This PDF is a selection from an out-of-print volume from the National Bureau of Economic Research

Volume Title: The Measurement of Capital

Volume Author/Editor: Dan Usher, editor

Volume Publisher: University of Chicago Press

Volume ISBN: 0-226-84300-9

Volume URL: http://www.nber.org/books/ushe80-1

Publication Date: 1980

Chapter Title: The Measurement of Capital Aggregates: A Postreswitching Problem

Chapter Author: Murray Brown

Chapter URL: http://www.nber.org/chapters/c7170

Chapter pages in book: (p. 377 - 432)

7 The Measurement of Capital Aggregates: A Postreswitching Problem

Murray Brown

7.1 Introduction

The problem of capital measurement is a postreswitching problem in the sense that the literature that centered on reswitching and attendant perversities contributed little to our knowledge of the conditions for capital aggregation. But it did serve to focus attention on the problem itself —to motivate the inquiry by indicating in no uncertain terms that the failure to satisfy certain aggregation conditions (namely, the Gorman conditions) could lead to results qualitatively different from those one would expect from the so-called neoclassical parables (those based on aggregate neoclassical production specifications). There is now a considerable consensus on this point, and so the inquiry must proceed beyond reswitching into more detailed and empirically oriented analyses of capital aggregation. That is the principal concern of the present paper. But before taking this up, it may be useful to give a brief review of the reswitching phenomenon. Its implications, presented after the review, should be examined closely because they motivate further work and also embody a critique of what has been done using aggregate capital measures.

The misspecification that may result from using improper aggregates is not negligible. It affects the empirical foundations of production and distribution analyses (and all the spinoff implications these have for pricing, productivity, etc.). It will not go away if we merely look it in the eye and pass on, and hence it is bound to return in devilishly unpredictable forms to render those analyses unacceptable.

The body of this chapter is a critical examination of conditions that must be satisfied for the empirical specification of capital aggregates at

Murray Brown is associated with the Department of Economics, State University of New York at Buffalo, Amherst campus.

various levels of aggregation. Here one must turn away from the capital-theoretic reswitching literature and look critically at the aggregation conditions associated with functional form and relative prices, inter alia. These include the Leontief-Solow conditions, Fisher's aggregation analysis, the Gorman conditions, the Houthakker-Sato approach, Hicksian composite commodity aggregation, the Brown-Chang conditions, and the statistical case for production aggregation.

On the basis of this review, I feel that two recommendations can be adequately defended. The first is addressed to the development of the output, capital, and labor data used in production analysis. Since the very existence of stable aggregates is questionable, one must at the least suspend judgment on studies using such aggregates. It follows that problems requiring aggregates must be treated in such a way that the aggregates are justified empirically. To do that requires that sufficiently disaggregated data be available upon which aggregation conditions can be tested. Thus, the first recommendation is that these be made available to allow for such tests. For output and labor, the data requisite can be reasonably satisfied; with respect to capital, it may be very costly to develop sufficiently disaggregated data on the numerous physical capital items to allow for acceptably rigorous applications of the aggregation conditions. Of course there are many data, already developed, that can be made available for aggregation analysis; where the confidentiality rule is not violated, they may be found useful to this end.

In view of the problem of the intractability of the data with respect to capital, and in view of certain theoretical problems, the second recommendation concerns the kind of tests one can reasonably hope to apply. I argue at some length below that composite commodity aggregation is the approach that requires our attention at the moment. For the reasons, I am afraid one has to read on.

7.2 A Brief Review of Reswitching and Capital Aggregation

The possibility of reswitching was originally discovered by Sraffa, who published his results in 1960. Apparently, members of Sraffa's seminars at Cambridge University were aware of the phenomenon well before the results appeared in print. In 1956 Joan Robinson published a version of the reswitching phenomenon called the Ruth Cohen Curiosum. After Sraffa's publication, Samuelson (1962) showed the conditions under which aggregate neoclassical analysis (parable) is possible; these conditions assumed reswitching away. This was related to Champernowne's (1953–54) excellent treatment of chain indexation of capital and since this was the first published demonstration of the reswitching difficulties encountered by aggregate neoclassical-type production analysis, we shall begin there.[1]

Consider an economy in which two commodities are produced in fixed proportions: a consumption good, say corn, produced by means of labor and capital, and capital, produced by means of labor and itself. There are many techniques, and each technique is associated with a particular specification of the capital good. For n heterogeneous capital goods (heterogeneous either in the physical sense as in Samuelson's model or in the sense of different lengths of time required in producing particular capital as in Champernowne's model), the technology of the economy can be described by a book of "blueprints" that is simply a set of the following technique matrixes:

$$\underbrace{\alpha = \begin{array}{c c} & \begin{array}{cc} \text{Corn} & \text{Capital type } \alpha \end{array} \\ \begin{array}{c} \text{Labor} \\ \text{Capital type } \alpha \end{array} & \begin{bmatrix} a_{lo} & a_{l1} \\ a_{1o} & a_{11} \end{bmatrix} \end{array}, \quad \beta = \begin{array}{c c} & \begin{array}{cc} \text{Corn} & \text{Capital type } \beta \end{array} \\ \begin{array}{c} \text{Labor} \\ \text{Capital type } \beta \end{array} & \begin{bmatrix} b_{lo} & b_{l1} \\ b_{1o} & b_{11} \end{bmatrix} \end{array}, \gamma, \delta, \ldots}_{n \text{ techniques}}$$

Let capital be infinitely durable.[2] In a competitive equilibrium there are zero profits and hence the value of the output must equal the cost of production:

(1) $$P_oY_o = W_oL_o + r\,K_{1o}(\alpha)\,P_1(\alpha)$$

(2) $$P_1(\alpha)\,Y_1(\alpha) = W_oL_1 + r\,K_{11}(\alpha)\,P_1(\alpha),$$

where P_o = price of consumption goods,
$P_1(\alpha)$ = price of capital (denoted by subscript 1) good type α,
Y_o = output of consumption good,
$Y_1(\alpha)$ = output of capital good (by subscript 1) type α,
r = rate of profit,
W_o = nominal wage rate,
$K_{1j}(\alpha)$ = amount of capital good type α used in producing one unit of good j, $j = o$, 1, and
L_j = amount of labor employed in producing one unit of good j, $j = o,1$.

Dividing (1) and (2) by Y_o and $Y_1(\alpha)$, respectively, the price equations are obtained:

(3) $$P_o = a_{lo}W_o + r\,a_{1o}\,P_1(\alpha)$$

(4) $$P_1(\alpha) = a_{l1}W_o + r\,a_{11}\,P_1(\alpha),$$

where $a_{lj} = L_j/Y_j$ and $a_{1j} = K_{lj}/Y_j$, $j = o,1$.

Therefore, for the particular technique matrix α, solving (3) and (4) for the wage rate and capital price, both normalized by the consumption good price, gives

$$(5) \qquad W_o/P_o = \frac{1 - a_{11}r}{(1 - a_{11}r)\, a_{lo} + a_{1o}a_{l1}r}$$

$$(6) \qquad P_1(\alpha)/P_o = \frac{a_{l1}}{(1 - a_{11}r)\, a_{lo} + a_{1o}a_{l1}r}\,.$$

Note that, if $m = \dfrac{a_{1o}/a_{lo}}{a_{l1}/a_{11}} > 1$, the consumption section is the more capital intensive, whereas if $m < 1$, the capital sector is more capital intensive. One must keep that in mind for what is to follow.

Equation (5) is the well-known wage curve or wage-profit relationship; (6) relates relative prices to the profit rate. Much of the story turns on the properties of these two equations.[3]

Motivated by Joan Robinson, Champernowne (1953) tried to find a unit in which capital goods can be measured such that the conventional production function can be constructed and marginal productivity theory can be preserved. To do this, he proposed the Divisia type of chain index. An example shows how the chain index of capital is found and how the conventional production function emerges. In the model above, assume that the economy's technology consists of three techniques, where each requires a different capital good.

Figure 7.1*a* depicts the $W_o/P_o - r$ relationship, and figures 7.1*b*, 7.1*c*, and 7.1*d* show the price-profit rate relationships for each technique. The intercepts of the wage curves on the ordinate are $1/a_{lo}(\alpha)$ for the α technique, $1/a_{lo}(\beta)$ for the β technique, and $1/a_{lo}(\gamma)$ for the γ technique; on the abscissa, they are $1/a_{11}(\alpha)$, $1/a_{11}(\beta)$, and $1/a_{11}(\gamma)$. The α technique is more capital intensive (higher a_{11} and lower a_{lo} coefficients) than the β technique, which in turn is more capital intensive than the γ technique. Except at the switch points, r_1 and r_2, economy-wide forces will select that single technique that yields the highest real wage rate for a given profit rate. Thus, in this simple example, α is selected from zero to r_1 profit rate, β from r_1 to r_2, and γ for $r > r_2$. Clearly, as r increases, capital intensity falls. And that is the "well-behaved" case that underlies the aggregate neoclassical postulate.

When one compares (5) and (6) across techniques, one must take care. It is meaningful to compare the $W_o/P_o - r$ relations, but it is illegitimate to compare the price-profit equations across techniques in this model. The basic reason is that W_o/P_o for techniques α, β, and γ have the same dimensions, while $P_1(\alpha)/P_o$, $P_1(\beta)/P_o$ and $P_1(\gamma)/P_o$ have different dimensions.[4]

However, the ratios of the capital values in terms of the consumption good *at equal-profit* points are comparable and this is what is required for the chain index of capital. Their capital values in terms of the price of the consumption good at $r = r_1$ where techniques α and β are equally profitable can be obtained by substituting r_1 into their respective equa-

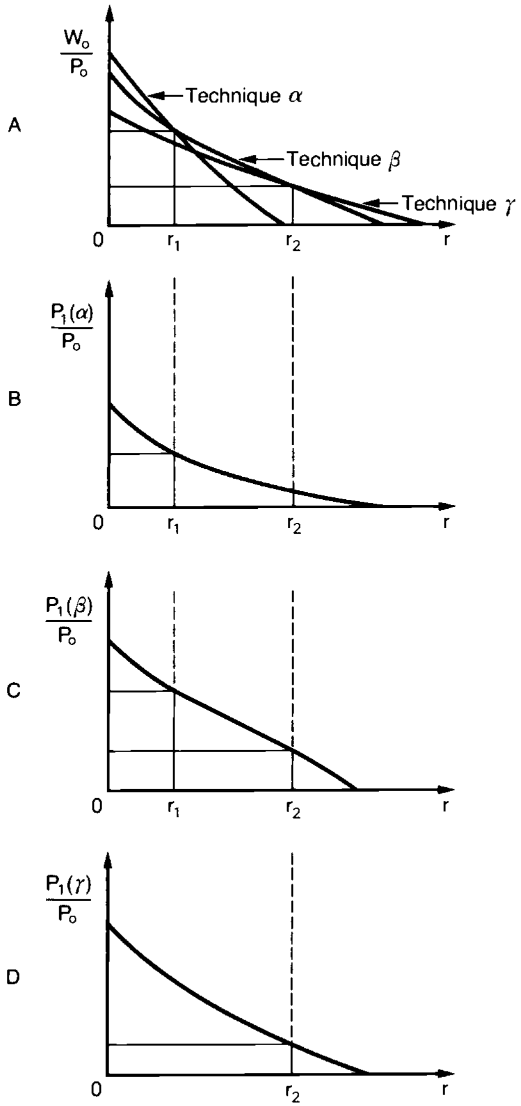
A
$\frac{W_o}{P_o}$
Technique α
Technique β
Technique γ
0
r_1
r_2
r
B
$\frac{P_1(\alpha)}{P_o}$
C
$\frac{P_1(\beta)}{P_o}$
D
$\frac{P_1(\gamma)}{P_o}$

Fig. 7.1

tions (6). Pairs of comparable ratios can be found, and, consequently, a chain index of capital can be erected. Let the base of the index be the real value of γ equipment at r_2, which can be derived from (1) and (2); call it $K(\gamma)$. Suppose the ratio of capital costs $\left[\frac{P_1(\beta)}{P_o}\Sigma K_{1j}(\beta) \Big/ \frac{P_1(\gamma)}{P_o}\Sigma K_{1j}(\gamma)\right]$ of the β to the γ technique at r_2 is 3 : 1, and that the ratio of the capital costs of the α to β technique at r_1 is 6 : 5. Then a chain index of these three heterogeneous capital goods would be $K(\gamma)\cdot(1 : 3 : 3\frac{6}{5})$. Thus, as the interest rate falls, the quantity of capital rises. Champernowne is clearly able to arrange all the alternative techniques of production in a "chain" for some "predetermined" rates of profit (chosen at equal-profit points). Different capitals are larger than others in an unambiguous manner. The conventional production function in which output is expressed as a unique relationship between labor and capital (here representing quantities of different capitals) can be traced out by parametric variations of the profit rates, and one can go on to do straightforward marginal productivity theory.

Of course the example is a special one. Champernowne himself showed that reswitching will destroy the whole sequence. For if the same technique is selected at two different intervals of the profit rate or, stated in another way, if a technique is equiprofitable to another technique at more than two given rates of profit, it is impossible to arrange the alternative techniques in the way required by Champernowne's chain index. (In a different type of model, one can show that if more than one [heterogeneous] capital good is allowed to be used in any technique matrix, then in general there is no way to find such a chain index.)

It is easy to see that reswitching prevents the unambiguous ordering of techniques in terms of capital intensity and the profit rate. The simplest way to show that is to focus on two techniques yielding the two wage curves depicted in figure 7.2. For $0 < r < r_1$, technique α is adopted; for $r_1 < r < r_2$, technique β is the more profitable, and for

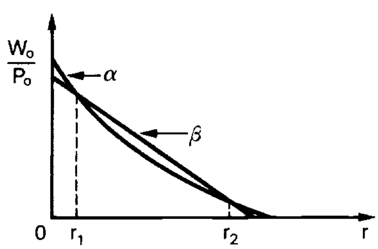

Fig. 7.2

$r > r_2$, technique α comes back or reswitches. Since $1/a_{lo}(\alpha) > 1/a_{lo}(\beta)$ and $1/a_{11}(\alpha) > 1/a_{11}(\beta)$, as the profit rate rises monotonically from $r = 0$ the economy adopts the less capital-intensive technique; but, as r continues to rise, there comes a point where it readopts the less capital-intensive technique. That is one of the reasons Champernowne's index breaks down. Another difficulty—called capital reversal—results when the wage frontier (the envelope of the wage curves) is concave from below. But the reswitching phenomenon is enough to show us that the chain index solution to the capital measurement problem is unacceptable. Note that the reason for the so-called perverse reswitching case is that the coefficient ratio is not unity, or more generally that it is such that it allows two intersections of the wage curves along the frontier.

Samuelson, in his well-known "surrogate production" model (1962), defended aggregate neoclassical production theory. He compared the simple heterogeneous capital model given above (which, as he said, is more realistic) with the neoclassical smooth, malleable-capital model. By a very special assumption that the $W_o/P_o - r$ relation for each technique is linear, the simple neoclassical malleable-capital model, in which output and capital are "jelly," can be a good approximation to the more realistic heterogeneous capital model given above.

Suppose the economy's technology implies the factor-price frontier derived from the wage curves in figure 7.3. Each segment on the frontier is associated with a specific method of production (and therefore a specific capital good). By increasing the number of techniques, a continuous frontier is generated, and hence a continuous switch from one technique to another will be expected as the rate of profit changes. Samuelson then argues that a general good, K, called jelly, can be found such that the factor price trade-off relation generated by the conventional neoclassical production function (with capital jelly as an in-

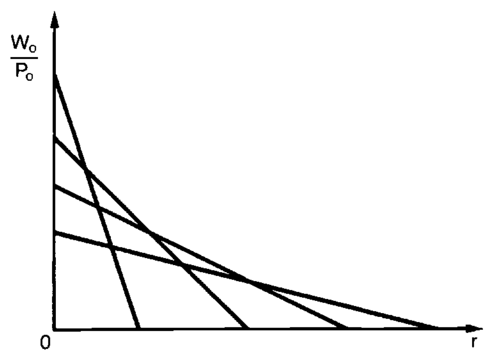

Fig. 7.3 The factor price frontier is the envelope of these wage curves.

put) is a good approximation to the factor-price frontier obtained from the simple heterogeneous capital model. The more realistic model can thus be represented by a neoclassical production function with all the usual aggregate neoclassical properties (i.e., differentiability, positive marginal products, constant returns, etc.). By means of the invisible hand of competition, the marginal product of the capital jelly equals the reward to capital jelly, and the marginal product of labor equals the real wage rate. Duality theory permits one to show the following: $Y \equiv C + PrK = F(L,J) = LF(1,J/L) = Lf(J/L)$; in a perfectly competitive economy, we have

$$W^* = \frac{\partial Y}{\partial L} = f(J/L) - \frac{J}{L} \cdot f'(J/L)$$

$$r = \frac{\partial Y}{\partial J} = f'(J/L),$$

where $W^* = W_o/P_o$; the assumptions of positive marginal products and diminishing or constant returns implies

$$\frac{dW^*}{d(J/L)} = -\frac{J}{L} \cdot f''(J/L) > 0 \tag{7}$$

$$\frac{dr}{d(J/L)} = f''(J/L) < 0. \tag{8}$$

Thus W^* is an increasing function of J/L and r is a decreasing function of J/L. The factor-price relation (trade-off) of the production function can be traced out by parametric variations of J/L. Graphically, this trade-off is given in figures 7.4 *a*, *b*, *c*, and *d*, where figure 7.4*d* clearly mimics figure 7.3. That is, the more realistic heterogeneous capital model can be approximated as closely as we like by increasing the density of techniques, which allows us to employ the neoclassical single-malleable-capital model. Samuelson further shows that the simple Marshallian elasticity of the factor-price frontier is a measure of the distribution of income. By equations (7) and (8), we have

$$\frac{dW^*}{dr} = -\frac{(J/L) \cdot f''(J/L)}{f''(J/L)} = -\frac{J}{L}.$$

Therefore the simple Marshallian elasticity $= -\dfrac{dW^*}{dr}\dfrac{r}{W^*} = \dfrac{J\,r}{LW^*} =$ ratio of relative shares.

Finally, one can show that $C/L = \gamma(r)$ is monotone decreasing, that is, $\gamma' < 0$.

All of Samuelson's aggregation results rest on the assumption of linear factor-price relations. That is, the m ratio must equal unity. (The equality of sectoral factor ratios satisfies the Gorman conditions; see below.) This assumption completely excludes reswitching. Being linear,

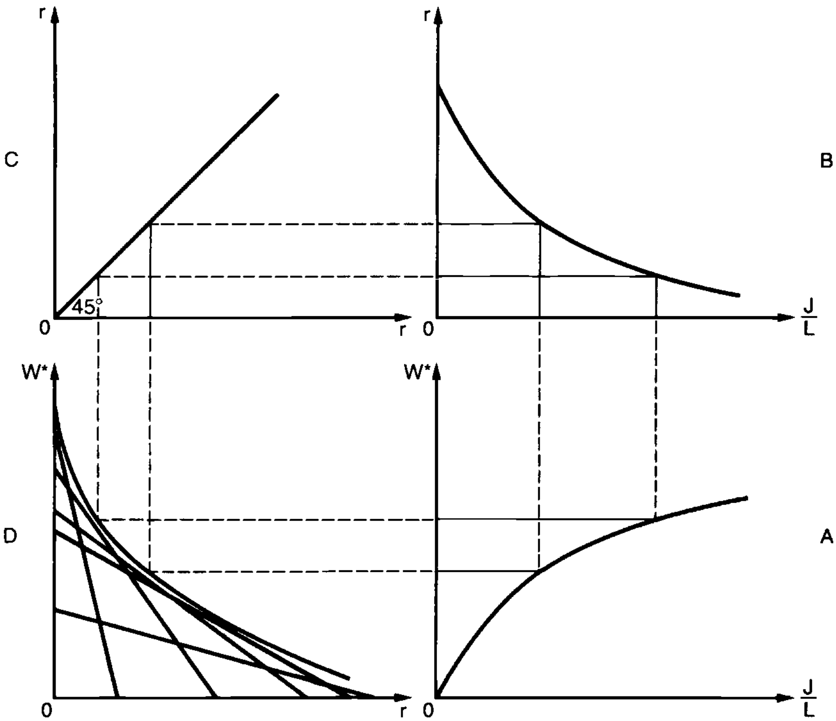

Fig. 7.4

each factor-price curve intersects another at the most only once. The technique will never come back again at different intervals of the rate of profit. The assumption of no reswitching is crucial to the development of a surrogate production function.[5]

Given that assumption, one arrives at the simplest neoclassical (Clarkian) parable, in which there is one homogeneous malleable physical capital (actually, one can measure capital in value terms in this case, but the value capital behaves like a physical quantity), no joint production, and smooth substitutability of labor and the capital aggregate. The marginal productivity relationships determine the functional income distribution and all the other variables in the general equilibrium system upon which the parable is based.

After the Samuelson article appeared, Levhari published a paper that attempted to show that reswitching was not possible in an economy in which the technique matrix is indecomposable. There was a flurry of effective refutations of that theorem, and in November 1966 a symposium in the *Quarterly Journal of Economics* presented them and also forced agreement on a large number of problems. Reswitching and other perversities are potentially present in models containing heterogeneous capital items of the circulating capital or fixed capital type, many consumption goods or only one consumption good, Austrian production processes, Walrasian production processes, decomposable and indecomposable technique matrixes, and smooth as well as discrete technologies. Reswitching, however, is associated only with discrete technologies, but other perversities such as capital reversal are relevant to smooth production technologies.

The second phase of the so-called reswitching controversy (at this point it is no longer a controversy in the literal sense) was taken up with spelling out the nature of the phenomenon. In 1969 I showed that, in a model of the type given above, if the technology is such that the substitution effects between labor and capital outweigh the change in composition or the change in the weighting of the two sectors, then a general type of perversity cannot occur (also see Brown 1973). This result has been confirmed by Hatta (1974) and by Sato (1976*b*) using a more general model. Burmeister (1977) focuses on the concept of a regular economy showing that it is necessary and sufficient to preclude paradoxical aggregate consumption behavior. The duality between the wage frontier and the technology frontier has been investigated (Sato 1974; Burmeister and Kuga 1970; Bruno 1969). Finally, different types of models have been examined; these range from different characterizations of steady-state models (Cass 1976; Zarembka 1976) to dynamic models (Oguchi 1977).

7.3 The Implications

One way to spell out the implications of what has been presented above is to compare the neoclassical parable to the intertemporal general equilibrium model containing many heterogeneous capital goods (Samuelson 1976; also see Nuti 1976). The following is a list of some steady-state properties of the neoclassical parable, some of which have been indicated above but do not hold generally:

a) $-\partial C_{t+1}/\partial C_t = 1 + r_t$,
b) $\partial^2 C_{t+1}/\partial C^2_t \leq 0$,
c) $W_o/P_i = f_i(r) = f(r)$, factor-price frontier trade-off,
d) $r = f'(K/L)$, marginal productivity, $f'' < 0$,
e) $C/L = \gamma(r)$, monotone decreasing, $\gamma'(\ ) < 0$,
f) $C/L = \theta(K/L)$, monotone increasing, $\theta'(\ ) > 0$,
g) K/Y or capital-output ratio declining with profit rate,
g') K/L declining with profit rate,

h) no reswitching possible,
i) no capital reversals,
j) elasticity of (r,w) frontier $=$ wage share/profit share.[6]

Clearly, not all of these hold generally. It has been stressed repeatedly that (*h*) does not hold in general and therefore the neoclassical parable goes by the way. But (*a*), (*b*), and (*c*) do hold up in very general circumstances. Even if joint production is present, one can still accept the wage-profit trade-off that is dual to the consumption-growth rate relation just as it is in the nonjoint production case (Burmeister and Kuga 1970). Continuing, (*e*) does not hold in general, nor do (*f*), (*g*), and (*g'*). The neoclassical parable and its implications are thus generally untenable.[7]

What does this mean for those who want to measure capital at various levels of aggregation? If the conditions for no reswitching and no capital reversal ($m = 1$ covers both, but the conditions, $m \neq 1$ and the wage-profit frontier concave from below, permit capital reversals), then the capital aggregates are unstable. This means they are not invariant to changes in relative prices (Brown 1973). One may construct them as is usually done, but it is unlikely that they do not change with changes in the profit rate as Robinson has noted. Of course, that in turn means that the production function estimated on the basis of those capital aggregates is no longer a physical-technical relationship, for it now contains market variables. One cannot have much confidence in predictions based on such an unstable relationship.

7.4 Separability, Duality, Price, and Quantity Capital Indexes

We begin the discussion of the conditions underlying capital aggregates with those that require restrictions on functional form. For most of the exposition, we need treat only two sectors, in each of which there are three factors of production, two physically heterogeneous kinds of capital (x_{1j} and x_{2j}; $j = 1,2$), and labor (x_{oj}; $j = 1,2$). The original statement of this type of aggregation is attributable to Leontief (1947). The theorem is applicable to a partial equilibrium approach (analyzing the behavior of a single sector while treating the other sectors as exogenous) as well as to a general equilibrium analysis (in which feedback effects are permitted between sectors). In all the models, the capital goods are thought to be produced within the economy. They are akin to intermediate goods, but they are not "netted out" as is often done with inputs of materials. In many applications, the latter are indeed netted out so that these models refer to value-added magnitudes. Of course, as will become clear, the aggregation theorems based on the

Leontief results can encompass all types of goods. Finally, we abstract from depreciation and joint production in the initial exposition, returning to it briefly at a later point.

Suppose we focus on two production functions:

$$y_j = f^j\,(x_{0j}, x_{1j}, x_{2j}), \quad j = 1{,}2, \tag{9}$$

where y_j are the outputs of the two sectors which we can take to be value-added measures for the moment. The functions f^j can be taken to have strictly positive marginal products (i.e., $f^j_i = (\partial f^j/\partial x_{ij}) > 0$; $i = 0{,}1{,}2$. For the Leontief theorem, the production functions f^j are taken to be strictly quasi-concave over the economic region.[8] They can be characterized by any degree of returns to scale; the freedom allowed by the Leontief theorem in this respect is one of its main advantages.

The Leontief theorem itself simply states that the necessary and sufficient condition to write f^j in equation (9) as

$$y_j = f^j(x_{0j}\,, x_j), \quad j = 1{,}2, \tag{10}$$

where $$x_j = g(x_{1j}\,, x_{2j}),$$

is that

$$\frac{\partial}{\partial x_{0j}}\left[\frac{\partial f^j}{\partial x_{1j}} \bigg/ \frac{\partial f^j}{\partial x_{2j}}\right] = 0. \tag{11}$$

(For a simple proof. see Green 1964.) This condition, meaning that the marginal rate of substitution between the capital items is independent of labor, is called weak separability.[9] Note that it allows for aggregation of capital inputs within each sector; in other words, it permits intrasectoral aggregation.

Since weak separability is the basis for many of the aggregation results in this particular area of aggregation theory, it is worthwhile to interpret its meaning here. In the first place, it requires that changes in the labor (or any noncapital input) not affect the substitution possibilities between the capital inputs. Suppose labor input is ten, and the two capital substitution possibilities are, say, three x_1 to one x_2 and two x_1 to three x_2, both combinations yielding one hundred units of output. Now let labor input increase to twenty, which, combined with the same capital ratios, yields two hundred units of output. In this case the Leontief condition holds. (This example is based on Green 1964, pp. 11–12.) As Solow indicates (1955, p. 103), the condition will not often be satisfied, even approximately, in the real world. Some examples such as brick buildings and wooden buildings or aluminum fixtures and steel fixtures turn out to be cases where the capital items are homogeneous except in name. For more complex cases—bulldozers and trucks or sound amplification equipment and desks in a classroom—the technical substitution possibilities will probably depend on the amount of the labor input.

Yet there is a class of situations, according to Solow, in which the weak separability condition may be expected to hold. Suppose the production of y_j can be decomposed into two stages, one in which something called x_j is produced out of x_{1j} and x_{2j}, alone, and the second stage requiring this substance in combination with labor x_{0j} to produce y_j. More specifically, suppose that the "production" of x_j is given by

$$x_j = g^j(x_{1j}, x_{2j}); \tag{12}$$

for example, if x_{1j} and x_{2j} are two kinds of electricity-generating equipment and x_j is electric power, then generating capacity would be an index of the capital inputs. Clearly, the functions g^j in (12) are capital index functions, and it is important to know their properties. One way to do that is to follow Solow's article, where he shows that the g^j functions are linearly homogeneous (given that the F^j functions are linearly homogeneous and that the weak separability condition applies). Green (1964, chap. 4) does the same; but now an additional problem must be considered.

Examine (10) again, and see that the three factors of production are partitioned into two groups, a labor "group" x_{0j} and a capital group x_j. When there are only two groups, the weak separability condition is sufficient to allow for that decomposition and to yield price and quantity indexes for each group.[10] That is, if there are only two groups, it is sufficient (see Green 1964, p. 21) for there to exist a quantity index (12) in each sector and a sectoral capital price index: $p_{x_j} = p_{x_j}(p_{1j}, p_{2j})$. Moreover, it can be shown that, if the production function is homothetic,[11] the expenditure on the capital aggregate in each sector is $p_{x_j} x_j$, which, when added to the expenditure on the labor input, $p_{0j} x_{0j}$, adds up to total expenditure.

But when there are more than two groups, and of course that is probably the case, weak separability is no longer sufficient. Strotz (1959) and Gorman (1959) show that not only must the weak separability condition hold, but, in addition, each quantity index must be a function homogeneous of degree one in its inputs. These conditions, called homogeneous functional separability by Green (1964, p. 25), are necessary and sufficient[12] for each group expenditure to equal the sum of the expenditures on each item in the group; that is,

$$E^j{}_r = p_{i_r j} \, x_{i_r j}, \quad r = 1,2, \ldots, S,$$

where S is the number of groups into which the factors of production are partitioned.

It is customary to prove the above results by using duality theory. (See Shephard 1953.) Let us partition the inputs of (9) into labor and capital groups for each sector; that is, let x_{0j} and g^j (x_{1j}, x_{2j}) be the two groups.[13] Suppose that the g^j are homogeneous functions (they are quantity index functions) and that corresponding price indexes (homogeneous functions of prices) can be specified: $p_{x_{0j}}$ *and* p_{x_j} (p_{1j}, p_{2j}). Then, following Shephard (1953), the following aggregation conditions must apply:

(*a*) $\sum_{i=0}^{3} p_{x_{ij}} x_{ij} = p_{x_{0j}} x_{0j} + p_j g^j$

(*b*) $F^j(x_{0j}, x_{1j}, x_{2j})$ can be expressed as $F^j[x_{0j}, g^j(x_{1j}, x_{2j})]$, where g^j are homogeneous functions.

(*c*) Minimum cost, C^j, can be expressed as a function, $C^j(y_j, p_{x_{0j}}, p_{x_j})$; that is, as a function of the sectoral output rate and the price indexes.

(*d*) The aggregate cost function, $C^j(y_j, p_{x_{0j}}, p_{x_j})$ may be derived from the aggregate production function, $F^j(x_{0j}, g^j(\cdot))$ as $C^j(y_j, p_{x_{0j}}, p_{x_j}) = \min_{x_{0j}, g^j} (p_{x_{0j}} x_{0j} + p_j g^j)$, where g^j is given above and the prices are taken as parameters from the firm or sector's point of view.

If these conditions apply, then clearly each sector need concern itself with only two factors of production, and one can obtain all the information from the two-factor formulations that one does from the formulation involving all the elementary factors (in our simple exposition, there are only three).

The aggregation problem is solved if the production functions are such that F^j are arbitrary increasing functions of x_{0j} and g^j (j = 1,2), and the g^j are homogeneous of degree one in their respective arguments; in other words, the production functions are homothetic.[14]

Duality between cost and production functions is involved precisely here, yielding information on the implied price indexes. For it is one of the enduring results of duality theory that if the production function in each sector is homothetic, then the sector's cost function factors into a product of $f^j(y_j)$, which is the inverse function of F^j (recall that the F^j are assumed to be monotonically increasing and also assume that $F^j(0) = 0$), and a function

$$\Gamma^j(p_{x_{0j}}, p_{x_{1j}}, p_{x_{2j}})$$

that is homogeneous of degree one in the prices; that is,

$$C^j(y_j, p_{x_{0j}}, p_{x_{1j}}, p_{x_{2j}}) = f^j(y_j) \cdot \Gamma^j(p_{x_{0j}}, p_{x_{1j}}, p_{x_{2j}}).$$

This considerably simplifies the cost function; it is worth repeating that, to do this, homotheticity of the production function and the independence of prices and quantities are required.

Using this well-known result, it is a simple matter to form subindexes for the two groups. Thus, in terms of costs for each group:

$$\min_{x_{0j}} (p_{x_{0j}} x_{0j}) = x_{0j}\Gamma^{0j}(p_{x_{0j}})$$

(13) and

$$\min_{x_{ij}} \left(\sum_{i=1}^{2} p_{x_{ij}} x_{ij} \right) = x_j \Gamma^{1j} (p_{x_{1j}}, p_{x_{2j}}).$$

where, recall, x_j is given by $g^j(x_{1j}, x_{2j})$ for g^j homogeneous of degree one; and Γ^{0j} and Γ^{1j} are homogeneous functions of degree one. Putting the two together, we can write:

$$C^j(y_j, p_{x_{0j}}, P_{x_{1j}}, p_{x_{2j}}) =$$

(14)

$$= \min_{x_{0j}, x_j} [x_{0j}\Gamma^{0j}(p_{0j}) + x_j\Gamma^{1j}(p_{1j}, p_{2j})],$$

where x_{0j} and x_j are restricted by

$$y_j = F^j(x_{0j}, x_j).$$

As Strotz and Gorman have demonstrated, the procedure for obtaining subindexes can be thought of as occurring in two stages: the first minimizes total costs by choosing the optimal proportion of each group of factors, whereas the second stage involves the minimization problems for each of the subgroups in (13) in which group costs are minimized separately given total costs.

The result is that $g^j(\cdot)$ and $\Gamma^{1j}(\cdot)$ are quantity (of capital) and capital price index numbers—they are the aggregates we seek—that simultaneously accomplish four things: the first is that they reflect the optimal inputs obtained from minimizing cost with respect to homothetic production surfaces; second, they are generalized index numbers that satisfy three fundamental Fisherian properties;[15] third, they satisfy the aggregation conditions (a)–(d), and, finally, they are consistent with the two-stage Gorman-Strotz optimization procedure. This is an extraordinary list of accomplishments, obtained at the cost of two seemingly harmless assumptions.

But there are limitations, and they are not negligible. The basic limitation of the duality theory and the resulting indexes can be seen from a simple example. Consider a firm using only two factors of production, x_0 and x_1, whose production follows the homogeneous of degree one CES form, which is obviously homothetic; that is,

$$y = \gamma(\kappa_0 x_0{}^{-a} + \kappa_1 x_1{}^{-a})^{-\frac{1}{a}} .$$

The first-order conditions can be written in marginal rate of substitution form:

$$\frac{\kappa_1}{\kappa_0}\left(\frac{x_0}{x_1}\right)^{1+a} = \frac{E_1}{E_0}\frac{p_1}{p_0},$$

where $E_i = 1 + 1/e_i$, e_i is the elasticity of supply of the ith factor, and the p_i are the factor prices. The cost function is

$$C = \left(\frac{1}{\gamma}y\right)\left[\kappa_0 p^a{}_0 \left[1 + \underline{a}EP^{-\frac{a}{1+a}}\right]^a + \kappa_1 p^a{}_1 \left[1 + \underline{a}^{-1}E^{-1}P^{\frac{a}{1+a}}\right]^a\right]^{\frac{1}{a}},$$

where $\underline{a} = \kappa_0/\kappa_1$, $E = E_0/E_1$, and $P = p_0/p_1$. Clearly, if $\partial E/\partial x_0 = \partial E/\partial x_1 = 0$, the cost function factors into two expressions, one in output that is homogeneous of degree one and the other in the p_i that is also homogeneous of degree one (note that P is homogeneous of degree zero in the p_i).

However, if factor-supply elasticities are related to the quantities of the factors themselves, and hence to the output, then the cost function does not factor into two terms that are homogeneous of degree one. This means, inter alia, that price and quantity of factor input indexes computed as expressions homogeneous of degree one misspecify the actual price and quantity changes, not because of the usual index number problems but because of the distortions introduced by imperfect competition in factor markets. (It can be shown that quantity output and price indexes would suffer a similar fate as a result of the presence of imperfect competition on the output side.) One can expect this to occur in those industries largely controlled by few firms, in time periods over which the factor supply elasticities are likely to change, and between firms in industries largely controlled by a few firms that coexist with smaller firms.

Suppose that industry price and quantity indexes homogeneous of degree one are constructed and that an analyst, using that data, aims to test hypotheses related to the degree of imperfect competition in that industry. That is, the data are constructed on the assumption that the firm or industry is competitive in factor markets,[16] and the analyst uses that data to test the degree to which that firm or industry is competitive. Clearly, the outcome must be biased. Or suppose a productivity analysis were undertaken using price and quantity indexes constructed as above; the productivity measure is clearly affected.

A practical difficulty with the approach based on weak separability and homotheticity is that it requires microdata on physical inputs and outputs to test the aggregation conditions. We do not have measures in physical units of the numerous capital items that enter production processes at even the most disaggregated level of production. But, even were they available, it may be difficult to accept the assumption of competition that underlies the construction of this type of aggregate.

7.5 Fisher's Extensions of Functional Form Conditions: Intersectoral and Intrasectoral Aggregation

Perhaps the most extensive analysis of aggregation conditions focusing on functional form has been done by Frank Fisher (1965; 1968*a*,*b*; 1969). Not only does he consider capital, labor, and output aggregation, but he also includes the difficult problems of fixed and movable factors.

Fisher introduces optimizing conditions for the economy into aggregation analysis. Thus, suppose the production functions are

$$y_j = F^j(x_{0j}, x_{1j}), \tag{15}$$

where capital may differ from firm to firm and for simplicity all firms' outputs are indistinguishable and there is only one kind of labor. Under what condition is it possible to write total output Y as given by the aggregate production function:

$$Y = \sum_j y_j = F(x_0, x_1), \tag{16}$$

where $x_0 = x_0(x_{01}, x_{02}, \ldots, x_{0n})$ and $x_1 = x_1(x_{11}, x_{12}, \ldots, x_{1n})$ are indexes of aggregate labor and capital, respectively. This, then, is solely a question of intersectoral aggregation.

If only restrictions on functional form were considered, the necessary and sufficient conditions for intersectoral aggregation are that every firm's production function be additively separable in capital and labor; that is, each F^j be of the form: $F^j(x_{0j}, x_{1j}) = Q^j(x_{0j}) + \Psi^j(x_{1j})$. That these conditions are extremely restrictive has been noted by Fisher and others.

Here Fisher notes that these conditions are answers to the wrong question. A production function, he states, describes the maximum level of output that can be achieved if the inputs are efficiently employed. Accordingly, one should ask not for the conditions under which total production can be written as (16) under any economic conditions, but rather for the conditions under which it can be written once production has been organized to get the maximum output achievable with the given factors. Thus, efficient production requires that Y be maximized given x_0 and x_1, a circumstance that introduces allocative decisions into the aggregate procedure. If one wishes to analyze production within a market system (or a centrally controlled one), then it does not seem reasonable to ignore the optimizing conditions for aggregation purposes.

Suppose in the simplest case that labor is movable and only labor can be allocated to firms to maximize total output. Letting Y^* be that maximal output, one can evidently write $Y^* = G(x_0, x_{11}, x_{12}, \ldots, x_{1n})$, there being no labor aggregation problem, since the values of x_{0j} are determined in the course of the maximizing procedure. The entire problem is the existence of a capital aggregate. Recalling that the weak separability condition (that MRS between x_{1i} and x_{1j} be independent of x_0 in G) is both necessary and sufficient for the existence of a group capital index, Fisher proceeds to draw the implication of this condition for the form of the original firm production functions in (15). He finds that under the assumption of strictly diminishing returns to labor ($F^j_{0j,0j} < 0$; $j = 1,2, \ldots, n$; where the subscripts denote differentiation), a necessary and sufficient condition for capital aggregation is that every firm's production function satisfy a partial differential equation in the form $F^j_{0j,1j}/F^j_{1j}F^j_{0j,0j} = g(F^j_{0j})$, where g is the same function for all firms. Further, assuming constant returns to scale, capital augmenting technical differences turns out to be the only case under constant returns in which a capital aggregate exists.

This means that each firm's production function be written as $F^j(x_{0j}, x_{1j}) = F^1(x_{0j}, b_j x_{1j})$, where the b_j are positive constants. Such a requirement is highly restrictive, since a different capital good is equivalent in all respects to more of the same capital good. For example, sound amplification equipment in a classroom is considered to be three times the number of desks in the same classroom. One requires a very complicated transformation scheme that somehow allows the varied and myriad capital goods to be accounted for in the same units.

The capital augmentation result and the notion of capital generalized constant returns[17] are important contributions of Fisher's analysis. He utilizes these basics in more complicated models, some of which are discussed below. But the general message that comes out of the work is that the conditions for output, capital, and labor aggregation are unlikely to be satisfied exactly.

Are they likely to be approximated? All we really care about is whether aggregate production functions provide an adequate approximation to reality in terms of the empirical values of the output, labor, and capital variables. Thus, for approximate capital aggregation it would suffice for technical differences among firms to be approximately capital augmenting.

But this is not a useful result. "The reason for being unhappy with capital aggregation, for example, is not merely that one thinks technical differences are not likely all to be exactly capital augmenting but that one thinks there are some differences that are not anything like capital augmenting" (Fisher 1969, p. 570). The interesting question is whether an aggregate production function gives a satisfactory approximation in a bounded region defined by the empirical values of capital and labor. Clearly, one must define what one means by a satisfactory approximation and also decide how badly the conditions are violated.

Fisher arrives at a generally negative conclusion: it appears that the only way to accept such approximations would be to admit certain well-defined irregularities in production functions, irregularities that are not exhibited by the aggregate production function in practice. Such an escape from the stringent conditions for aggregation will be available, if at all, only in rather special cases.

In view of this, Fisher asks why production functions with parameters estimated from factor payments turn out to fit input and output data so well. Since the matter is too complicated to treat analytically, he suggests experimenting with constructed data in which the aggregation conditions are known not to be satisfied. Aggregate production functions are then estimated on these data (in the latest study, the CES is used; see Fisher, Solow, and Kearl 1974). It turns out that the aggregate Cobb-Douglas predicts wages well whenever labor's share is roughly constant; with the CES, generalizations are more difficult to obtain. In spite of the special nature of the constructed data (all micro-units exhibit constant returns to scale), several other suggestive results emerge from this Monte Carlo type of study: composition effects can seriously distort aggregate elasticity of substitution estimates; and the wage equation estimates are more reliable than the production function estimates, though combining the two allows one to track output and factor prices closely. Thus aggregate production functions can work in special cases. And that is precisely their problem. We would require a catalog of unknown proportions to indicate their areas of applicability. Even then, one could not allay the doubt that the results are special in one way or another, and it may be difficult to specify which way it is.

7.6 The Gorman Aggregation Conditions

The Gorman (1953) conditions are developed along lines similar to those followed by Fisher. It is assumed that the optimal conditions for the distribution of given totals of moveable inputs among firms are satisfied. These efficiency conditions (which imply Pareto optimality) require that the marginal rates of substitution (MRS) between the ith and jth factors be the same for all firms:

$$\frac{\partial F^i}{\partial x_{ki}} \Big/ \frac{\partial F^i}{\partial x_{hi}} = \frac{\partial F^j}{\partial x_{kj}} \Big/ \frac{\partial F^j}{\partial x_{hj}}, \qquad \begin{array}{c} i,j = 1,\ldots,n; \\ k = 1,\ldots,n; \\ h = 1,\ldots,m. \end{array} \tag{17}$$

If all firms use some labor input, the given totals of the factors of production must be well-defined aggregates:

$$x_r = \sum_{s=1}^{n} x_{rs}. \tag{18}$$

These, together with (17) imply a transformation surface: $G(y_1, y_2, \ldots, y_n; x_0, x_1, \ldots, x_m)$.[18]

Given that (17) holds and that isoproduct surfaces are convex, then Gorman shows that intersectoral aggregation of the production functions (equation 9) requires that the expansion paths for all firms be parallel straight lines through their respective origins. There will then exist functions, h^1, and h^2 and F, such that $Y = \sum_{j=1}^{2} h^j(y_j) = F(x_0, x_1, x_2)$, where F is homogeneous of degree one in its inputs. Hence, each F^j will be expressible as a function of F. Also, if the expansion paths are required to be parallel, the optimal ratios of the factors must be the same for all firms. An example may be useful here.

Suppose the F^j were CES, that is,

$$y_j = \left[b_{1j} x_{1j}^{(\sigma_j - 1)/\sigma_j} + b_{2j} x_{2j}^{(\sigma_j - 1)/\sigma_j} + b_{3j} x_{0j}^{(\sigma_j - 1)/\sigma_j} \right]^{\sigma_j/(\sigma_j - 1)},$$

then the marginal rates of substitution equilibrium conditions would be

$$\frac{\partial F^j}{\partial x_{1j}} \Big/ \frac{\partial F^j}{\partial x_{2j}} = \frac{b_{1j}}{b_{2j}} \left[\frac{x_{2j}}{x_{1j}} \right]^{1/\sigma_j} = \frac{p_1 (r + \delta_1)}{p_2 (r + \delta_2)} \tag{19}$$

$$\frac{\partial F^j}{\partial x_{1j}} \Big/ \frac{\partial F_j}{\partial x_{0j}} = \frac{b_{1j}}{b_{3j}} \left[\frac{x_{0j}}{x_{1j}} \right]^{1/\sigma_j} = p_1 (r + \delta_1), \quad j = 1,2, \tag{20}$$

where the b_{ij} and σ_j are constants, the prices are normalized by the wage rate, and δ_j are constant declining-balance type depreciation coefficients. It is readily seen that the expansion paths are straight lines through their origin; moreover, the two conditions in (19) and the two in (20) imply parallel expansion paths if $b_{11}/b_{21} = b_{12}/b_{22}$, $b_{11}/b_{31} = b_{12}/b_{32}$, $\sigma_1 = \sigma_2$. Thus, under these conditions, intersectoral aggregation is possible. Note that, in the example above, satisfaction of the conditions entails that the production function F^1 has the form AF^2 with A an arbitrary positive constant. The two capital goods can be regarded as identical except for a choice of units. This ensures the feasibility of aggregation, but the requirement is so stringent that it is not likely to be satisfied in practice.

7.7 Economywide and Sectoral Weights in Divisia Input Price Indexes: The Gorman Conditions Again

The Gorman conditions turn up in unexpected places,[19] and one of those is in the weights on the Divisia indexes of capital inputs in a sectoral context. I shall show here that the practice of using economywide deflators to obtain real capital measures within a sector requires that the Gorman conditions be satisfied for all sectors in the economy. That is a patent impossibility, and hence that procedure involves a misspecification of unknown proportions.

Consider the value of capital used in the jth sector:

$$v_j = \sum_{i=1}^{2} q_i x_{ij}, \quad j = 1,2. \tag{21}$$

Take its total differential[20] and express it in relative terms:

$$\hat{v}_j = \sum_i w_{ij}\hat{q}_i + \sum_i w_{ij}\hat{x}_{ij}, \quad j = 1,2, \tag{22}$$

where the "hatted" variables represent relative changes, that is $\hat{v}_j = dv_j/v_j$, and so on, and $w_{ij} = q_i x_{ij}/v_j$, which is simply the costs of the ith capital item in the jth sector as a proportion of the sector's total capital costs.

The two components of $\hat{v}_j$ in equation (22) are called Wicksell effects, the first being the price Wicksell effect (PWE) while the second is the real Wicksell effect (RWE).

Suppose the two capital items in the jth sector (say, shearing machines and lathes) are to be treated as an aggregate. For several reasons (see Usher 1973, chap. 7), one must start with the value magnitudes (21) and (22) and derive the real aggregate from them. Referring to (22), it is thus necessary to eliminate the PWE. This is usually done by deflating the value of the capital aggregate (i.e., v_j) by a Divisia or chain index. In relative change terms, a commonly used index is

$$\hat{q}^* = \sum_{h=1}^{2} w_{h.}\hat{q}_h, \tag{23}$$

where

$$w_{h.} = \frac{\sum_{j=1}^{2} q_h x_{hj}}{\sum_{j=1}^{2} v_j}.$$ [21]

The $w_{h.}$ are (possibly) changing Divisia weights; that is, $w_{h.}$ represents the economywide costs of the hth capital item as a proportion of total costs of all capital items. Note that $\hat{q}^*$ is an economywide measure that corresponds to an economywide Divisia or chain index. If the Bureau of Labor Statistics (BLS) wholesale price index (WPI) (or some variant of it) is used, an economywide input index is implied.[22] However, note that relative changes in the BLS WPI and $\hat{q}^*$ differ unless all depreciation rates are zero or the same.

Now, "deflate" (22) by (23)—that is, deduct $\hat{q}^*$ from $\hat{v}_j$; this yields $\hat{x}^*_j$, say: $\hat{x}^*_j = \sum_i (w_{ij} - w_{i.})\hat{p}_i + \sum_i w_{ij}\hat{x}_{ij}$. Recall that the deflation procedure, to be successful, must make the PWE vanish, leaving only $\sum_i w_{ij}\hat{x}_{ij}$. This implies that $\sum_i (w_{ij} - w_{i.})\hat{p}_i = 0$. Since $\hat{p}_i > 0$,[23] a necessary and sufficient condition for the PWE to vanish by this deflation procedure can be shown to be $x_{11}/x_{21} = x_{12}/x_{22}$. In turn, this can be shown to be identical to the Gorman conditions (parallel, straight-line expansion paths) if the production functions are homogeneous of degree one.

How does one interpret this result? Someone analyzing production in a single sector that uses two types of capital to produce it deflates the total cost of these two capital inputs by a price index of the two items that contains weights representing the proportions of costs of each item in the total costs of all capital produced. In doing that, the analyst has assumed (whether knowingly or not) that the Gorman conditions are satisfied. Clearly, they cannot be satisfied in realistic situations, and hence the PWE is not eliminated. A price effect remains in the "real capital aggregate," and every function specifying that aggregate must be unstable. (Clearly, one can isolate the direction of the resulting bias by an analysis of the sectoral and economywide weights.) Though the result above is subject to several qualifications,[24] one arrives at the discomforting conclusion that using an economywide index to deflate capital costs within a sector to derive a real measure of capital almost inevitably fails to purge the price Wicksell effect completely, and thus the resulting data fluctuate with prices. Since data estimation procedures are often used to derive data on which production functions are estimated, misspecifications are bound to be present.

7.8 Houthakker-Sato Aggregation

In a paper having a succès d'estime,[25] Houthakker (1955–56) found a way around the difficulties encountered by Solow-Fisher and Hicks by postulating that factor proportions are distributed in a certain way among the firms over which the aggregation is to take place. The introduction of the distribution function is novel, though there is an analogue from consumption theory on the distribution of income among consumers (see Katzner 1970, pp. 139 ff.). In subsequent work, the Houthakker idea was taken up by Levhari (1968) and Sato (1975), the latter developing it very fully.

In Houthakker's paper, each firm is assumed to operate under two-factor (labor and capital) fixed coefficients production conditions:

$$y_j = \alpha_j k_j = \beta_{0j} x_{0j}, \tag{24}$$

where k_j is the jth firm's capital-labor ratio. Efficiency conditions require that the firm is above its shutdown point if its quasi rent is nonnegative; that is, $y_j - p_0 x_{0j} = y_j(1 - p_0/\beta_j) \geq 0$ or $\beta_j \geq p_0$, where the labor input is taken to be homogeneous among firms so that all firms face the same wage rate. The distribution of capacity output of the firms is determined by the α_j and k_j and one can define a capacity distribution function as

$$\phi(\beta) = \sum_{\alpha} \alpha k(\alpha,\beta). \tag{25}$$

The right-hand side is clearly the total productive capacity of firms with the labor efficiency level of β. To find the total productive capacity over all firms, one must integrate over β; thus

$$Y(p_0) = \int_{p_0}^{\beta_0} \phi(\beta) d\beta, \tag{26}$$

and total employment is

$$L(p_0) = \int_{p_0}^{\beta_0} \frac{\phi(\beta)}{\beta} d\beta, \tag{27}$$

where β_0 is the supremum of β (clearly, the β are taken to be bounded from above). Suppose the density function follows a Pareto distribution:

$$\phi(\beta) = C\beta^{-1/(1-a)}. \tag{28}$$

Inserting this into $Y(p_0)$ and $L(p_0)$ and eliminating p_0 from the two equations, one obtains the aggregate production function:

$$Y = J^{1-a} L^{a}, \quad 0 < a < 1, \tag{29}$$

where the aggregate capital, J, can clearly be found from

$$J = \int_{0}^{\beta_0} \phi(\beta) d\beta. \tag{30}$$

Thus, in the Houthakker model, if all firms operate according to Leontief production functions and if the β_j are distributed according to Pareto, the aggregate production function is Cobb-Douglas. Clearly, the weak separability property (in any of its variants) is unnecessary here.

Sato's procedure is only slightly different, but it yields far more general results. He begins with the micro production functions and the productive capacity function associated with the labor coefficient. He then derives necessary and sufficient conditions for the existence of an aggregate production function with capital and labor aggregates. That is, the following equalities must hold:

$$(31)\quad \begin{aligned} &i)\ L(p_0) = \int_{p_0}^{\beta_0} \frac{\phi(\beta)}{\beta} d\beta = \frac{1}{\beta_0} H\ \frac{p_0}{\beta_0}\ J \\ &ii)\ Y(p_0) = \int_{p_0}^{\beta_0} \phi(\beta) d\beta = G[H\ \frac{p_0}{\beta_0}\ ]J, \end{aligned}$$

where J is given by (30) and the H and G functions satisfy the middle equalities of equations (31.i) and (31.ii), respectively.

Using this procedure, a host of results can be obtained. The micro production functions can now be allowed to have elasticities of substitution exceeding zero, and the distribution functions need no longer be of the Pareto form. Levhari (1968) had already derived an aggregate CES function using the Houthakker procedure, specifying zero elasticity of substitution micro production functions. Sato is able to treat this and the original Houthakker result as special cases of his more general approach.

The aggregate production function derived in this manner is a short-run relationship, since it describes the employment-output relation given the efficiency distribution. If the efficiency distribution shifts, the short-run aggregate production function also shifts, but the resulting factor proportions may not lie along the ex ante production function.[26] Generally, one considers the elasticity of substitution of the ex post or clay production process to be less than that of the ex ante production function. Sato shows (1975, pp. 134 ff.) that, if the efficiency distribution is stable in form, the resulting estimates should reveal the ex ante production function. Thus, the burden of the analysis that generates the desired aggregates shifts from the underlying production functions themselves to the stability of the distribution function.

Is the distribution function inherently unstable when the variables vary? (Sato's estimates, 1975, p. 205, are not uniformly acceptable.) Do firms entering the industry have the same distribution of productive capacities as those leaving? (See Sato 1975, p. 30.) Clearly, the presence of nonneutral technical change implies a change in the slope of the Pareto curve, since capacity will be added at the low end of the scale of input ratios (see Sato 1975, p. 140). At the very least, the estimation problems associated with the distribution function are just as formidable as those of the production function itself. Moreover, recall that one must estimate the distribution function in addition to the production function, thus compounding the difficulties.

There is one further estimation problem with the Houthakker-Sato approach that requires some discussion. The distribution of productive capacities does not appear to be independent of the macro production function. The disturbances on each of the econometric forms are probably correlated (certainly shocks in the distribution function affect aggregate output); thus there is a simultaneous equation estimating problem that differs from that treated in the literature on error specification in production models (see, e.g., Zellner, Kmenta, and Dreze 1966). This problem does not appear to be recognized, much less resolved. One may wish to classify the simultaneity problem as another practical difficulty (see Sato 1975, pp. 201–2).

Glancing back at (24) and (29), one notices that the original Houthakker problem was the intersectoral aggregation of two-factor production functions. When more than two factors are considered, one has to invoke the familiar separability conditions (Sato 1975, pp. 65 ff.) in order to do intrasectoral aggregation. The addition of the distribution function is useful only in intersectoral aggregation; nothing is added to the traditional analysis of indexes of capital goods and prices. Hence, the national income statistician interested in the theoretical foundations of those indexes would not turn to the Houthakker approach.

The question of whether the Houthakker-Sato procedure is more or less restrictive than either those based on the weak separability property or the composite commodity condition is a difficult one to handle. The introduction of the distribution function complicates any comparison, since one has little basis for knowing if its specification and estimation is more or less restrictive than the requirements of the alternative aggregation procedures. That in one respect it allows for a (possibly) more limited range of possibilities (e.g., micro elasticities of substitution greater than unity are ruled out [Sato 1975, p. 61]) than weak separability, and so on, is clear. That it is an essentially short-run analysis puts it on the same footing as composite commodity aggregation but makes it less desirable than the Gorman theorem. That it is difficult to test empirically gives it the same grades on this account as the weak separability approach. That it allows for more general micro production processes than the Gorman theorem (except for the elasticity of substitution restriction above) is a significant point in its favor. That it requires fairly restrictive assumptions on stability of the distribution function detracts from the previous point.[27] And that intrasectoral aggregation requires some sort of restrictive weak separability or composite property as well as the somewhat restrictive stability conditions of the distribution function—that also is clear. Thus, much is clear yet, nevertheless, a comparison cannot yield an unambiguous answer on which procedure is preferable. It remains to say that the Houthakker-Sato approach must be subjected to further work to resolve some of the outstanding problems indicated above.

7.9 Commodity Aggregation Approach

Up to this point I have focused on the conditions for aggregation that arise out of the form of the production function (weak separability, homotheticity, etc.). The Hicks (1946) commodity aggregation approach that I now consider sidesteps those considerations of functional form. Hicks writes: "a collection of physical things can always be treated as if they were divisible into units of a single commodity so long as their relative prices can be assumed to remain unchanged, in the particular problem at hand" (1946, p. 33). Thus, let q_{ij} be the capital user costs; that is, $q_{ij} = p_{ij}(r_{ij} + \delta_{ij})$, $(i = 1,2)$ in its simplest form, where p_{ij} is the price of the ith capital good in the jth sector, r_{ij} is the net own interest rate, and δ_{ij} is the depreciation rate on the ith capital good. If the system is in equilibrium and depreciation is independent of the output rate, then the variables defining user costs are independent of the sector with which the capital is associated and the net own interest rates for all capitals are the same; therefore,

$$(32) \qquad q_i = p_i(r + \delta_i).$$

For our purposes, we can use (32) to illustrate the present aggregation procedure.

Now, define the value of capital in the jth sector as

$$(33) \qquad v_j \equiv \sum_i^2 q_i x_{ij} = q_1 \sum_i^2 \frac{q_i}{q_1} x_{ij}, \quad j = 1,2,$$

where the last equality would hold just as well were we to replace q_1 by q_2. Hicks proves that if $\left[\frac{q_1}{q_i}\right] d \left[\frac{q_i}{q_1}\right] = 0$, one can decompose v_j into two components, a "price" component, q_1, and a quantity component, $\sum_i^2 \frac{q_i}{q_1} x_{ij}$. With a slight modification, these components serve as price and quantity indexes of aggregate capital in the jth sector. Clearly, any number proportional to q_1, say $q^* = \alpha q_1$, would serve as the capital price index. Thus, the factor reversal test for price-quantity indexes is satisfied. Moreover, one can obtain "real" sectoral aggregate capital by deflating v_j by q^* and economywide "real" aggregate capital by deflating $\sum v_j$ by q^*. Finally, note that both the quantity and price indexes are homogeneous of degree one.

There are several reasons why prices of goods within a group may move in proportion to each other. Suppose certain prices are administered (fixed) over a relevant time period under conditions of monopoly (Fisher 1969, p. 572). Conversely, goods that are within a competitive industry or group would tend to move together in the long run. They may move together because of governmental price or incomes policy. Or, if the economy were in a balanced, steady-state growth or if it were stationary, prices would be constant and of course proportional to each other. Finally, if the labor shares in all firms are equal, then relative prices are constant (see below).

This is a very simple aggregating device, yet its exact form requires the stringent proportionality condition. However, approximations do not wreak havoc with the composite commodity conditions as they do with the functional form procedure. Clearly, commodity aggregation is unlikely to hold in general, but it may hold approximately for certain subgroupings (see Diewert 1974) and for some groups for certain periods and cycles but not for others. I shall elaborate upon this in a forthcoming study. Note that it has been used for theoretical purposes to justify partial and comparative static equilibrium analyses (see Arrow and Hahn 1971, pp. 7, 253).

It may also be the case that some prices are proportional to each other over certain time periods and not over others. For example, the trend and eight-year cycle could be the same for two prices, but they may differ over shorter-run cycles. Does this mean that the prices fail to satisfy the commodity aggregation theorem? Not at all, for a long-term grouping of the corresponding quantities is possible, whereas that grouping would make no sense in the short run.

Following this line of thought, one can consider the possibility that there is a systematic lead-lag relationships between the two prices but that, aside from that, they are proportional to each other. Does the lead or lag prevent the application of the commodity aggregation theorem? Again, the answer depends on the use to which the grouping is to be put.

There are many problems with this approach, the main one being that the q_j are not prices—rather, they are per unit capital rental values. Thus they are conglomerates of several factors that may change in various ways. Another problem is that published prices generally refer to total output, whereas a value-added price is the more appropriate concept. I shall elaborate on these and other matters relating to commodity aggregation in a forthcoming paper.

7.10 The Brown-Chang General Equilibrium Approach

The principal shortcoming of the preceding aggregation approaches is that they are done in a compartmentalized manner. That is, the restrictions on functional form required by the Gorman theorem are discussed in abstraction from their effect on prices; and the composite commodity theorem is derived without reference to economywide forces affecting prices of the items in the composite.

The Brown-Chang analysis remedies that deficiency in the literature by treating prices as endogenous within a simple general equilibrium model. Recall that the Gorman conditions focus on factor proportions while the composite commodity theorem emphasizes relative factor prices. But both factor proportions and factor prices are endogenous in a general equilibrium context, and hence any restriction on, or requirement of, one set of variables must affect the other.

Refer back to the little two-sector production function model (eq. 9), taking each to be homogeneous of degree one. Now add a zero profit condition for each sector:

$$p_j y_j / x_{0j} = 1 + q_1 x_{1j}/x_{0j} + q_2 x_{2j}/x_{0j}, \quad j = 1,2. \tag{34}$$

All prices are measured in terms of the nominal wage. For simplicity, assume for the moment that machines last forever (or one could also assume that depreciation rates are the same for both equipments) and thus the gross rental rates become: $q_i = p_i r$, where r is the only exogenous variable in the system. Parenthetically, we remark that one can "close" the system completely by specifying a relationship between net own rates of productivity and the rate of time preference (see Solow 1963) or by postulating that the fiscal monetary authorities control r to obtain a distributional objective (see Sraffa 1960). We require one other set of conditions—the marginal productivity equilibrium conditions must hold for each sector, that is,

$$\begin{aligned} f^j_0 &= \frac{\partial f^j}{\partial x_{0j}} = 1/p_0, \quad j = 1,2 \\ f^j_i &= \frac{\partial f^j}{\partial x_{ij}} = p_i r / p_j, \quad i,j = 1,2. \end{aligned} \tag{35}$$

These, together with (34) place us firmly within the world of perfect competition.

Brown and Chang now set out to find the conditions for the aggregation of the capital aggregates in a composite commodity sense; that is, under what conditions can one specify a $q^1{}_j$ and q such that

$$\begin{aligned} &i)\ x_j = \sum_{i=1}^{2} \frac{q_i}{q^1_j} x_{ij}, \quad j = 1,2 \\ &\text{and} \\ &ii)\ \mathrm{x} = \sum_{h=1}^{2} \frac{q^1_h}{q} x_h? \end{aligned} \tag{36}$$

Clearly, this requires

$$\frac{d(q_1/q_2)}{dr} = \frac{d(p_1/p_2)}{dr} = 0.$$

Thus, solving (34) for the p_j in terms of r, taking the ratio of the two expressions and then differentiating with respect to r, one finds

$$\text{(37)} \qquad \frac{d(p_1/p_2)}{dr} = \frac{1}{p_2^2 r \Delta}\left(P_1 \frac{y_1}{x_{01}} - p_2 \frac{y_2}{x_{02}}\right),$$

where $$\Delta \equiv \left(\frac{y_1}{x_{01}} - \frac{x_{11}}{x_{01}} r\right)\left(\frac{y_2}{x_{02}} - \frac{x_{22}}{x_{02}} r\right) - \frac{x_{12}}{x_{02}} \frac{x_{21}}{x_{01}} r^2.$$

The term in parentheses in (37) represents the difference in the wage shares in the two sectors.

Thus we arrive at the principal Brown-Chang result: commodity aggregation is assured, (36) holds, if the labor shares in the two sectors are equal. As an example, suppose the production functions are CES; aggregation of the two capital goods is possible—both intersectorally and intrasectorally—if

$$a_1 \left(\frac{y_1}{x_{01}}\right)^{\frac{1}{\sigma_1} - 1} = \frac{x_{01}}{p_1 y_1} = a_2 \left(\frac{y_2}{x_{02}}\right)^{\frac{1}{\sigma_2} - 1} = \frac{x_{02}}{p_2 y_2},$$

where a_i and σ_i $(i = 1,2)$ are constants. This condition requires that the production functions be restricted in a particular way, but the restriction appears to be weaker than the Gorman conditions, since factor ratios do not all have to be the same at all rates of interest.

In the general equilibrium model of production presented above, the equal-labor share condition guarantees the constancy of relative commodity prices (provided depreciation rates are equal), thereby permitting intrasectoral as well as complete aggregation of the capital items in the system. The condition can be applied to models with joint products and can be generalized to models with many primary factors; it can be applied to the capital goods as a group, to all sectors in the economy including the consumption good sector, and to a subgroup of capital goods sectors—the decomposable case—that does not require inputs from sectors outside the group. The decomposable case is particularly useful, since no matter how many groups of capital goods are in the economy, as long as there is a particular group whose production does not require capital inputs from sectors outside the group, the commodity aggregation condition applies, provided labor shares in that group are equal. In the decomposable case, however, the equal-labor share condition alone is not sufficient to guarantee aggregation of a proper subgroup of capital goods. If capital goods are divided into more than two groups, it is possible to derive more detailed conditions for aggregation. But these conditions are expected to be more stringent and therefore less likely to hold.

Intuitively, the equal-labor share condition amounts to equal capital/labor cost ratios in each industry—a variant of Marx's case of equal organic composition of capital. When the depreciation rates are the same for each capital good, this condition is equivalent to the equal factor intensity condition in value terms; that is, the *aggregate value* of capital/labor ratios are the same in every industry.

When there is an increase in the rate of interest, the increase in rental costs of capital in each industry depends upon (with identical rates of depreciation for each capital good) the *aggregate* value of capital employed in that industry. This explains why it is the capital intensity in value terms, not in physical terms, that is crucial in determining the effect on prices resulting from a change in the rate of interest. When every industry has the same capital intensity in value terms, an increase in the rate of interest will increase the cost of every commodity in the same proportion, thereby maintaining constant relative prices.

If the depreciation rates of different capital goods are different, the commodity price ratios and rental rates will not remain fixed as a result of a change in the rate of interest, even though the wage shares are equal in all sectors. However, Sato (1976*a*) has shown that if production is taken net of depreciation, the equal-labor share condition still applies.

The logic of the Brown-Chang and the Gorman aggregation conditions of a many-sector, many-capital, equal-depreciation rate model can now be examined. Suppose that there are n capital goods sectors and m consumption goods sectors. When the equal-labor share condition together with equal rates of depreciation for each capital good is satisfied, the relative commodity prices and the relative rental rates of capital are always constant. We can then apply the Hicksian aggregation theorem to perform intrasectoral aggregation over all capital inputs. When we appropriately choose the same units of measurement for the aggregated capital inputs in each sector, the system is reduced to one with n capital goods sectors and m consumption goods sectors, each with two factors of production, aggregated capital and labor. Now the equal-labor share condition in this two-factor model amounts to equal capital intensities in value terms as well as in physical terms. At this point the Gorman conditions are satisfied and intersectoral aggregation is possible. If all wage shares are equal in the capital good sectors, all these capital goods may be aggregated into a single capital good sector. A similar argument applies to the aggregation of the consumption goods sectors. The resulting system becomes the familiar two-sector, one-capital-good model.

Further aggregation of the two-sector model into a one-sector aggregative model can be achieved if the wage shares are equal in the two sectors. This implies the satisfaction of the Gorman conditions, and the production functions in the two sectors clearly differ only by an efficiency unit.

The commodity aggregation approach depends on the constancy of relative prices. In this case aggregation is completely independent of demand conditions. Clearly, this is not the only case allowing aggregation. For example, based on demand or utility functions, one can carry out aggregation of several outputs and production functions into an economywide production function (see Sato 1975).

In the general equilibrium model, it is not possible for the conditions of the Gorman theorem to be satisfied without simultaneously satisfying the condition for commodity aggregation (see Zarembka 1976). If the production functions happen to have a form such that the expansion paths for both firms are parallel straight lines through their respective origins, then (as noted above) the optimal factor ratios must be the same. Since all firms face the same factor prices, labor shares clearly must be the same in both sectors, and thus the commodity aggregation condition is satisfied in this model.

This leads to the unexpected result that the satisfaction of the Gorman conditions allows only intersectoral aggregation (see above), but if those conditions are satisfied in a general equilibrium type model, then the conditions for commodity aggregation are also satisfied, which means that intersectoral, intrasectoral, and full aggregation are permitted. In other words, if the very stringent Gorman conditions restricting the form of the production function are met (allowing only intersectoral aggregation), then the capital goods prices are proportional to each other (and to a capital price index), so that one need not stop with intersectoral aggregation but can proceed to aggregate all the capital items in each capital good sector and all the capital items among the many capital goods sectors.

The converse of this proposition clearly does not hold. For, even if all labor shares were to be equal, all optimal factor ratios do not have to be the same, so that if conditions for commodity aggregation hold, the Gorman conditions do not have to hold. It is possible for sectoral labor shares to be equal without all factor ratios being equal. Thus the conditions for commodity aggregation are somewhat weaker than that required by the theorem that focuses on the form of the production functions.

The equal-labor share condition is weaker than the Gorman condition in another respect. The latter requires that all capital items be used in each sector for intersectoral aggregation to hold, while the equal-labor share condition allows for the absence of capital items, since we have assumed only that labor is indispensable in each line of production. Certainly, in many sectors there are many corners, that is, numerous capital items not actually employed in production, and hence the Brown-Chang condition is weaker than the Gorman condition.

How do the Brown-Chang results relate to the other principal restriction on functional form procedure developed by Fisher? Fisher's discussion of the case of two-factor (fixed capital, movable labor) constant returns to scale contains a result similar to that found by Brown and Chang. There, a necessary and sufficient condition for the existence of an aggregate capital stock (obtained from all vintages) is that the average product of labor shall be the same for all vintages. Also, in his discussion of aggregation of fixed and movable capital items in the constant returns to scale case, Fisher (1968*b*, p. 422) shows that a necessary condition for total capital aggregation is that the average product of every kind of labor be the same in every firm whenever all movable factors (labor and movable capital) are optimally allocated. In the two cases cited above, he assumes that only one homogeneous output is produced, which must be the same for all firms. Therefore, with identical output prices in Fisher's models, equal average product of labor in every sector is equivalent to the equal-labor share condition. The Brown-Chang model assumes that all factors are movable, thus assuming away the problem of aggregating fixed factors.

Perhaps the most relevant way to compare the Brown-Chang model and Fisher's is to examine his aggregation of movable factors. Fisher discusses extensively full aggregation as well as subaggregation. In particular, he finds that what is required for subaggregation is that given the relative wages (relative prices) of the labor inputs (outputs) to be included in the aggregate, every firm employs those inputs (produces those outputs) in the same proportion. This is very close to parallel expansion paths. Note, however, that some fixed factors are assumed to exist and are left out of the aggregate in the case above. When there are no fixed factors, the condition that every firm employ all movable factors in the same proportion naturally implies the equal-labor share condition for the Brown-Chang model. The converse, of course, is not true, which implies that the equal-labor share condition may be more general than Fisher's result in this context. It is important to note that in Fisher's analyses, he uses conditions of technical efficiency (maximize the last output, given the amounts of the other outputs and the amounts of the inputs) while assuming the existence of some fixed factors that cannot be moved around over different uses so as to equalize marginal products. In the Brown-Chang analysis, it is assumed that all factors are movable so as to satisfy the equilibrium conditions. Clearly, the use of the equilibrium conditions makes aggregation easier. Finally, in the Brown-Chang model capital outputs are produced to be used as capital inputs. This is closely related to the capital aggregation problems associated with the recent reswitching debate.

What are the shortcomings of the Brown-Chang results? Certainly the general equilibrium model is not as general as one would like (such as that proposed by Arrow and Hahn 1971, chap. 5), but that is not a fundamental problem, since one would conjecture that many of the results would hold in a more general model. The real problem is the same as that encountered by the aggregation theory described above, which is based on duality theory, and that is the necessity for assuming competition in factor markets. The equal-labor share condition, though less restrictive than the Gorman conditions in certain respects, is still a stringent one. This is coupled with the fact that it cannot be applied in an economy where competition is suspected of being imperfect.

7.11 Structural versus Nominalistic Aggregation and a Paradox in Aggregation Analyses

We can now introduce an important distinction in aggregation theory. Consistent aggregates can be specified for essentially two reasons. The first is associated with the restrictions on functional form based in one way or another on the weak separability property of the underlying production functions. These give rise to proportional factor inputs. We know that, if this property is present, then whatever the behavior of the myriad aspects of the economy, consistent aggregation is preserved. That is, prices can change in a proportional or nonproportional manner because of supply shifts, say, and the aggregates would be unaffected. This says that if the physical-technical properties of production that manifest this property remain unchanged, the aggregates are preserved whether or not monopoly forces are present, whether relative supplies of factors change, whether disequilibrium effects are present, whether the economy is a steady-state growth path, whether an incomes policy is enforced, and so on. In short, knowing that the aggregates are conditional upon the properties of the production function is enormously economical. (Only nonneutral technical change would offset the aggregates.) We call this "structural" aggregation.

Going to commodity aggregation, we cannot infer this. Prices could be proportional to each other for a variety of reasons, and the resulting aggregates are subject to change owing to changes in any one of them. Thus, observing constancy of relative prices and basing the aggregation procedure on them would be questionable. If one is able to derive aggregates in this way at all, they are less likely to be stable than those derived from a knowledge of the properties of the production functions. Obviously, the resulting aggregates are real groupings in name only—hence we call this nominalistic.

It is one of the paradoxes of aggregation analysis that it may not be possible to derive the structural type of aggregates until and unless one can first obtain the nominalistic type. The reasoning is as follows. In order to test for weak separability (inter alia) underlying the structural aggregates, one must have data on the myriad physical capital items used in a given production unit. These data do not exist, nor are they likely to become available. Data on expenditures exist for many categories, but to obtain the physical data on the items within those categories, price indexes must be used to deflate them. But the very existence of price indexes that allow for consistent aggregation is the point in question; for one does not know that the production functions are homogeneous of degree one in the elementary inputs in each group index, since that is the object of the test. Mention could also be made of the current impossibility of estimating production functions with thousands of inputs even were the physical data available.

But one can test for nominalistic factor aggregation, and though that does not yield inferences directly with respect to the production function, one could use the resulting aggregates in estimating the aggregated production function over the sample for which the nominalistic aggregates hold. Aside from engineering approaches, that seems to be the only feasible way of going about it, thus giving rise to the paradox.

7.12 The Statistical Case (?) for Aggregative Analysis

The argument for specifying aggregative relationships directly rather than focusing on micro aspects of economic and technological behavior finds expression in the econometric literature. It is necessary to examine it to determine if and when it can be applied to the production and capital aggregation problems under discussion here. The intent is to make precise an aspect of the crude notion that macro relationships are preferable because of offsetting errors among the micro components. This is a purely statistical approach, and if it could be implemented it would afford a means of bypassing the difficulties of satisfying the aggregation conditions noted above.

In Theil's original work along this line, he found that the micro equations are the more appropriate ones to estimate under the assumptions of perfect (micro) model specification and nonstochastic regressors. This was developed further by Grunfeld and Griliches (1960), who indicate that there are circumstances in which an aggregate variable may be forecast with more precision than an aggregate of forecasts from the micro equations. Such a result arises from the possibility that micro equations are less well specified than the macro equation. The problem with the Grunfeld-Griliches analysis is that it is very difficult to be precise about when and the extent to which the micro equations are less well specified than the macro equation. In fact, one is reduced to articulating special cases and examples of the alleged specification bias rather than a general analysis. But, as we shall see, that is not the main objection to the whole procedure.

Rejecting the Grunfeld-Griliches approach, Orcutt, Watts, and Edwards (see Edwards and Orcutt 1969 for references) focus on the difficulty of obtaining suitable micro data. In general they find that it is better to forecast on the basis of an aggregate of micro forecasts rather than doing a macro forecast. In any event, the focus of the literature has switched to the measurement errors attached to micro and macro data. If micro data is subject to more measurement error than aggregate data, there is the trade-off of the loss from the specification bias resulting from aggregation and the potential gain from the reduction in the inaccuracy of the measured aggregate data.

Aigner and Goldfeld (1974) consider the problems of estimation and prediction when the data on independent variables contain less measurement error than the micro data. Greene (1975), correcting an error in their model, does the same. The last reference I make to this literature is Welsch and Kuh (1976), who employ a general random coefficient model to determine how the variances of the coefficients behave as the number of micro units increase.

To give an idea of one line of development, let us follow Greene's analysis of the very simple offsetting errors case. Suppose there are only two micro equations,

$$\begin{aligned} Y_1 &= \beta_1 X_1 + u_1 \\ Y_2 &= \beta_2 X_2 + u_2, \end{aligned} \tag{38}$$

where the measured values of the micro variables are

$$\begin{aligned} y_1 &= Y_1 + w \\ y_2 &= Y_2 - w \end{aligned} \tag{39}$$

and

$$\begin{aligned} x_1 &= X_1 + v \\ x_2 &= X_2 - v. \end{aligned} \tag{40}$$

The macro variables, $(Y_1 + Y_2) = (y_1 + y_2)$ and $(X_1 + X_2) = (x_1 + x_2)$, are clearly assumed to be measured without error.[28]
After combination,

$$(41)\qquad \begin{aligned} &i)\ y_1 = \beta_1 x_1 + \epsilon_1 \\ &ii)\ y_2 = \beta_2 x_2 + \epsilon_2, \end{aligned}$$

where $\epsilon_1 = u_1 + \beta_1 v - w$ and $\epsilon_2 = u_2 - \beta_2 v + w$. The macro equation is simply the addition of these, which is,

$$(42)\qquad (y_1 + y_2) = \beta(x_1 + x_2) + \epsilon,$$

where $\epsilon = u_1 + u_2 + (\beta_1 - \beta)\, x_1 + (\beta_2 - \beta)\, x_2 + (\beta_1 - \beta_2)\, v$. All variables are taken to be independent and normally distributed with zero means; that is, $(X_1,X_2) \sim N(0,\Sigma_x)$, $(u_1,u_2) \sim N(0,\Sigma_u)$, $w \sim N(0,\sigma^2_w)$ and $v \sim N(0,\sigma^2_v)$, where

$$\Sigma_x = \begin{bmatrix} \sigma^2_{x_1} & \sigma_{x_1x_2} \\ \sigma_{x_1x_2} & \sigma^2_{x_2} \end{bmatrix}; \qquad \Sigma_u = \begin{bmatrix} \sigma^2_{u_1} & \sigma_{u_1u_2} \\ \sigma_{u_1u_2} & \sigma^2_{u_2} \end{bmatrix}.$$

The error terms in (41) involve the βs, and thus the estimation problem is akin to a classic errors-in-variables problem. The limiting values to which the least-squares estimates of β_i tend in probability can be shown to be

$$\text{plim}\ \hat{\beta}_i = \beta_i \bigg/ 1 + \frac{\sigma^2_v}{\sigma^2_{x_i}}\ .$$

This gives the familiar result that the micro parameter $\hat{\beta}_i$ is in fact an underestimate of β_i.

Now, going to the macro equation, (42), one finds that

$$\text{plim}\ (\hat{\beta}) = \gamma\beta_1 + (1 - \gamma)\, \beta_2,$$

where

$$\gamma = (\sigma^2_{x_1} + \sigma_{x_1x_2})/(\sigma^2_{x_1} + \sigma^2_{x_2} + 2\sigma_{x_1x_2}),$$

which is also inconsistent. Hence, in deciding whether to use the micro or the macro equations when measurement errors offset each other in the micro variables, the choice devolves upon two sets of inconsistent estimators. A preliminary conclusion can be reached here without further analysis; and that suggests that if β_1 and β_2 are close together and if the measurement error is large, one would be advised to estimate the macro equation (42) and use $\hat{\beta}$ as an estimate of each β_i rather than estimate each β_i separately from (41). This result is made more precise by specifying the mean square errors of the estimators and comparing them. However, this large sample result is not an interesting one, since an instrumental variable estimator, using $(x_1 + x_2) = (X_1 + X_2)$, can be shown to dominate both micro and macro least-square estimators (Greene 1975).

Even in the small sample case, there is an opportunity for the macro estimator to outperform the micro in spite of the presence of "aggregation bias." Here aggregation bias is represented by $\beta_1 \neq \beta_2$. It is found that, though there are exceptions, in most cases in which there is aggregation bias, the micro estimator is superior. However, when observation errors are introduced, there is an inevitable trade-off between aggregation bias and measurement errors of the micro data. It is found that as the error variance increases relative to the "true" variables, the advantage of the micro estimator declines, which is not an unexpected result.

There are many different results in this literature. Some studies support prediction in certain circumstances from disaggregated data (e.g., Edwards and Orcutt 1969). Using different models, Grunfeld and Griliches (1960) and perhaps Aigner and Goldfeld (1974) and Greene (1975) find superiority in certain circumstances in macro analysis. In the Welsch and Kuh (1976) analysis, it is difficult to say which is superior.[29] A summing-up results in the characterization of the glass as half-full or half-empty and hence is not very informative.

There are two immediate problems with this analysis. The first is that the model itself is extremely simple: linear specifications, variances of the exogenous variables identical, and no measurement error on the macro variables. The last assumption is quite restrictive, for there is no reason to believe that, in general, the macro variables are free of measurement error if the micro data from which they are computed are not. If and when this type of analysis proceeds to examine macro measurement error, there will be three elements to the trade-off that will have to be considered: micro and macro measurement error and aggregation bias. Whether the resulting analysis will be more than impressionistic remains to be seen.

The second problem with this statistical approach is a fundamental one within the context of aggregation of production and capital. It devolves upon the notion of aggregation bias, which here simply means $\beta_1 \neq \beta_2$. But, in the reswitching literature, not only may the form of the macro relationship differ from the micro equations, but the macro equation probably will contain different variables. The difference between the macro equation and the micro equation is not only that the former is in some sense an aggregate of the micro relationships, but that the two types of specifications may differ. Moreover, it is very, very hard (if not impossible) to know even the approximate specification of the macro relationship without knowing the properties of the micro equations. Therefore one is forced to treat the micro relationships directly; for, in spite of the alleged measurement error, there is simply no way of proceeding. The statistical approach provides us with interesting and suggestive, though tentative, results; but it is at present irrelevant to the problem of aggregation of production and capital.

7.13 Conclusions

The review of aggregation theory leads to two general conclusions: the first is that if one ignores the potential aggregation bias, qualitatively incorrect predictions result, which indicates the importance of the problem for the areas of production, income distribution, productivity, and pricing; the second is that at present there is only one basis, flawed though it is, for testing capital aggregates, and that is the commodity aggregation approach. None of the procedures that focus on functional form is feasible because, aside from the stringency of the conditions they require, they need an inordinate amount of micro physical data that simply are not available. Thus, the only feasible procedure is commodity aggregation, for the requisite data appear to be available and the conditions do not rule out imperfect competition. The principal shortcoming is that it allows for the specification only of nominalistic aggregates. But if they exist, then at least one could test for more enduring aggregates based on restrictions on functional form. That seems to be the appropriate course of action, given our review of the theoretical and statistical bases for aggregation.

Notes

1. This review does not pretend to be a definitive statement of the problem. In fact, no such thing exists, though reviews from one or another point of view are available; see Harcourt (1972), Blaugh (1974), Samuelson (1976), and Burmeister (1976).

2. Bruno, Burmeister, and Sheshinski (1966) show that there is no essential difference between the circulating-capital and the fixed-capital models as far as the important capital-theoretic issues are concerned.

3. Before doing any analysis with this model, it is necessary to ensure that the technique is feasible, which means that P_1/P_0 must be positive. See Hicks (1965, pp. 97–98). If $r < 1/a_{11}$, one can show by differentiating (6) that $P_1(\alpha)/P_0$ will be an increasing or decreasing function of the rate of profit, depending on whether the capital good sector is more or less capital intensive than the consumption good sector. The curves in figures 7.1*a–d* are drawn with the assumption that the consumption sector is more capital intensive than the capital sector. Note that, if $m = 1$, relative prices are independent of the profit rate. But this is unimportant for Champernowne's chain index.

4. See Brown (1969). W_0/P_0 has the dimension:

$$\dim \frac{\$\text{ labor}}{\$\text{ corn}} = \frac{\text{Labor}}{\text{Corn}},$$

which is invariant to changes in technique. However,

$$\dim \left[\frac{P_1(\alpha)}{P_o}\right] = \frac{\$/\text{capital type } \alpha}{\$/\text{corn}} = \left[\frac{\text{capital } \alpha}{\text{corn}}\right]^{-1},$$

which does not have the same dimension as $P_1(\beta)/P_o$ and $P_1(\gamma)/P_o$.

5. The condition that the factor-price relation should be linear in Samuelson's surrogate production is stronger than that required for the construction of Champernowne's chain index. To find a chain index, only one intersection between any pair of techniques is required. Straight-line factor-price relations are not necessary in Champernowne's construction. However, if these ratios are identical for all techniques, there is only one intersection between any pair of techniques (Hicks 1965, p. 154). It is in this sense that the condition for constructing a chain index is somewhat weaker than that for the surrogate production function. Yet, if the simple Marshallian elasticity at each point on the frontier is to be used to measure the distribution of income, it is necessary and sufficient that the factor-price curves should be linear.

6. Initial and terminal capital stocks and all other consumptions are understood to be held constant in (a) and (b). If there are many capital goods with no joint production, the wage rate in terms of every good's price forms a factor-price or wage-profit frontier in (c); this is a generalization of the model used above to illustrate the chain index and the surrogate production function. If the rate of growth, g, is positive, then the monotone relations in (e) and (f) are taken to hold only for $r > g$ and for K/L less than the golden rule capital-labor ratio associated with $r = g$. If there are many capital goods (the general model), then in (g) and (g'), $\Sigma P_i K_i / P_i Y_i$ is to be expressed in terms of some numeraire.

7. It may be useful to indicate what reswitching does not imply. It does not imply that marginal analysis is silly; one can use smoothly differentiable production functions or specify a production possibility set that is closed and convex (this is more general in one sense but it rules out increasing returns), and there is a considerable intersection of implications that results from the two specifications; the choice should be empirically determined. It does not imply that there are inherent contradictions in capitalistic production; one has to refer to an entirely different literature to try to show that. Finally, it does not imply that general equilibrium theory is silly because in that theory one can specify as many heterogeneous capital items as one wishes, treating each as a separate good with its own market, etc.; no aggregates need be involved.

8. For any two distinct points in input space, x_{ij} and x'_{ij}, strict quasi-concavity is defined by $f^j[(1-\alpha)x_{0j} + \alpha x'_{0j}, (1-\alpha)x_{1j} + \alpha x'_{1j}, (1-\alpha)x_{2j} + \alpha x'_{2j}] > \min\,[f^j(x_{0j}, x_{1j}, x_{2j}), f^j(x'_{0j}, x'_{1j}, x'_{2j})]$, where $0 < \alpha < 1$.

9. There are essentially two concepts of separability that are important here. The first is weak separability, already defined. The other is strong separability: consider a function F of n variables; it is called strongly separable with respect to a partition $\{N_1, \ldots, N_S\}$ if the marginal rate of substitution between two inputs i and j from different subsets of inputs N_S and N_t, respectively, does not depend upon inputs outside of N_S and N_t; that is, let $y = f[\phi^1(x^{(1)}) + \phi^2(x^{(2)}) + \ldots + \phi^S(x^{(s)})]$, where ϕ^i $(i = 1, \ldots, S)$ are functions of the subsectors $x^{(i)}$ and f is a monotone-increasing function of the ϕ_i; then

$$\frac{\partial \left[\dfrac{\partial \phi^i}{\partial x_t^{(i)}} \bigg/ \dfrac{\partial \phi^j}{\partial x_s^{(i)}}\right]}{\partial x_k} = 0, \quad i \epsilon N_S,\ j \epsilon N_T;\ k N_S \cup N_T (S \neq T).$$

If the function is weakly separaоle, it can be written as $y = f[\phi^1(x^{(1)}),$ $\phi^2(x^{(2)}), \ldots, \phi^S(x^{(s)})]$. For proofs of these propositions, see Goldman and Uzawa (1964).

10. Weak separability is not both necessary and sufficient to accomplish this, since indexes can be formed if the weights on the inputs within a group index behave in a certain way (see below). The behavior of the weights can be independent of the form of the production function.

11. A function of $x_1, x_2, \ldots, x_n$ factors of production is homothetic if it can be written as $\phi^j(\sigma^j(x_1, x_2, \ldots, x_n))$, where σ^j are homogeneous of degree one and ϕ^j are continuous monotonically increasing functions of σ^j. Actually, less restrictive ϕ^j functions (namely, upper semicontinuous functions; see Shephard 1970, pp. 92 ff.) produce similar results.

12. These conditions are necessary and sufficient if one focuses on the form of the production function alone (Green 1964, p. 28). They are merely necessary if one allows for the commodity aggregation condition to hold (see below) or if something is going on "behind the scene" (see Brown and Chang 1976).

13. This can be generalized in several directions when n-factors ($n > 3$) are considered. First, the number of partitions can obviously be extended; second, the number of items in each group can be variable provided no item is allowed to be in more than one group; third, one can allow for several groups and many individual factors (in the text, we have one grouping of capital items and another "group" consisting of the labor input); finally, the F^j and g_j need not be continuous; they can be finite, nondecreasing, nonnegative upper semicontinuous functions (Shephard 1970, pp. 20 ff.). The last extension allows for discontinuities, provided the functions are continuous only from the right.

14. Note that weak separability does not require the g^j function to be homogeneous.

15. They are: (*i*) if all prices and quantities are fixed between time point t_1 and t_0, then $g^j_{t_1} = \Gamma^{1i}_{r_1 t_0} = 1$; (*ii*) if all prices (quantities) at t_1 are proportional to those at t_0, then $g^j_{t_1 t_0} = \Gamma^{1j}_{t_1 t_0} = \alpha$, where α is the factor of proportionality; and (*iii*) $g^j_{t_1 t_0} \Gamma^{1j}_{t_1 t_0} = \sum_{i=1}^{2} p^1_{ij} x^1_{ij} \Big/ \sum_{i=1}^{2} p^0_{ij} x^0_{ij}$.

16. This assumption is tenuous at best in light of the commonly held view that capital markets are notoriously imperfect.

17. Fisher (1969, p. 560) finds that capital aggregation is possible under somewhat less restrictive conditions than under capital augmenting technical differences. This involves the case in which each firm's production function becomes one of constant returns after a transformation of the capital inputs; that is, $F^j(x_{0j}, x_{1j}) = F^j[H^j(x_{1j}), x_{0j}]$, where the f^j are homogeneous of degree one and the H^j are monotonic. Despite the fact that this is more general than capital augmentation, it is itself very restrictive.

18. It is known that if all firm's production functions are quasi-concave and homogeneous of degree $\lambda(0 < \lambda \leq 1)$, then G is convex.

19. For example, the Hicksian condition, $m = 1$, which rules out reswitching (1965), and Zarembka's conditions for aggregation (1975) are simply the Gorman conditions in two of their guises.

20. To be precise, one must take the differential in terms of the exogenous variable. Here we can treat the interest rate as exogenous (see discussion of the Brown-Chang conditions below).

21. Actually, the Bureau of Labor Statistics wholesale price index follows the Lespeyres formula, but in a modified form. When new weights are introduced, the chain-link device is employed, at the linkage points, implies a Divisia index.

22. To see what is involved, note that the WPI weights represent the total net selling value of commodities produced, processed, or imported into the United States and flowing into primary markets; see BLS (1971, pp. 103–4). The weights in the WPI correspond to the $w_{.h}$ and hence are economywide weights.

23. We assume that the system satisfied the Hawkins-Simon conditions (see Brown and Chang 1976).

24. The equal or zero depreciation assumption is one. Another is the assumption that production functions are homogeneous of degree one.

25. See for example, Solow (1967, pp. 46–48).

26. The ex ante production function describes the factor substitution possibilities before capital is installed; that is, it is the putty part of the putty-clay appelation representing the whole range of blueprints for the production process available to the firm in its planning stage.

27. Using a different approach, Sonnenschein (1973) has shown that (*a*) for a given aggregate expenditure and for prices and aggregate demands—all of which satisfy the aggregate budget constraint—and (*b*) any set of rates of change of aggregate demands with respect to prices and total expenditure (these must satisfy the homogeneity constraint), then there is a finite collection of utility maximizing consumers with equal total expenditures. An aggregate demand system is the result, but it is peculiar to the point (prices, aggregate demands, etc.) that is initially taken as given. Thus the implied aggregate demand systems (the functions themselves) are conditional upon prices, etc. This difficulty is similar to that affecting the Houthakker-Sato distribution function.

28. The case where the macro variables contain some measurement error has yet to be worked out in an acceptable manner.

29. In analyzing the affect of aggregation on the reduction in the variance of parameter estimates, Welsch and Kuh use a model that is similar to Green's except for one important characteristic. The former model allows for random coefficients; that is, $E(\beta_i) = \beta$. This is seemingly more general than Green's model, but in fact Welsch and Kuh have assumed away one of the most interesting aspects of the analysis; namely, aggregation bias. For they do not allow the β_i to vary in a nonstochastic manner among the micro units. However, they advise that the relative efficiency of aggregation could be severely reduced by differences in micro behavior (1976, p. 362).

References

Aigner, D. J., and Goldfeld, S. M. 1974. Estimation and prediction from aggregate data when aggregates are measured more accurately than their components. *Econometrica* 47:113–34.

Arrow, K. J., and Hahn, F. H. 1971. *General competitive analysis.* San Francisco: Holden-Day.

Blaugh, M. 1974. *The Cambridge revolutions: Success or failure: A critical analysis of Cambridge theories of value and distribution.* London: Institute of Economic Affairs.

Brown, M. 1969. Substitution-composition effects, capital intensity, uniqueness and growth. *Economic Journal* 79:334–47.

——— 1973. Toward an econometric accommodation of the capital perversity phenomenon. *Econometrica* 41:937–54.

Brown, M., and Chang, W. W. 1976. Capital aggregation in a general equilibrium model of production. *Econometrica* 44 (November): 1179–1200.

Brown, M.; Sato K.; and Zarembka, P. 1976. *Essays in modern capital theory*. Amsterdam: North-Holland.

Bruno, M. 1969. Fundamental duality relations in the pure theory of capital and growth. *Review of Economic Studies* 36:39–53.

Bruno, M.; Burmeister, E.; and Sheshinski, E. 1966. The nature and implications of the reswitching of techniques. *Quarterly Journal of Economics* 80:526–53.

Burmeister, E. 1977. On the social significance of the reswitching controversy. *Revue d'Economie Politique*. In press.

Burmeister, E., and Kuga, K. 1970. The factor-price frontier in a neoclassical multi-sector model. *International Economic Review* 11:162–74.

Cass, D. 1976. The Hamiltonian representation of static competitive or efficient allocation. In *Essays in modern capital theory,* ed. M. Brown, K. Sato, and P. Zarembka. Amsterdam: North-Holland.

Champernowne, D. G. 1953. The production function and the theory of capital: A comment. *Review of Economic Studies* 21:112–35.

Diewert, W. E. 1974. A note on aggregation and elasticities of substitution. *Canadian Journal of Economics* 12:12–20.

Edwards, B., and Orcutt, G. H. 1969. Should aggregation prior to estimation be the rule. *Review of Economics and Statistics* 51:409–20.

Fisher, F. 1965. Embodied technical change and the existence of an aggregate capital stock. *Review of Economic Studies* 32:263–88.

———. 1968*a*. Embodied technology and the existence of labour and output aggregates. *Review of Economic Studies* 35:391–412.

———. 1968*b*. Embodied technology and the aggregation of fixed and movable capital goods. *Review of Economic Studies* 35:417–28.

———. The existence of aggregate production functions. *Econometrica* 37:553–77.

Fisher, F. M.; Solow, R. M.; and Kearl, J. R. 1974. Aggregate production functions: Some CES experiments. Working paper, Department of Economics, M.I.T. Cambridge, Mass.

Goldman, S. M., and Uzawa, H. 1964. A note on separability in demand analysis. *Econometrica* 32:387–98.

Gorman, W. M. 1953. Community preference fields. *Econometrica* 21:63–80.

———. 1959. Separable utility and aggregation. *Econometrica* 27:469–81.

Green, H. A. J. 1964. *Aggregation in economic analysis: An introductory survey*. Princeton: Princeton University Press.

Greene, W. H. 1975. Estimation and prediction from aggregate data when aggregates are measured more accurately than their components: Some further results. Discussion Paper 7513, Social Systems Research Institute, University of Wisconsin, Madison, Wisc.

Grunfeld, Y., and Griliches, Z. 1960. Is aggregation necessarily bad. *Review of Economics and Statistics* 42:1–13.

Harcourt, G. C. 1972. *Some Cambridge controversies in the theory of capital*. Cambridge: Cambridge University Press.

Hatta, T. 1974. The paradox in capital theory and complementarity of inputs. Department of Economics, Ohio State University.

Hicks, J. 1946. *Value and capital*. London: Oxford University Press.

———. 1965. *Capital and growth*. New York: Oxford University Press.

Houthakker, H. S. 1955–56. The Pareto distribution and the Cobb-Douglas production function in activity analysis. *Review of Economic Studies* 23:27–31.

Katzner, D. W. 1970. *Static demand theory*. New York: Macmillan.

Leontief, W. W. 1947. Introduction to a theory of the internal structure of functional relationships. *Econometrica* 15:361–73.

Levhari, D. 1968. A note on Houthakker's aggregate production function in a multifirm industry. *Econometrica* 36:151–54.

Nuti, D. M. 1976. On the rates of return on investment. In *Essays in modern capital theory*, ed. M. Brown, K. Sato, and P. Zarembka. Amsterdam: North-Holland.

Oguchi, N. 1977. *Dynamic growth models and capital perversities*. Ph.D. Diss., Department of Economics, SUNYAB.

Robinson, J. 1956. *The accumulation of capital*. London: Macmillan.

Samuelson, P. A. 1962. Parable and realism in capital theory: The surrogate production function. *Review of Economic Studies* 29:193–206.

———. 1976. Interest rate determinations and oversimplifying parables: A summing up. In *Essays in modern capital theory*, ed. M. Brown, K. Sato, and P. Zarembka. Amsterdam: North-Holland.

Sato, K. 1974. The neoclassical postulate and the technology frontier in capital theory. *Quarterly Journal of Economics* 88:353–84.

———. 1975. *Production functions and aggregation*. Amsterdam: North-Holland.

———. 1976*a*. A note on capital and output aggregation in a general equilibrium model of production. Department of Economics, SUNYAB.

———. 1976*b*. Perversity in a multisector neoclassical production system. Discussion Paper number 380. Department of Economics, SUNYAB.

Shephard, R. 1953. *Cost and production functions*. Princeton: Princeton University Press.

———. 1970. *Theory of cost and production functions*. Princeton: Princeton University Press.

Solow, R. M. 1955. The production function and the theory of capital. *Review of Economic Studies* 23:101–8.

———. 1963. *Capital theory and the rate of return*. Amsterdam: North-Holland.

———. 1967. Some recent developments in the theory of production. In *The theory and empirical analysis of production*, ed. M. Brown. Studies in Income and Wealth, no. 31. New York: National Bureau of Economic Research.

Sonnenschein, H. 1973. The utility hypothesis and market demand theory. *Western Economic Journal* 11:404–10.

Sraffa, P. 1960. *Production of commodities by means of commodities: A prelude to a critique of economic theory*, Cambridge: Cambridge University Press.

Strotz, R. 1959. The utility tree: A correction and further appraisal. *Econometrica* 27:482–88.

U.S. Department of Labor. Bureau of Labor Statistics. 1971. *BLS handbook of methods*. Bulletin 1711. Washington, D.C.

Usher, D. 1973. The measurement of economic growth. Discussion Paper no. 131, Institute for Economic Research, Queen's University, Kingston, Ontario.

Welsch, R. E., and Kuh, E. 1976. The variances of regression coefficient estimates using aggregate data. *Econometrica* 44:353–64.

Zarembka, P. 1975. Capital heterogeneity, aggregation, and the two-sector model. *Quarterly Journal of Economics* 79:103–14.

———. 1976. Characterizations of technology in capital theory. In *Essays in modern capital theory*, ed. M. Brown, K. Sato, and P. Zarembka. Amsterdam: North-Holland.

Zellner, A.; Kmenta, J.; and Dreze, J. 1966. Specification and estimation of Cobb-Douglas production function models. *Econometrica* 34:784–95.

Comment Edwin Burmeister

Murray Brown's paper provides a comprehensive survey of the theoretical and practical problems associated with capital aggregation. It should be required reading for every econometrician before he is allowed access to his computer! The paper is long, and I shall have space only to briefly summarize some of the results, along with a few comments of my own.

Brown begins with a review of the neoclassical parable and the so-called reswitching controversy that reached its peak with the November 1966 *Quarterly Journal of Economics* Symposium, "Paradoxes in Capital Theory" (1966). The primary issue can be easily explained. Suppose there exist two alternative Leontief-Sraffa production techniques, a and b, both using a homogeneous labor input and n types of heterogeneous capital inputs. Suppose also that there is a single consumption good. In a steady-state equilibrium the technique employed will maximize the real wage W/P_c or minimize P_c/W, the price of final output in terms of the single primary factor, labor. Thus in figure C7.1*a* and *b*, technique a will be used for $0 \leqq r < r_1$ and $r_2 < r \leqq r^*$, while technique b will be used for $r_1 < r < r_2$. Both techniques are viable and may coexist at the *switch points* r_1 and r_2.

The crucial observation is that all physical quantities, for example, the stocks of capital goods and the output of the final consumption good, depend only on the technique employed. Thus, suppose we define any indexes of "capital" for techniques a and b, say K^a and K^b, that depend only upon the technique employed. Clearly the existence of reswitching makes it obvious that the techniques cannot be ordered in terms of such indexes and the steady-state profit rate because, when there is reswitching, at least one technique in employed for two disjoint intervals of the profit rate.

That physical quantities depend only on the technique employed is illustrated in figure C7.1*b* where equilibrium consumption is plotted for alternative steady-state profit rates. Note that $C = C^a$ when technique a is employed ($0 \leqq r < r_1$ and $r_2 < r \leqq r^*$), while $C = C^b$ when technique b is employed ($r_1 < r < r_2$). If the technology consists of smooth neoclassical production functions, figure C7.1*b* is replaced by figure C7.2. Although reswitching is precluded in these circumstances,[1] *paradoxical consumption behavior* may still exist; that is, steady-state consumption

Edwin Burmeister is Commonwealth Professor of Economics at the University of Virginia.

Research support from the Center for Advanced Studies at the University of Virginia and the National Science Foundation is acknowledged with thanks. I am also grateful to John Whitaker for helpful comments.

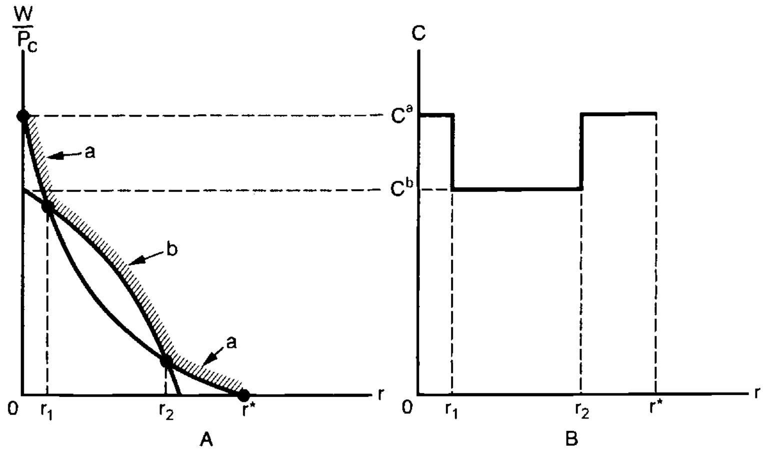

Fig. C7.1

may rise with the profit rate, as illustrated in figure C7.2. Thus there cannot exist a well-behaved neoclassical production function (across steady states),

$$C = F(K,L),$$

where C is consumption, L is the fixed labor supply, and K is some index of capital that always falls with an increase in the steady-state profit rate. Brown refers to my own result that such an aggregate production function can be defined if, and only if,

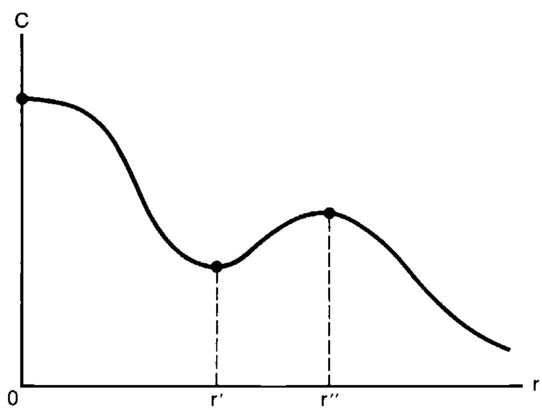

Fig. C7.2

$$\sum_{i=1}^{n} p_i \frac{dk_i}{dr} < 0 \text{ for all feasible } r,$$

where p_i is the price of the ith capital good in terms of any numeraire and k_i is the capital-labor ratio for capital of type i, $i = 1, \ldots, n$.[2]

I agree fully with Brown's stated conclusion that "the neoclassical parable and its implications are generally untenable" (p. 15). Freak cases such as Samuelson's surrogate production function example are of little comfort.

In section 7.4 Brown reviews capital aggregation theorems that work because the form of the production function is restricted. The original Leontief (1947) theorem concerns

$$Q = F(L, K_1, K_2);$$

if

$$\frac{\partial}{\partial L}\left[\frac{\partial F}{\partial K_1} \bigg/ \frac{\partial F}{\partial K_2}\right] = 0,$$

then there exists an aggregate production function

$$Q = F(L, K),$$

where "aggregate capital" is given by some function

$$K = G(K_1, K_2).$$

Brown then discusses additional problems that arise when there are more than two groups of inputs, and he summarizes the results due to Strotz (1959) and Gorman (1959), as well as the application of the duality between cost and production functions stemming from Shephard's work (1953). He concludes with two objections to this approach.

1. Price and quantity indexes are constructed *assuming* perfect competition in factor markets, and in many instances this assumption is obviously false.

2. The basic micro data required for the construction of capital indexes, even when they conceptually exist, generally are unobserved. I should like to add a third problem.

3. Suppose K_1 and K_2 are two physically different types of capital goods. Using the production function for industry 1, F^1, we construct an index for "capital" in that industry, say $G^1(K_{11}, K_{21})$. But, using a different production function for industry 2, having the same physical inputs, in general we have a different index $G^2(K_{12}, K_{22})$. Thus even if we consider *all* points for which the quantities of the physical inputs are the same in both industries and $K_{11} = K_{12}$, $K_{21} = K_{22}$, in general $G^1 \neq \lambda G^2$ for any scalar $\lambda > 0$. This means that an aggregate production function for the whole economy need not exist, even when sectoral capital aggregation is possible.

The problem of intersectoral aggregation—the basis for my last objection—has been studied in a series of papers by Frank Fisher (1965, 1968*a*, *b*, 1969). We now have n industry production functions

$$Q_j = F^j(L_j, K_j), \quad j = 1, \ldots, n,$$

and ask when we can find an aggregate production function

$$Q = \sum_{j=1}^{n} Q_j = F(L,K),$$

where

$$L = L(L_1, \ldots, L_n),$$
$$K = K(K_1, \ldots, K_n).$$

Essentially the Fisher sufficient condition is that, when labor is optimally allocated, every production function must differ only by the degree of capital augmentation or, as Brown states, "For example, sound amplification equipment in a classroom is considered to be three times the number of desks in the same classroom." I agree with the negative feelings expressed by both Fisher and Brown for progress along this line.

The Gorman (1953) aggregation conditions require that all firms have homothetic production functions with parallel expansion paths through the origin, in which case we may express

$$Q_j = F^j(L_j, K_{1j}, \ldots, K_{nj}), \quad j = 1, \ldots, n,$$

in the aggregate form

$$Q = \sum_{j=1}^{n} h^j(Q_j) = F(L, K_1, \ldots, K_n),$$

where

$$L = \sum_{j=1}^{n} L_j$$

and

$$K_i = \sum_{j=1}^{n} K_{ij}, \quad i = 1, \ldots, n.$$

If in addition the production functions $F^j(\bullet)$ exhibit constant returns to scale, then by renumbering the isoquants the production functions may be made identical; that is, outputs are identical except for the units in which they are measured. Brown proceeds to show how such unrealistic conditions are often assumed implicitly when one uses a price index for capital goods as a deflator to measure "real capital" in an industry.

The Houthakker (1955–56) approach discussed in section 7.8, although ingenious, seems to me of little relevance for the primary issue at hand. Its usefulness is limited to the intersectoral aggregation of production functions with two factor inputs, and even then there are formidable estimation difficulties stated by Brown.

All the aggregation procedures discussed so far rely on functional form restrictions. Alternatively, the Hicks (1946) composite commodity theorem allows aggregation of heterogeneous commodities if their relative prices remain constant for the problem under consideration. The relevant question is then which hypothetical alternatives are to be investigated.

The Hicks theorem is the basis for the Brown and Chang (1976) general equilibrium aggregation results. This model requires the following assumptions:

1. There is no joint production.
2. The rate of profit r is exogenous; for example, r may be determined exogenously by the rate of time preference or by fiscal and monetary policy (section 7.10).
3. There is no technological change.
4. Steady-state equilibrium always prevails so that we may express capital net rentals rates as

   $$q_i = p_i r$$

   rather than the more general form

   $$q_i = p_i r - \dot{p}_i$$

   that allows for capital gains and losses.
5. There is perfect competition in factor markets (section 7.10).

I fear that any one of these five reasons is sufficient to reject the model as empirically unrealistic; but suppose we accept it. For such a model Marxians know that relative prices are constant if, and only if, there is "equal organic composition of capital" and the "cost of labor/value of capital" ratio is the same function of r for every industry:

$$\frac{WL_j}{\sum_{i=1}^{n} P_i K_{ij}} = \psi(r), \quad j = 1, \ldots, n.$$

This condition, of course, leads to a labor theory of value in which relative prices reflect the ratio of *total* embodied labor; that is,

$$\frac{p_i(r)}{p_j(r)} = \frac{l_i}{l_j} = constant \quad \text{for all } r,$$

where the vector of total embodied labor is given by

$$l = (l_1, \ldots, l_n) = \frac{P}{W}(r = 0) = p(0)$$
$$= [p_1(0), \ldots, p_n(0)].$$

It is also a theorem that such a labor theory of value is valid *if, and only if,* prices are a markup on unit labor costs:

$$\frac{P_i}{W}(r) \equiv p_i(r) = \alpha(r) a_{Li}(r),$$

where

$$a_{Li}(r) = \frac{L_i}{Q_i}, \quad i = 1, \ldots, n,$$

and where $\alpha(r)$ is the same markup function for all industries. As Brown and Chang state the result,

$$\text{labor's relative share} = \frac{WL_i}{P_i Q_i} = \frac{a_{Li}}{P_i}$$
$$= \text{the same function of } r \text{ alone}$$
$$= \frac{1}{\alpha(r)}, \quad i = 1, \ldots, n.$$

I am afraid few econometricians would be willing to assume such stringent conditions. The Marxian case of equal organic composition is precisely that freak situation in which capital theoretic problems due to heterogeneity do not arise![3]

Moreover, the condition that labor's relative share be the same function of r for every industry surely imposes *some* restrictions on the production functions. For example, it is certainly sufficient that

$$Q_j = F^j(L_j, K_{1j}, \ldots, K_{nj}) =$$
$$L_j^{\beta}[\phi^j(K_{1j}, \ldots, K_{nj})]^{1-\beta}, \quad j = 1, \ldots, n,$$

where $0 < \beta = \mathit{constant} < 1$ and $\phi^j(\cdot)$ is concave and homogeneous of degree one. In this case, of course,

$$\text{labor's relative share} = \frac{1}{\alpha(r)} = \beta, \quad j = 1, \ldots, n.$$

In general, when β may vary with r, one wonders what *necessary* functional form restrictions are implied by the Brown-Chang-Marx condition and how these restrictions relate to those of Solow, Gorman, Fisher, and others.

In section 7.11 Brown makes the important distinction between nominalistic and structural aggregation. He points out the paradox that nominalistic aggregation based upon the observed constancy of relative prices is usually a prerequisite for determining whether the underlying functional forms themselves allow aggregation. There is an "uncertainty principle" at work.

To estimate functional forms directly, we would require unavailable microeconomic data. Thus, as a practical necessity, nominalistic aggregation is required to estimate a set of production functions and to ask whether they satisfy any known sufficient conditions allowing additional aggregation. But, even if we discover that the answer is yes, we can never be certain that this affirmative conclusion is true in general because the conditions that permitted nominalistic aggregation in the first place may not remain valid over time. Moreover, the Hicks composite commodity approach yields an aggregate function that is related in a very complex way to the underlying micro functions; there might be an identification problem whereby certain specific restrictions on the micro functions cannot be tested using the aggregate function.

I do not think we should be too apologetic about this result. After all, economists are confronted with an impossible task when they are asked to estimate production functions without all the microeconomic input data! It *is* progress to recognize logical impossibilities, even when they are distressing.

It is difficult to find an optimistic note on which to close. My conclusion is that, given the current state of the art, the real-world facts contradict every set of conditions that would allow for theoretically rigorous capital aggregation. So where do we go from here? Three avenues of research remain relatively unexplored:

1. Further analysis of the Fisher type of Monte Carlo experiments may at least help us to understand more precisely the reasons why an aggregate production function sometimes "works," at least for tracking wages and, to a lesser degree, output. Although research in this direction is probably tedious, presumably some approximation theorems can be proved that would indicate error bounds on aggregate production function predictions for certain specified microeconomic structures.

2. Statistical cases can be made for aggregation in some instances, as Brown discusses in section 7.12, and perhaps further research in this direction will yield fruitful results.

3. Derivation of production functions from underlying engineering data remains an unexplored area, although it is unclear whether such derivations will yield results that permit aggregation of heterogeneous inputs.

I am not very optimistic about success along any of these roads; one must ponder what to do if we are dissatisfied with the theoretical foundations of current econometric work. I have one revolutionary suggestion: Perhaps for the purpose of answering many macroeconomic questions—particularly about inflation and unemployment—we should disregard the concept of a production function at the microeconomic level. The economist who succeeds in finding a suitable replacement will be a prime candidate for a future Nobel prize.

Finally, I turn to an additional difficulty that precludes aggregation of many multisector models into a *dynamic* one-sector Solow-Swan model. First, consider a dynamic multisector model in which prices are predicted with perfect short-run foresight; that is, $E(\dot{p}_i/p_i) = \dot{p}_i/p_i$ for all commodities. The work by Hahn (1966), Samuelson (1967, 1972*a*), Kuga (1977), myself,[4] and others shows that the rest point or steady-state equilibrium for such a model is not stable, but rather it is a saddle point in the space of capital-labor ratios and relative prices. Thus *any* aggregation procedure that gives rise to a dynamically stable evolution for an index of the capital-labor ratio incorrectly reflects the inherent instability of the underlying microeconomic model.[5]

Second, suppose we follow the Burmeister and Graham (1974, 1975) adaptive type of price expectations mechanism. Then stability is possible, but so far we know of only very stringent sufficient stability conditions; for example, the inverse of the input coefficient matrix must have a negative diagonal and positive off-diagonal elements at all feasible factor price ratios.[6] In addition to the restrictions imposed by aggregation, we now must *assume* that some such stability conditions hold, for the microeconomic data required to test for stability conditions are unavailable.

This is an especially unhappy state of affairs because it is completely unrealistic to assume that our observed data always are generated by steady-state equilibria. We must look at dynamic microeconomic structure; if aggregation to a stable one-sector model is possible, it is probably necessary that unstable microeconomic components of an aggregate index would cancel out to yield dynamic stability of the index.

I conjecture that something close to the converse also is true. That is, if a multisector model admits aggregation to a stable one-sector model without assuming that the economy is always in steady-state equilibrium, then most likely the underlying microeconomic model is stable.

In closing I note that, if this conjecture is correct, then there is a serious conceptual difficulty. The problem is not merely aggregation, but the fact that we do not yet have any satisfactory theoretical justifications for supposing stability of disaggregated dynamic models with heterogeneous capital goods.

Notes

1. See Burmeister and Dobell (1970, theorem 5, p. 279).
2. See Burmeister (1977). Also see Burmeister and Dobell (1970, pp. 282–94), Burmeister and Turnovsky (1972), Burmeister (1974), Brock and Burmeister

(1976), and Burmeister (1976, 1979). The work of Burmeister and Hammond (1977) proves that points for which the expression

$$\sum_{i=1}^{n} p_i \frac{dk_i}{dr}$$

is positive are dynamically unstable rest points if the economy follows a max-min rule.

3. These results about "equal organic composition of capital" are stated and proved in Burmeister (1979). The so-called transformation problem between Marxian values and competitive prices arises because "equal organic composition of capital" is a freak case; see, for example, Samuelson (1972*b*, *c*).

4. See Burmeister and Dobell (1970, pp. 297–306), Burmeister and Graham (1974, 1975), Brock and Burmeister (1976), and Burmeister et al. (1973).

5. Preliminary computer simulations suggest that divergence away from steady-state equilibrium may be quite rapid; see Burmeister et al. (1973).

6. Stability of a heterogeneous capital good model with technological change is another formidable problem, except in the special case when there is labor-augmenting technical progress at the *same rate* in every sector.

References

Brock, William A., and Burmeister, Edwin. 1976. Regular economies and conditions for uniqueness of steady states in optimal multi-sector economic models. *International Economic Review* 17 (February): 105–20.

Brown, M., and Chang, W. W. 1976. Capital aggregation in a general equilibrium model of production. *Econometrica* 44: (November): 1179–1200.

Burmeister, E. 1974. Synthesizing the neo-Austrian and alternative approaches to capital theory: A survey. *Journal of Economic Literature* 12 (June): 413–56.

———. 1976. Real Wicksell effects and regular economies. In *Essays in modern capital theory*, ed. M. Brown, K. Sato, and P. Zarembka, pp. 145–64. Amsterdam: North-Holland.

———. 1977. On the social significance of the reswitching controversy. *Revue d'Economie Politique*. March–April.

———. 1979. Critical observations on the labor theory of value and Sraffa's standard commodity. Discussion paper, University of Virginia.

Burmeister, E.; Caton, C.; Dobell, A. R.; and Ross, S. A. 1973. The saddlepoint property and the structure of dynamic heterogeneous capital good models. *Econometrica* 41:79–96.

Burmeister, E., and Dobell, A. R. 1970. *Mathematical theories of economic growth*. New York: Macmillan.

Burmeister, E., and Graham, D. A. 1974. Multi-sector economic models with continuous adaptive expectations. *Review of Economic Studies* 41:323–36.

———. 1975. Price expectations and global stability in economic systems. *Automatica* 11:487–97.

Burmeister, E., and Hammond, P. J. 1977. Maximin paths of heterogeneous capital accumulation and the instability of paradoxical steady states. *Econometrica* 45:853–71.

Burmeister, E., and Turnovsky, Stephen J. 1972. Capital deepening response in an economy with heterogeneous capital goods. *American Economic Review* 62 (December): 842–53.

Fisher, F. 1965. Embodied technical change and the existence of an aggregate capital stock. *Review of Economic Studies* 32:263–88.

Fisher, F. 1968*a*. Embodied technology and the existence of labour and output aggregates. *Review of Economic Studies* 35:391–412.

———. 1968*b*. Embodied technology and the aggregation of fixed and movable capital goods. *Review of Economic Studies* 35:417–28.

——— 1969. The existence of aggregate production functions. *Econometrica* 37:553–77.

Gorman, W. M. 1953. Community preference fields. *Econometrica* 21:63–80.

———. 1959. Separable utility and aggregation. *Econometrica* 27:469–81.

Hahn, F. H. 1966. Equilibrium dynamics with heterogeneous capital goods. *Quarterly Journal of Economics* 80:633–46.

Hicks, J. 1946. *Value and capital.* London: Oxford University Press.

Houthakker, H. S. 1955–56. The Pareto distribution and the Cobb-Douglas production function in activity analysis. *Review of Economic Studies* 23:27–31.

Kuga, K. 1977. General saddlepoint property of the steady state of a growth model with heterogeneous capital goods. *International Economic Review* 18 (February): 29–59.

Leontief, W. W. 1947. Introduction to a theory of the internal structure of functional relationships. *Econometrica* 15:361–73.

Paradoxes in capital theory: A symposium. 1966. *Quarterly Journal of Economics* 80, no. 4 (November): 503–83.

Samuelson, P. A. 1967. Indeterminacy of development in a heterogeneous-capital model with constant saving propensity. In *Essays on the theory of optimal economic growth*, pp. 219–31. Cambridge, Mass.: MIT Press.

———. 1972*a*. The general saddlepoint property of optimal-control motions. *Journal of Economic Theory* 5:102–20.

———. 1972*b*. The "transformation" from Marxian "values" to competitive "prices": A process of rejection and replacement. In *The collected scientific papers of Paul A. Samuelson*, ed. Robert C. Merton, vol. 3. Cambridge: M.I.T. Press.

———. 1972*c*. Understanding the Marxian notion of exploitation: A summary of the so-called transformation problem between Marxian values and competitive prices. In *The collected scientific papers of Paul A. Samuelson*, ed. Robert C. Merton, vol. 3. Cambridge: M.I.T. Press.

Shephard, R. 1953. *Cost and production functions*. Princeton: Princeton University Press.

Strotz, R. 1959. The utility tree: A correction and further appraisal. *Econometrica* 27:482–88.

This Page Intentionally Left Blank