $K(t)$ in the latter represents a stock of actual goods. It is therefore possible to speak of a true balance sheet in this case, with the value of the stock on one side of the balance sheet and the wealth claims against this stock on the other.

This balance sheet exists along side the income and product account of the circular flow model, raising the question of where to locate the balance sheet in the CFM diagram. By its very nature, the CFM portrays the flows of goods and payments into and out of the production and consumption sectors of the economy, and there is no provision for a stock of goods linking one accounting period to the next. One resolution of this problem is to append the items on the
balance sheet account to the relevant sectors. The *capital stock account*, \( P^C(t)K(t) \), can be attached to the producer sector and the *wealth account*, \( NW(t) \), to the consumer sector. The linkage between the two stock accounts and the flows of the CFM occurs via a *saving and investment account*, in which investment in inventories, \( I(t) \), flows into the capital account and the saving, \( S(t) \), flows into the wealth account.

V. Capital as a Produced Means of Production

The capital accounts described up to this point are missing one essential thing: the entity that most people intuitively regard as “capital” — something rather solid and durable like machines and buildings. Unlike the preceding inventory case, where the capital stock is an inventory of consumption goods that has already been produced, this sort of capital is an input that is used to produce future output, and, at the same time, is itself produced. The essential analytical difference between the two cases thus lies in the structure of production. In the inventory case, the production function takes the form \( C(t) + I(t) = F(L(t), t) \), with a subsidiary storage function implicit in the analysis. In the case of productive capital, the structure of production must reflect the fact that investment is a separate good with its own production function. The various equations of the preceding sections must be modified accordingly:

\[
I(t) = F^I(L^I(t), K^I(t), t) \quad \text{and} \quad C(t) = F^C(L^C(t), K^C(t), t) \quad (6)
\]

with adding up conditions \( L = L^I + L^C \) and \( K = K^I + K^C \); the accumulation equation (5) remains \( K(t) = I(t) + (1-\delta)K(t-1) \). This structure differs from the preceding case in two regards:
investment \( I(t) \) and consumption \( C(t) \) are now distinct goods and are not perfect substitutes as before; and, the stock of capital \( K(t) \) is now an intertemporal factor of production.

The difference in the structure of production between the two cases carries over to the valuation of capital goods. The unit of value associated with the inventory stock is the price of consumption goods, \( P^C(t) \), which is related to the utility function. In the case of productive capital, there are two prices: one associated with the output of the investment good, \( P^I(t) \), which is related to the marginal cost of producing the good; and, one associated with the use of the good as an input in production, \( P^K(t) \), which is related to the marginal productivity of the capital input and is the rent that the good could command for use in annual production. The two price concepts are not independent, since the willingness to pay for a unit of new capital stock must be related to the future stream of rents generated by that good, \( P^K(t+\tau) \), over its economic life, \( N \).

Under the assumption that investment will continue in any year up to the point at which the cost of the last unit just equals the discounted present value of the rental income it generates:

\[
P^I(t,0) = \sum_{\tau=0}^{N} \frac{P^K(t+\tau,\tau)}{(1+r)^{\tau+1}}.
\]  

(7)

The discount rate is, again, denoted by \( r \). We will assume for the remainder of this section that all productive capital is rented in formal markets, so that \( P^K(t) \) is an observable price. The situation in which capital is producer-utilized is deferred to the following section.

The circular flow diagram corresponding to the case of productive capital is shown in Figure 2. The flows around the diagram have been adjusted to account for the output of investment goods by the producer sector and the flow of rental income out of the sector to
consumers. The basic GDP identity is determined by the products of the two sectors: \( P^i(t)I(t) = P^i(t)L^i(t) + P^k(t)K^i(t) \), and \( P^c(t)C(t) = P^i(t)L^c(t) + P^k(t)K^c(t) \). The aggregate identity is then the sum of the two:

\[
P^c(t)C(t) + P^i(t)I(t) = P^i(t)L(t) + P^k(t)K(t).
\] (8)

The flows in the upper right-hand side of Figure 2 show the product flow from the standpoint of the consumer, for whom the acquisition of the capital good is an act of saving:

\[
P^c(t)C(t) + P^i(t)I(t) = P^c(t)C(t) + S(t).
\] (9)

This is leads to the saving account introduced in the preceding section, which now has the form \( P^i(t)I(t) = S(t) \). Productive capital is stored by consumers in the producers’ sector, as before. However, producers are now assumed to purchase units of new capital up to the point that marginal cost equals the discounted present value of the stream of future rents, as per equation (7). The resulting \( P^i(t)I(t) \) flows into the investment account from the producer sector, and \( S(t) \) flows in from the consumers’ saving account, which is connected to the investment account via the financial market for debt and equity instruments. The \( P^i(t)I(t) \) is added to the producers’ capital account, and the \( S(t) \) to the consumers’ wealth account, as before. The new accounting element in the Figure 2 variant of the CFM is the disposition of the rental payments, \( P^k(t)K(t) \). They can be thought of as flowing through the rental market into a capital payment account attached to the producers’ capital account, from which they flow into the consumers’ capital income account in the form of the return to the debt and equity instruments held in the wealth account.
The analogue to the balance sheet equation, (4') above, is derived from the expanded income identity by taking present values. In the current case, it takes the form:

\[
NW(0) = \sum_{t=1}^{\infty} \frac{P^C(t)C(t)}{(1+r)^t} - \sum_{t=1}^{\infty} \frac{P^I(t)L(t)}{(1+r)^t} = \]

\[= \sum_{t=1}^{\infty} \frac{P^C(t)K(t)}{(1+r)^t} - \sum_{t=1}^{\infty} \frac{P^I(t)L(t)}{(1+r)^t} = P^I(0)K(0) \quad (10) \]

The last equality above follows because the terms involving \(K(t)\) and \(I(t)\) cancel out except for the value of the initial endowment of capital carried forward from the past. This is, of course, exactly the same result previously obtained for inventory capital and for net foreign debt. The income and wealth accounts in the case of productive capital are thus an extension of the accounts described in the preceding sections, revealing that the architecture of the accounts is not fundamentally changed when the concept of capital is extended to include capital as a produced means of production.

VI. Producer-Operated Fixed Capital

Accounting for fixed assets would still be relatively easy if only they were rented in active markets. All the prices in the income and product accounts would correspond to observed data based on actual market transactions. Unfortunately, this is not the way capital markets normally operate. Only a small fraction of the productive capital stock flows through a formal
rental market, so there is thus no explicit rental price \( P^K(t) \), nor a rental flow \( P^K(t)K(t) \), for the national accountant to observe. The lower left “cost” branch in Figure 2 is therefore an empirical void, and, as a result, current accounting practice has traditionally ignored this cost branch, and has thereby lost sight of the structure of the circular flow model. Fortunately, this situation is beginning to change.

The absence of formal rental markets unquestionably creates a serious empirical challenge to the task of implementing the full CFM, but economic theory provides a way of using the observable aggregate data to impute the unobserved rental price. The solution is based on the user cost of capital approach, pioneered by Jorgenson (1963) and Hall-Jorgenson (1967), in which the asset pricing equation (7) is solved to yield an expression for the implicit rental price:

\[
P^K(t) = \left[ r(t) - \rho(t) + (1 + \rho(t))\delta \right] P^I(t) \quad (11)
\]

This expression is the user cost for a new asset without provision for taxes or other complications. It is the opportunity cost that the producer-user of the capital must recover in each year, and it is equivalent to the rental price of capital. Estimates of the user cost can be constructed for each type of asset if the elements on the right hand side of (11) are measurable: the rate of asset price revaluation, \( \rho(t) \), the rate of economic depreciation, \( \delta \), the rate of return, \( r(t) \), and the acquisition price of the asset, \( P^I(t) \). The result is an imputed value of the rental price for each type of capital.

Estimation of the components of the user cost imputation varies in the degree of difficulty. Data on the acquisition price of the asset of tangible capital, \( P^I(t) \), is readily obtainable, yielding estimates of asset price revaluation, \( \rho(t) \). Estimation of the rate of economic
depreciation, \( \delta \), presents the greatest difficulties, and is discussed in a separate section below. Estimating the rate of return, \( r(t) \), is also a difficult issue (see Schreyer (2004) for a recent discussion of this problem). One approach is to base the estimate of \( r(t) \) on the rate of interest used to finance the acquisition of the capital good, under the assumption that arbitrage will drive the rate of return into line with this interest rate. Another option is to use a weighted average of the return to debt and corporate equity. The use of an \textit{ex ante} measure of \( r(t) \) in one form or another has the virtue of tying the user cost to the financial costs that investors face when contemplating the acquisition of a capital good. On the other hand, there is no guarantee that the investor’s decision was in fact based on any particular financial rate of return. The presence of risk and liquidity constraints might cause the investor to use a higher rate of discount in assessing the costs and benefits implicit in the asset pricing equation (7). Moreover, there is no guarantee that any \textit{ex ante} measure of \( r(t) \) will lead to an imputed estimate of \( P^k(t) \) that satisfies the requirement that the value of input should equal the value of output in the fundamental income and product accounting identity (8).

Jorgenson and Griliches (1967), and Christensen and Jorgenson (1969, 1970), develop an alternative approach to imputing the user cost based on an \textit{ex post} rate of return to capital that insures that the right-hand side of (8) equals the left-hand side. The first step in this procedure is to estimate the total payment to capital input, \( \Pi(t) \), by measuring the residual revenue not paid out as labor income, i.e., \( \Pi(t) = P^c(t)C(t) + P^l(t)L(t) - P^k(t)K(t) \). This expression is also equal to \( P^k(t)K(t) \) given the basic accounting identity, which results in

\[
\Pi(t) = P^k(t)K(t) = [r(t) - \rho(t) + (1 + \rho(t))\delta] P^l(t)K(t), \tag{12}
\]
when the user cost formula is inserted in place of $P^K(t)$. Once the other elements on the right-hand side of (12) have been estimated, (12) can be solved to yield an estimate of $r(t)$. As an added bonus, Berndt and Fuss (1986) show that this ex post rate of return can be interpreted as a Marshallian quasi-rent that embodies a correction for changes in capital utilization. On the negative side, this procedure assumes that all the relevant capital has been accounted for in (12), and that there are no other sources of rent (an issue taken up in greater detail below). An empirical difficulty also arises because this procedure can lead to the imputation of negative user costs during periods of high asset-price inflation.

VII. Producer-Constructed Fixed Capital

The preceding section explored the implications for national income and wealth accounting of indirect ownership, where the problem was the absence of an explicit rental price. A further problem arises when capital stock is not only producer-operated, but also producer-constructed. When capital assets are constructed and used by their operators, not only is there no market rental price $P^K(t)$ to estimate, there is no observable market transactions on the value of the asset acquired this way (i.e., on the implied $P^I(t)l(t)$). The cost of constructing this type of asset can sometimes be determined, but a firm’s internal construction costs may not always equal

---

When there are multiple types of productive capital, with different user costs, the analogue of (12) is

$$I(t) = \sum_{i=1}^{N} P^K_i(t)K_i(t) = \sum_{i=1}^{N} \left[ r(t) - \rho_i(t) + (1 + \rho_i(t))\delta_i \right] P^I_i(t)l_i(t)K_i(t)$$

As with a single asset, once all of the other elements on the right-hand side of this expression have been estimated, it can be solved to yield an estimate of the common $r(t)$. 
the value of the asset it constructs, due to lumpiness and other factors. Moreover, even when
they are equal, a problem still arises because there is no price deflator, $P(t)$, with which to isolate
the real quantity of the investment, $I(t)$.

The problem of self-construction is of limited quantitative significance for business
tangible capital, arising mainly in the case of certain types of maintenance and repair. However,
it is the dominant situation for investment in business intangibles like research and development,
computer software developed within firms, human competencies, and product marketing. Many
items in these categories are firm specific, in the sense that they are of value only (or mainly) to
the firm that makes them, and also because firm-specific expertise is required for their
production. Other items, like R&D, may also have a value outside the firm, but are closely held
because of appropriability problems. All share the feature that it is almost impossible to define
the units of measurement in which the quantity, $I(t)$, might be measured in principle, much less
observed in practice. Indeed, their very presence may be a matter of some dispute because they
have no tangible embodiment. Expenditures on intangible capital are thus potentially subject to
manipulation by firms seeking to “improve” balance sheets and income statements. As a result,
accountants at both micro and macro levels have historically been reluctant to treat them as
capital expenditures, though this is beginning to change at the macro level with the capitalization
of software expenditures in the NIPA and the move toward a satellite account for R&D.

The fact that this type of expenditure presents measurement difficulties has no direct
bearing on the question of whether or not the expenditure should be treated as capital. This is a
matter of the intrinsic nature of the good in question and, specifically, whether or not a current
expenditure is made in order to increase future consumption (or to prevent a decrease in future
consumption). If it passes this test, standard intertemporal economic theory unambiguously
implies that it should be regarded as saving and treated as capital (Corrado, Hulten, and Sichel (forthcoming)). When applied to specific cases, this rule suggests that those maintenance and repair expenditures made with the expectation that they will prevent a reduction of consumption in the future should largely be treated as capital. R&D spending is also capital formation under this rule, as are many other intangible business expenditures made with the intention of increasing future output and thus potential consumption. Outside the business sector, the opportunity cost of delaying entry into the labor force in order to acquire additional education should be treated as an investment in human capital, as should many health-related expenditures.  

VIII. Depreciation and Obsolescence

The problem of economic depreciation has troubled the field of national accounting for many years and therefore deserves special attention. The essential characteristic of depreciable assets is that they are “used up” in the process of production through wear and tear, causing the productive efficiency of an asset to erode as it ages. This erosion was dealt with in the accumulation (5) above by the simplifying assumption that it occurs at a constant rate $\delta$. A more

---

7 Corrado, Hulten, and Sichel (forthcoming) find that more than one trillion dollars of business intangibles are currently excluded from U.S. investment spending each year, a sum approximately equal to the amount business spends annually on tangible fixed capital. Jorgenson and Fraumeni (1989, 1992) also argue that large amounts of human capital spending are ignored by current practice.

8 A full account of this history is beyond the scope of this paper, as is the algebraic derivation of the link between asset deterioration, depreciation, and asset valuation. More complete accounts can be found in Hulten and Wykoff (1981, 1996) and Triplett (1996), and a mathematical formulation of the issues is provided in Hulten (1990).
The relative efficiency terms, $N_i$, are actually the marginal rate of technical substitution between a new asset and an asset of age $i$. The relative efficiency of a new asset is therefore one (i.e., $N_0 = 1$). The $N_i$'s are assumed to be constants in order to make the model empirically operational, but could be allowed to vary according to economic conditions in a more general model.

The general specification, which includes this simple case, defines the stock of capital as the sum of all surviving past investments weighted by their remaining productive efficiency, $\phi_i$:

$$K(t) = \phi_0 I(t) + \phi_1 I(t-1) + \ldots + \phi_N I(t-N).$$  \hspace{1cm} (13)

The relative efficiency terms, $\phi_i$, of this stock accumulation equation decline over time until they become zero at the time, $N$, when the asset is finally retired from service (the stock accumulation equation of the preceding sections assumes the special case $\phi_i = (1-\delta)^i$).

The decline in the $\phi_i$ weights leads, ceteris paribus, to a decline in the quantity of capital services. This, in turn, leads to a decline in the value of the capital asset, partly because of the loss in productivity before retirement and partly because each year that passes moves the asset closer to the end of its productive life, thus shortening the remaining stream of income in the asset valuation equation (7) above. The decline in value is termed “economic depreciation,” and is conceptually distinct from deterioration ($\Delta \phi_i$), though it should be noted that when depreciation follows a geometric pattern, the rates of depreciation and deterioration are identical (i.e., $\Delta \phi_i / \phi_i = \delta$). Given equation (7), economic depreciation $\delta(t)$ can be shown to be the partial derivative of the asset’s price, $P(t,s)$, with respect to age, $s$, while asset revaluation (the term $\rho(t)$ in the user cost expression (11)) is the partial derivative of price with respect to time. Under certain assumptions, an estimate of the pure age effect can be obtained by measuring the price differential between two similar assets of different ages at the same point in time. This is the basis for the Hulten-Wykoff measures of economic depreciation embodied in the U.S. National

---

9 The relative efficiency terms, $\phi_i$, are actually the marginal rate of technical substitution between a new asset and an asset of age $i$. The relative efficiency of a new asset is therefore one (i.e., $\phi_0 = 1$). The $\phi$’s are assumed to be constants in order to make the model empirically operational, but could be allowed to vary according to economic conditions in a more general model.
This analysis has the following implications for the measurement of income and wealth. The depreciation portion of the annual gross return to capital (i.e., the $\delta$ term in the user cost) must be considered to be the cost associated with maintaining the value of the original investment intact, and not as income to the owner. This principle implies that the gross value of the goods emanating from the production side of the CFM is equal to the net income accruing to the consumer sector plus depreciation, and not net income alone. This does not disturb the equality of the circular flow of payments, since the flows in the top quadrants of the CFM are the gross value of goods produced and purchased, and they are equal to the flow of gross payments from producers to consumers in the bottom quadrants. However, the difference between net and gross income suggests that there may be two separate measures of total economic activity, one appropriate to the production side of the CFM, the other to the consumer side, and that GDP is not necessarily the appropriate indicator of annual output’s contribution to intertemporal utility.

This idea is rigorously developed in Weitzman (1976), who shows that the optimal solution to the problem of maximizing the intertemporal utility function is $C(t)+p(t)\Delta K(t)$, where $p(t)$ is the investment good price relative to the price of consumption. This expression is Haig-Simon income, consumption plus change in net worth, but it is also equal to factor income less depreciation as well as to net domestic product (Hulten (1992)). The Weitzman result can be

---

10 It is worth noting that the depreciation experience of a single asset is not necessarily the same as the average experience of the whole cohort of similar assets that were put in place in the same year, and it is the whole cohort of assets that matters for national accounting purposes. The members of the cohort will generally be retired at different ages, and if the retirement pattern is normally distributed, the average rate of depreciation for the cohort will be close to the geometric rate (Hulten and Wykoff (1981)). This result seems to contradict the intuition that a physical asset like a chair retains most of its productive value up the point that it is retired from service (the one-hoss shay model), but this intuition is flawed by the fallacy of composition.
interpreted as implying that NDP is a better measure of aggregate economic activity than GDP, but the CFM makes clear that both are important. The production functions on the supply side of the CFM represent a transformation of labor and capital inputs into gross output — the output that actually leaves the factory doors. No one has ever, or can ever, see a unit of physical output net of depreciation leave a factory because it simply does not exist. GDP is the appropriate concept for studying the parameters of the production function and how the productivity of the inputs changes over time. On the other hand, NDP is the appropriate concept for studying consumer welfare, since, as Weitzman puts it, NDP “is a proxy for the present discounted value of future consumption (page 156).” This dichotomy points to the utility of the CFM as a way of classifying economic activity: once the issue of net versus gross output is framed in the context of the CFM, both concepts are seen to be important for their respective realms of economic activity. Moreover, the failure of accountants to maintain a clear view of the CFM architecture contributes to confusions like the net versus gross output debate.

Technological obsolescence is another aspect of asset valuation, and it greatly complicates the asset valuation model and has been the source of much confusion recently. This phenomenon occurs when improvements in technology are embodied in the design of new capital goods. These higher quality assets are often quite different from the older assets against which they compete (e.g., jet versus propeller-driven aircraft) and are, in principle, a separate type of capital that should be accorded its own production function. However, the data requirements of this approach are so great that they render it nonoperational. The embodied

---

11 If each new technological vintage of, say, machine tools were treated as part of a separate production process, data on output, labor, and material input, would have to be collected on a machine-by-machine basis. It is a major undertaking to assemble consistent production data at an establishment level of detail, and it is hard to expect that these data could be uniquely disaggregated to the level of individual machines.
This is not the only way to introduce embodied technical change into the capital accumulation model. An alternative approach would allow for an acceleration in the retirement of older vintages when the expected rate of embodied technical change accelerates. This is a more realistic approach, but it is also more complicated and is generally beyond the capacity of existing data to implement.

Technical change model can, however, be made empirically operational by assuming that differences in quality between old and new assets can be expressed as a difference in the effective quantity of the capital services that they represent. In terms of the preceding formulation in which quality differentials were absent, the relative efficiency parameters for new assets, $\phi_{t,0}$, is no longer constrained to equal one, but now increases over time at the rate of embodied technical change, $\lambda$. The efficiency function of new assets evolves, accordingly, as $\phi_{t,0} = (1+\lambda) \phi_{t-1,0}$. The efficiency function of existing assets of age $s$ in year $t$ takes the form $\phi_{t,s}$ and the vintage of that asset is denoted by $v = t-s$. This $\phi_{t,s}$ pattern is assumed to decline due to wear and tear alone, but not because of the arrival of new capital (thus, $\phi_{t+1,s+1} = (1-\delta)\phi_{t,s}$, when deterioration proceeds at a constant rate). The capital accumulation analogue to equation (13) for this case has the form

$$K(t) = \phi_{t,0} I(t) + \phi_{t,1} I(t-1) + \ldots + \phi_{t,N} I(t-N). \quad (13')$$

The essential feature of this approach is to characterize embodied technical change as an increase in the effective quantity of new capital, which is then added to the quantity of older capital adjusted for deterioration.\(^{12}\)

On the other hand, the value of older vintages of capital is driven down by the arrival of new capital even if their own productivity has not changed, because the value of the marginal product of older capital falls even if the marginal product itself is not affected: the superior

---

\(^{12}\) This is not the only way to introduce embodied technical change into the capital accumulation model. An alternative approach would allow for an acceleration in the retirement of older vintages when the expected rate of embodied technical change accelerates. This is a more realistic approach, but it is also more complicated and is generally beyond the capacity of existing data to implement.
efficiency of new capital translates into a fall in output price in competitive markets, which in turn lowers the marginal revenue earned by existing capital. Hall (1968) shows that the price of older assets, \( P(t,s) \), as defined by the asset pricing equation (7) now evolves according to three factors: the rate of embodied technical change, \( \lambda \), the average rate of wear and tear, \( \delta \), and the average revaluation effect, \( \rho \). He also shows that partial derivative of \( P(t,s) \) with respect to \( s \) is \( \delta + \lambda \), and the partial derivative with respect to \( t \) is \( \rho + \lambda \), and that it is impossible to disentangle the average rate of change of the three effects only using price data on used assets, \( P(t,s) \).

When economic depreciation is defined as the partial derivative of asset price with respect to age, it becomes apparent that depreciation includes the combined effects of wear and tear and obsolescence. Since the measures of economic depreciation currently in the U.S. NIPA are derived from the Hulten-Wykoff price-based estimates, these measures must be interpreted as the sum of \( \delta \) and \( \lambda \). The Hall 1968 result shows that these two effects cannot be separated, given used price data alone, but in another paper, Hall (1971) shows that the technique of price hedonics can be used to resolve the identification problem, that is, to provide separate estimates of the parameters \( \delta \) and \( \lambda \), as well as \( \rho \). These separate estimates can then be used to split price-based estimates of economic depreciation, \( \delta + \lambda \), into its components. The results can, in turn, be employed separately to measure the quantity of capital stock on the producer side of the CFM using the modified accumulation equation (13'), while, at the same time, the same estimates can be used to measure the change in the value of that capital, \( \delta + \lambda \).

IX. The “Nonzero-Rent Economy”

The accounting architecture described in the preceding sections has two levels: a
foundation based on the circular flow of goods and payments, and a superstructure based on the
application of economic theory. The first is rather general, but the latter employs specific
assumptions about technology and preferences and about the valuation of the stocks and flows,
assumptions that Hall (2001) collectively terms the “zero-rent economy.” That economy is
characterized by competitive markets, constant returns to scale, and the possibility that all factors
can be freely adjusted in the long run. There are thus no economic rents in that economy,
leading to the income identity (12) connecting the total gross return to capital assets, $\Sigma P_i K_i$,
and the flow of income to consumers generated by the assets. This is one of the main
assumptions of the accounting models of Jorgenson and Griliches (1967), and Christensen and

These identities require a different interpretation in an economy with economic rents
generated by monopoly power, intramarginal efficiency rents, persistent disequilibrium,
imperfect information, or uncertainty. In this more realistic world, some of the income thought
to accrue to the collection of capital assets, $\Sigma P_i K_i$, is in reality a return to entrepreneurship or
to the owners of the firm. Any attempt to impose the zero-rent-economy rules in this world
results in a biased estimate of the return to the specific capital assets included in the analysis.
Moreover, there is a potential disconnect between the value of capital stock and the amount of
wealth. However, this does not mean that the use of theory, per se, is at fault, but rather that it is
important to use the right theory. Nor does it mean that the zero-rent model is irrelevant. Given
the difficulty of adapting models of imperfect competition, Schumpetarian entrepreneurship, and
uncertainty to national income accounting problems, the zero-rent model is a logical and
important step along the way toward Koopmans’ vision of measurement with theory.
X. Summary and Conclusion

This paper has discussed the design principles of a set of income and product accounts based on the circular flow model of payments and goods. While these flows encompass much more than just capital, the CFM has been shown to be a useful framework for sorting out just what is meant by the term “capital.” In particular, the clear division on the CFM between the consumer and business sectors reveals the dual nature of capital: on one hand, it is a stock of inventories and productive assets held by producers, and, on the other hand, it is simultaneously a stock of wealth held by consumers. Moreover, by approaching the problem of capital in gradual stages, it has become clear that capital in its most primitive form originates with the decision by consumers to defer consumption, and that this aspect carries over to capital in all its forms.

These are general architectural principles. The analysis of capital within the context of the CFM also yields insights into specific accounting practices. First, these general principles suggest a criterion for determining what is and is not capital: if an expenditure is made with the intention of increasing future rather than current consumption, then it should be treated as capital. This rule clearly applies to most R&D and many other intangible expenditures, most of which are not treated as capital under prevailing practice. This situation is beginning to change, and the CFM analysis suggests that this trend needs to be sustained.

Second, the structure of the CFM also calls attention to the need for a full accounting for the cost of capital in the lower left-hand branch of the model. This part of the model is an integral part of the production account associated with the business sector of the CFM, but it is largely missing from conventional national accounts like the NIPA and the SNA. Extending these accounts to include a full production sector is currently under consideration, and the
analysis in this paper strongly supports the adoption of this approach.

Third, division of the CFM into two functional sectors helps resolve the debate over gross versus net product as a measure of aggregate output. The clear answer is that both measures are relevant statistics of an aggregate economy. Gross output (GDP) is the natural measure to apply to the producer sector, because the production functions of that sector transform input into output that is gross of depreciation. On the other hand, net output (NDP) is the more appropriate indicator of consumer welfare, both current and future. Many researchers have, in the past, lost sight of the circular flow organization of the economy and used net output in their analysis of productivity.

Fourth, the division of the CFM into production and consumption sectors helps resolve the question of asset assignment. This issue has arisen in the debate occasioned by the proposed revisions to the SNA. The CFM suggests that assets rented under very long-term leases or under sales-lease-back arrangements, or that have split ownership, should be accorded the same treatment as other assets: as productive capital, they should be assigned to the industry of the producer sector in which they are used to produce output. The ownership of the asset should then be traced through the process of financial intermediation to wealth accounts of the ultimate owners in the consumer sector.

The CFM, with its emphasis on production and consumption as separate economic activities, provides a useful architecture for sorting out many accounting problems associated with “capital.” It also provides a flexible infrastructure for incorporating theoretical developments that improve the value of the accounts to the various user communities. Moving current accounting practice toward the CFM structure should be a central goal of the field of national income and wealth accounting.
References


Figure 1
Circular Flow Model without Capital
Figure 2
Circular Flow Model with Capital