ESTUDIOS DE ECONOMÍA APLICADA VOL. 24-1, 2006. PÁGS. 127-163

Aggregate production functions, neoclassical growth models and the aggregation problem

JESUS FELIPE(*) y FRANKLIN M. FISHER(**)
(*) Macroeconomics and Finance Research Division. Economics and Research Department. Asian Development Bank. Manila, Philippines; (**) Massachusetts Institute of Technology. Cambridge, MA 02139-4307
E-mails: (*) jfelipe@adb.org - (**) ffisher@mit.edu

ABSTRACT

Lawrence R. Klein pioneered the work on aggregation, in particular in production functions, in the 1940s. He paved the way for researchers to establish the conditions under which a series of micro production functions can be aggregated so as to yield an aggregate production function. This work is fundamental in order to establish the legitimacy of theoretical (neoclassical) growth models and empirical work in this area (e.g., growth accounting exercises, econometric estimation of aggregate production functions). This is because these models depend on the assumption that the technology of an economy can be represented by an aggregate production function, i.e., that the aggregate production function exists. However, without proper aggregation one cannot interpret the properties an aggregate production function. The aggregation literature showed that the conditions under which micro production functions can be aggregated so as to yield an aggregate production function are so stringent that it is difficult to believe that actual economies can satisfy them. These results question the legitimacy of growth models and their policy implications. Scientific work cannot proceed as if production functions existed. For this reason, the profession should pause before continuing to do theoretical and applied work with no sound foundations and dedicate some time to studying other approaches to estimating the impact of economic policies in order to understand what questions can legitimately be posed to the empirical aggregate data.

Keywords: Aggregate Production Function, Aggregation Problem, Aggregation Theorems.

Funciones agregadas de producción, modelos neoclasicos de crecimiento y el problema de la agregación

RESUMEN

Lawrence R. Klein fue uno de los pioneros del campo de la agregación, en particular en el área de las funciones de producción, durante la década de los 40. Sus contribuciones ayudaron a definir el problema de la agregación para que investigadores posteriores establecieran formalmente las condiciones formales bajo las que funciones de producción microeconómicas con propiedades neoclásicas pudieran ser agregadas con el fin de generar una función de producción agregagada. Esto es fundamental a la hora de justificar la legitimidad de modelos teóricos neoclásicos, así como trabajos empíricos en el área de crecimiento (como, por ejemplo, la contabilidad del crecimiento, o las estimaciones econométricas de funciones agregadas de producción), los cuales dependen de la hipótesis de que la tecnología de la economía puede ser representada por una función de producción agregada (es decir, la hipótesis de que la función de producción agregada existe). La literatura sobre la agregación ha demostrado que las condiciones bajo las que una serie de funciones de producción microeconómicas pueden ser agregadas y así generar la función de producción agregada son tan sumamente restrictivas que es difícil creer que las economías reales las satisfacen. Estos resultados cuestionan la legitimidad de los modelos de crecimiento neoclásicos y sus implicaciones. La conclusión es que si la economía se precia de tener carácter cientifico alguno no puede evolucionar bajo la premisa falsa de que las funciones de producción agregadas existen. Por ello, la profesión debería reflexionar antes de continuar desarrollando modelos de crecimiento teóricos de corte neoclásico y haciendo trabajo empírico sin fundamento teórico sólido.

Palabras Clave: Funciones Agregadas de Producción, Problema de la Agregación, Teoremas de Agregación.

We are grateful to two anonymous referees for their comments. The usual disclaimer applies. This paper represents the views of the authors and should not be interpreted as reflecting those of the Asian Development Bank, its executive directors, or the countries that they represent.

JEL classification: C43, B41, E01, E1, E23

Artículo recibido en Septiembre de 2005 y aceptado para su publicación en octubre de 2005.
Artículo disponible en versión electrónica en la página www.revista-eea.net, ref.: e-24101.

ISSN 1697-5731 (online) – ISSN 1133-3197 (print)

AGGREGATE PRODUCTION FUNCTIONS, GROWTH MODELS AND THE AGGREGATION PROBLEM[1]

The conceptual basis for believing in the existence of a simple and stable relationship between a measure of aggregate inputs and a measure of aggregate output is uncertain at best. Yet an aggregate production function is a very convenient tool for theoretically exploring some of the determinants of economic growth, and it has served as a framework for some interesting empirical studies

Richard Nelson (1964, p.575)

I have never thought of the macroeconomic production function as a rigorous justifiable concept. In my mind, it is either an illuminating parable, or else a mere device for handling data, to be used so long as it gives good empirical results, and to be abandoned as soon as it doesn't, or as soon as something better comes along

Robert Solow (1966, pp.1259-1260)

Arguably the aggregate production function is the least satisfactory element of macroeconomics, yet many economists seem to regard this clumsy device as essential to an understanding of national income levels and growth rates

Jonathan Temple (1999, p.15)

1. INTRODUCTION

We are delighted to have been invited to write a paper on methodological issues relating to growth to honor the work and contributions of Lawrence R. Klein. The topic that we have chosen is one in which (as perhaps many people do not know) Lawrence Klein was a pioneer in the 1940s. This is the so-called *aggregation problem in production functions*. The work of Klein and others in the 1940s led to a rich literature that discussed the conditions under which micro production functions could be aggregated so as to yield an aggregate production function.

The reader will appreciate the importance of the question we are raising in the context of the general topic of the paper, i.e., methodological issues relating to growth. Production functions are the pillar of *neoclassical* growth models. In these models, the constraint on growth is represented, on the supply side, by the production function. Since most growth models refer to the economy as a whole (or at least to

[1] For a more detailed discussion of many of the matters treated here, see Felipe and Fisher (2003).

large sectors of the economy), the production function used is, in fact, an *aggregate* production function. This is defined as a function that maps aggregate inputs into aggregate output. But what exactly does this mean? Such a concept has been implicit in macroeconomic analyses for a long time. However, it has always been plagued by conceptual confusions, in particular as to the link between the underlying micro production functions and the aggregate macro production function, the latter thought to summarize the alleged aggregate technology.

To understand what an aggregate production function is one must understand what the aggregation problem involves. The issue at stake is how economic quantities are measured; in particular those quantities that represent by a single number a collection of heterogeneous objects; in other words, what is the legitimacy of aggregates such as investment, GDP, labor and capital? To take a simple example, the question is as follows. Suppose we have two production functions $Q^A = f^A(K_1^A, K_2^A, L^A)$ and $Q^B = f^B(K_1^B, K_2^B, L^B)$ for firms A and B, where $K_1 = K_1^A + K_1^B$, $K_2 = K_2^A + K_2^B$ and $L = L^A + L^B$ (K refers to capital – two types- and L to labor –assumed homogeneous) The problem is to determine whether and in what circumstances there exists a function $K = h(K_1, K_2)$ where the aggregator function h(•) has the property that $G(K, L) = G[h(K_1, K_2)\ L] = \Psi(Q^A, Q^B)$, and the function Ψ is the production possibility curve for the economy.[2]

The problem is, as this paper shows, that the aggregate or macro production function is a fictitious entity. At the theoretical level it is built by adding micro production functions. However, there is an extensive body of literature that has shown that aggregating micro production functions into a macro production function is extremely problematic. This is the subject of the so-called aggregation literature and the issue at hand is referred to as the *aggregation problem*. Its importance lies in the fact that without proper aggregation we cannot interpret the properties of an aggregate production function. And without the latter, therefore, it is impossible to build a *neoclassical* growth model. If this is the case, one wonders about Nelson's (1964), Solow's (1966) and Temple's (1999) comments above: it cannot be true that an aggregate production function is a *convenient* tool for theoretical work. These authors are

[2] It will be noted that above we have already assumed that a production function exists at the level of the firm. In one sense, this is guaranteed. If an entity assigns the use of its various factors to different techniques of production so as to maximize output, then maximized output will depend only on the total amount of such factors, and that dependence can be written as a functional relationship (differentiability, of course, will not be guaranteed. We assume differentiability below only for convenience). That does not mean that one can aggregate over factors, and that is one part of the aggregation problem; the other one is aggregation over firms –aggregation to the case where factors are not all efficiently assigned.

clearly thinking of neoclassical growth models. Moreover, if the production function is a tool without sound theoretical foundations, one cannot help but wondering about much of the empirical work undertaken estimating it econometrically.[3] From a methodological point of view, the issue is interesting, for during the 1970s and 1980s, macroeconomics changed radically, in the sense that many economists (certainly a very influential group) insisted that macroeconomics should have microeconomic foundations. However, the aggregate production function, one of the key tools, does not have such foundations. At times, authors argue in their theoretical models that their production functions are microeconomic. At the empirical level, however, all the applications use aggregates.

The purpose of this paper is to provide an up-to-date account of the not well know aggregation problem and its devastating implications for the concept of an aggregate production function. With the surge of the new neoclassical endogenous growth literature in the 1980s (Barro and Sala-i-Martin 1995; Aghion and Howitt 1998) there has been a renewed interest in growth and productivity studies, propagated by the development of new models, the availability of large data sets with which to test the new and the old growth theories (e.g., Mankiw et al.'s 1992 use of the Summers and Heston data set), and episodes of growth that need to be explained and which have led to important debates, e.g., the East Asian Miracle (Young 1995). However, the references to the aggregation problem are absent. Temple (1999) is an exception, although, as in the case of Nelson (1964), we wonder why the least satisfactory tool in macroeconomics is an essential device.[4] Hence the relevance of the topic we discuss.

As mentioned above, Klein (1946a, 1946b) was one of the first economists to offer a systematic treatment of the aggregation problem in production functions. However, the aggregation literature truly flourished during the 1960s, parallel to but completely disconnected from the developments in growth at the theoretical (e.g., Solow 1956,

[3] In broad terms, aggregate production functions are estimated empirically for the following purposes: (i) to obtain measures of the elasticity of substitution between the factors, and the factor- demand price elasticities. Such measures are used for predicting the effects upon the distribution of the national income of changes in technology or factor supplies (Ferguson 1968); (ii) to apportion total growth into the accumulation of factors of production and technical change between two periods (Solow 1957); (iii) to test theories and quantify their predictions (Mankiw et al. 1992); and (iv) to address policy issues (Jorgenson and Yun 1984). Thus, from this point of view the most important question is the following: is the aggregate production function a summary of the "aggregate" technology? That is, suppose one estimates econometrically an aggregate production function: are the estimated coefficients (i.e., input elasticities, elasticity of substitution) really interpretable as technological parameters?

[4] In a more recent paper, Temple (2003) analyzes the long-run implications of growth models. While he makes a number of references to the assumptions and abstractions made to construct neoclassical growth models (e.g., linearity, Cobb-Douglas technology), there is no single reference to the fact that that an aggregate production function is *assumed* to exist.

Cass 1965) and applied levels (e.g. Denison 1961, 1972a, 1972b, Jorgenson and Griliches 1967, Jorgenson 1972). As the reference to Nelson (1964) above shows, at the time, authors were well aware of the issues at stake (a different question is that of why, despite knowing the problems, they decided to use this tool). Although Lawrence Klein pioneered this work, his aggregation method was not successful for it ran into a series of problems. This does not undermine the importance of his seminal work and the fact that he understood the significance of the problem and set the pace for future research in this area. Today, unfortunately, most (growth) economists are not even aware of the aggregation problems, which make their models absolutely fictitious and irrelevant. Those problems imply that their models are built on false presumptions and, consequently, cannot have the policy implications the authors derive from them.

It is important to mention that the aggregation literature was developed under the shadow of the so-called Cambridge-Cambridge debates and even today, the very few references made to the problems underlying the notion of an aggregate production function are expressed in terms of the Cambridge debates, and not in terms of the aggregation problem.[5]

The above is important for two reasons. First, the issues dealt with in the Cambridge-Cambridge debates and in the aggregation literature, though related, were not the same. Perhaps a useful and clarifying way to think about the Cambridge debates and the aggregation problem is to consider whether the measurement of capital problem relates to the interdependence of prices and distribution (the Cambridge-Cambridge debates), or whether it emerges out of the need to justify the use of the neoclassical aggregate production function in building theoretical models, and in empirical testing (the aggregation problem). Both problems can be present at once, of course, but they are not the same.

Second, the irony is that, perhaps, the aggregation problem is more damaging to the notion of an aggregate production function that the Cambridge-Cambridge debates. This is because while many neoclassical economists feel uncomfortable discussing the Cambridge debates because they easily lead to unending disputes (e.g., about whether the distribution of the product between the social classes is determined by the marginal products), the aggregation problem is a technical question posed in terms that neoclassical economists are familiar with. As such, it has technical answers.

The rest of the paper is structured as follows. Section 2 discusses the seminal work of Lawrence Klein on aggregation. Section 3 discusses the also early work on aggregation by Leontief and Nataf. This is followed in section 4 by a more extensive discussion of Fisher's aggregation conditions as well as those of Gorman. Section 5 offers a brief summary of the Houthakker-Sato aggregation conditions. Section 6 discusses some implications of the aggregation results for growth analysis. Section 7

5 For a recent summary of the Cambridge-Cambridge debates see Cohen and Harcourt 2003).

discuses why economists continue using aggregate production functions and refutes the standard replies. Finally, section 8 offers some conclusions.

2. KLEIN'S AGGREGATION CONDITIONS

Klein (1946a) initiated the first debate on aggregation in production functions by proposing methods for simultaneously aggregating over inputs and firms regardless of their distribution in the economy. Klein wanted to establish a macroeconomic system consistent with, but independent of, the basic microeconomic system. He thus approached the problem assuming as given both the theory of micro- and macroeconomics, and then tried to construct aggregates (usually in the form of index numbers) which were "consistent with the two theories" (Klein 1946a, p.94). He argued that the aggregate production function should be strictly a technical relationship, akin to the micro production function, and objected to utilizing the entire micromodel with the assumption of profit-maximizing behavior by producers in deriving the production functions of the macro-model.

The question Klein posed was whether one could obtain macroeconomic counterparts of micro production functions and the equilibrium conditions that produce supply-of-output and demand-for-input equations in analogy with the micro system. Klein argued that "there are certain equations in microeconomics that are independent of the equilibrium conditions and we should expect that the corresponding equations of macroeconomics will also be independent of the equilibrium conditions. The principal equations that have this independence property in microeconomics are the technological production functions. The aggregate production function should not depend upon profit maximization, but purely on technological factors" (Klein 1946b, p.303).

Klein proposed two criteria that aggregates should satisfy: (i) if there exist functional relations that connect output and input for the individual firm, there should also exist functional relationships that connect aggregate output and aggregate input for the economy as a whole or an appropriate subsection; and (ii) if profits are maximized by the individual firms so that the marginal-productivity equations hold under perfect competition, then the aggregative marginal-productivity equations must also hold (this criterion cannot be satisfied without the first). The first criterion means that an aggregate output must be independent of the distribution of the various inputs, that is, output will depend only on the magnitude of the factors of production, and not on the way in which they are distributed among different individual firms, nor in the way in which they are distributed among the different types of factors within any individual firm. The second criterion is that the aggregate production function so constructed should indeed behave like a production function in all respects.

Algebraically, Klein's problem is as follows. Suppose there are M firms in a sector, each of which produces a single product using N inputs (denoted x). Let the technology of the ν-th firm be representable as $y^{\nu} = f^{\nu}(x_1^{\nu},...,x_N^{\nu})$. Klein's aggregation problem over sectors can be phrased as follows: what conditions on the firm production functions will guarantee the existence of: (i) an aggregate production function G; (ii) input aggregator functions $g_1,\ldots, g_N$; and (iii) an output aggregator function F such that the equation:

$$F(y^1,...,y^M) = G([g_1(x_1^1,...,x_1^M), g_2(x_2^1,...,x_2^M),...,g_N(x_n^1,...,x_N^M)] \qquad (1)$$

holds for a suitable set of inputs x_n^m $(m = 1,...M; n = 1...N)$.

Klein used Cobb-Douglas micro production functions. He suggested that an aggregate (or strictly, an average) production function and aggregate marginal productivity relations analogous to the micro-functions could be derived by constructing weighted geometric means of the corresponding micro variables, where the weights are proportional to the elasticities for each firm. The elasticities of the macro-function are the weighted average of the micro-elasticities, with weights proportional to expenditure on the factor. The macro revenue is the macro price multiplied by the macro quantity, which is defined as the arithmetic average of the micro revenues (similar definitions apply to the macro wage bill and macro capital expenditure).

Klein's treatment of the problem, however, was rejected altogether by May (1947), who argued that even the firm's production function is not a purely technical relationship, since it results from a decision-making process. May argued:

"...The aggregate production function is dependent on all the functions of the micromodel, including the behavior equations such as profit-maximization conditions, as well as upon all exogenous variables and parameters. This is the mathematical expression of the fact that the productive possibilities of an economy are dependent not only upon the productive possibilities of the individual firms (reflected in production functions) but on the manner in which these technological possibilities are utilized, as determined by the socio-economic framework (reflected in behavior equations and institutional parameters). Thus the fact that our aggregate production function is not purely technological corresponds to the social character of aggregate production. Moreover, if we examine the production function of a particular firm, it appears that it, too, is an aggregate relation dependent upon nontechnical as well as technical facts. It tells us what output corresponds to total inputs to the firm of the factors of production, but it does not indicate what goes on within the firm"

(May 1947, p.63).[6]

[6] The standard procedure in neoclassical production theory is to begin with micro production functions and then derive equilibrium conditions that equate marginal products of inputs to their real prices. The solution to the system of equations given by the technological relationship and the equilibrium price equations yields the supply-of-output and demand-for-input equations as functions of output and input prices. And adding these equations over all firms yields the macro demand and supply equations. Note, however, as May pointed out that not even micro production functions are simply technological relations but assume an optimization process by engineers and management.

Thus, the macro production function is a fictitious entity, in the sense that there is no macroeconomic decision-maker that allocates resources optimally. The macro function is built from the micro units assumed to behave rationally.[7] Moreover, Klein's approach ran into two additional obstacles. First, Klein's problem was not the same as that of deriving the macro-model from the micro model. In fact his macro model does not follow from the micro model. Both are taken as given, and it is the indices that are derived. A second problem was pointed out by Walters (1963, pp.8-9). Walters noted that Kleinian aggregation over firms has some serious consequences. The definition of the macro wage bill (i.e., the product of the macro wage rate times the macro labor) is $W\,L = \frac{1}{n}\sum_{i=1}^{n} W_i L_i$, where W_i and L_i are the wage rate and homogeneous labor employed in the i-th firm, and $L = \prod_{i=1}^{n} L_i^{\alpha_i / \sum \alpha_i}$ is the definition of the macro-labor input, a geometric mean, where α_i is the labor elasticity of the i-th firm. In a competitive market, all firms have the same wage rate $W^*=W_i$ for all i. Substituting the macro-labor into the definition of the macro wage bill, and substituting W^* for W_i yields $W = W^* \frac{\sum L_i}{n\prod_{i=1}^{n} L_i^{\alpha_i / \sum \alpha_i}}$. This implies that the macro-wage will almost always differ from the common wage rate of the firms (similar issues apply to the prices of output and capital). It is therefore difficult to interpret W and to see why it should differ from W^*.

Despite these problems, Klein paved the way for work in the area of aggregation and for researchers to establish the conditions under which micro production function can be satisfactorily aggregated.

[7] As May pointed out, the aggregate production function cannot be considered purely technological. Even micro production functions do not give the output that is produced with given inputs. Rather, they give the *maximum* output that can be produced from given inputs. As Pu (1946) indicated, the macroeconomic counterpart of the equilibrium conditions holds if and only if Klein's aggregates arise from micro variables, all of which satisfy equilibrium conditions. Otherwise, the equilibrium conditions do not hold at the macro level. Thus, Klein's aggregates cannot be independent of equilibrium conditions if they are to serve the intended purpose.

3. THE LEONTIEF AND NATAF AGGREGATION THEOREMS

The next major result on aggregation was provided by Leontief (1947a, 1947b).[8] It deals with aggregation of variables into homogeneous groups. Leontief's (1947a) theorem provides the necessary and sufficient conditions for a twice-differentiable production function whose arguments are all non-negative, to be expressible as an aggregate. The theorem states that aggregation is possible if and only if the marginal rates of substitution among variables in the aggregate are independent of the variables left out of it. For the three-variable function $g(x_1, x_2, x_3)$ Leontief's theorem says that this function can be written as $G[h(x_1, x_2), x_3]$ if and only if $\frac{\partial(g_1/g_2)}{\partial x_3} \equiv 0$ where g_1 and g_2 denote the partial derivatives of g with respect to x_1 and x_2, respectively. That is, aggregation is possible if and only if the marginal rate of substitution between x_1 and x_2 is be independent of x_3. In general, the theorem states that a necessary and sufficient condition for the weak separability of the variables is that the marginal rate of substitution between any two variables in a group be a function only of the variables in that group, and therefore independent of the value of any variable in any other group.

In the context of aggregation in production theory (in the simplest case of capital aggregation), the theorem means that aggregation over capital is possible if and only if the marginal rate of substitution between every pair of capital items is independent of labor. Think of the production function $Q = Q(k_1, \ldots k_n, L)$. This function can be written as $Q = F(K, L)$, where $K = \phi(k_1, \ldots k_n)$ is the aggregator of capital, if and only if $\frac{\partial}{\partial L}\left(\frac{\partial Q/\partial k_i}{\partial Q/\partial k_j}\right) = 0$ for every i≠j. That is, the theorem requires that changes in labor, the non-capital input, do not affect the substitution possibilities between the capital inputs. This way, the invariance of the intra-capital substitution possibilities against changes in the labor input is equivalent to the possibility of finding an index of the quantity of capital. This condition seems to be natural, in the sense that if it were possible to reduce the n-dimensionality of capital to one, then it must be true that what happens in those dimensions does not depend on the position along the other axes (e.g., labor).

Note that Leontief's condition holds for aggregation within a firm and also for the economy as a whole even assuming that aggregation over firms is possible. Is Leontief's condition stringent assuming aggregation over firms? It will hold for cases such as brick and wooden buildings, or aluminum and steel fixtures. But most like-

[8] Leontief dealt with aggregation in general rather than only with production functions. For proofs of Leontief's theorem see the original two papers by Leontief; also, Green (1964, pp.10-15); or Fisher (1993, pp.xiv-xvi).

ly this condition is not satisfied in the real world, since in most cases the technical substitution possibilities will depend on the amount of labor. Think for example of bulldozers and trucks, or one-ton and two-ton trucks. In these cases no quantity of capital-in-general can be defined (Solow 1955-56, p.103).

Solow argued that there is a class of situations where Leontief's condition may be expected to hold. This is the case of three factors of production partitioned into two groups. For example, suppose $y_j = f^j(x_{0j}, x_j)$, j=1,2 where x_j is produced as $x_j = g^j(x_{1j}, x_{2j})$, so that the production of y_j can be decomposed into two stages: in the first one x_j is produced with x_{1j} and x_{2j}, and in the second stage x_j is combined with x_{0j} to make y_j. An example of this class of situations is that x_{1j} and x_{2j} are two kinds of electricity-generating equipment and x_j is electric power. In this case, the g^j functions are capital index functions (Brown 1980, p.389).[9]

The second important theorem is due to Nataf (1948). Besides the problem of aggregation of variables into homogeneous groups, there is the problem of aggregating a number of technically different microeconomic production functions. Nataf pointed out that Klein's (1946a) aggregation over sectors was possible if and only if all micro production functions were additively separable in capital and labor.

The problem here is as follows. Suppose there are n firms indexed by ν=1....,n. Each firm produces a single output Y(ν) using a single type of labor L(ν), and a single type of capital K(ν). Suppose that the ν–th firm has a two-factor production function $Y(\nu) = f^{\nu}\{K(\nu), L(\nu)\}$. To keep things simple, assume all outputs are physically homogeneous so that one can speak of the total output of the economy as $Y = \sum_{\nu} Y(\nu)$, and that there is only one kind of labor so that one can speak of total labor as $L = \sum_{\nu} L(\nu)$. Capital, on the other hand, may differ from firm to firm (although it may also be homogeneous). The question is: under what conditions can total output Y be written as $Y = \sum_{\nu} Y(\nu) = F(K, L)$ where $K = K\{K(1), ..., K(n)\}$ and $L = L\{L(1), ..., L(n)\}$ are indices of aggregate capital and labor, respectively? Nataf showed that the aggregates Y, L, K which satisfy the aggregate production function Y=F(K, L) exist, where the variables K(ν) and L(ν) are free to take on all values, if and only if every firm's production function is additively separable in labor and capital; that is, if every f^{ν} can be written in the form $f^{\nu}\{K(\nu), L(\nu)\} = \phi^{\nu}\{K(\nu)\} + \psi^{\nu}\{L(\nu)\}$. Assuming this to be so, the aggregate production relation can be written as Y=L+K,

[9] However, if there are more than two groups, Gorman (1959) showed that not only must the weak separability condition hold, but also each quantity index must be a function homogeneous of degree one in its inputs. This condition is termed "strong separability".

where, $Y = \sum_{v} Y(v)$, $L = \sum_{v} \psi^{v}\{L(v)\}$, and $K = \sum_{v} \phi^{v}\{K(v)\}$. Moreover, if one insists that labor aggregation be "natural", so that $L = \sum_{v} L(v)$, then all the, where c is the same for all firms. Nataf's theorem provides an extremely restrictive condition for intersectoral or even interfirm aggregation.[10] It makes one rather nervous about the existence of an aggregate production function unless there are some further restrictions on the problem.[11]

4. FISHER's AGGREGATION CONDITIONS

Years later, Fisher, in a series of papers summarized in Fisher (1969a), and reprinted with others in Fisher (1993), took up the issue Klein, Leontief and Nataf had started working on, and reminded the profession that, at any level of aggregation, the production function *is not* a description of what output can be produced from given inputs. Rather, the production function describes the maximum level of output that can be achieved if the given inputs are efficiently employed.

Fisher (1969a, 1993) observed that, taken at face value, Nataf's theorem essentially indicates that aggregate production functions almost never exist. Note, for example, that Nataf's theorem does not prevent capital from being physically homogeneous. Likewise, each firm's production function could perfectly exhibit constant returns to scale, thus implying that output does not depend on how production is divided among different firms, or even have identical technologies with the same kind of capital. As indicated previously, identity of technologies (e.g., all of them Cobb-Douglas) and constant returns do not imply the existence of an aggregate production function. Yet intuition indicates that under these circumstances one *should* expect an aggregate production function to exist. Something is wrong here.

Fisher pointed out that one must ask not for the conditions under which total output can be written as $Y = \sum_{v} Y(v) = F(K, L)$ under any economic conditions, "but rather

[10] For a number of applications of this result see Green (1964, chapter 5).

[11] Nataf's result can be proved using Leontief's theorem. By the latter, aggregation is possible if and only if the ratio of marginal products of capital in two firms independent of all labor inputs. But in Nataf's non-optimizing setup, the amount of labor in a given firm cannot influence the marginal product of capital in any other. Hence, Leontief's condition requires that it not influence the marginal product of capital in the given firm either. This way one obtains additive separability. The conclusion that the marginal product of labor must be constant and the same in all firms follows from the requirement that the labor aggregate is total L, so that reassigning labor among firms does not change total output.

for the conditions under which it can be so written *once production has been organized to get the maximum output achievable with the given factors*" (Fisher 1969a, p.556; italics in the original). This was, of course, the problem with Klein's original formulation. Thus "the problem with Nataf's theorem is not that it gives the wrong answer but that it asks the wrong question. A production function does not give the output that can be produced from given inputs; rather, it gives the *maximum* output that can be so produced. Nataf's theorem fails to impose an efficiency condition" (Fisher 1993, p.xviii; italics in the original). Thus, efficient allocation requires that Y be maximized given K and L. This is why optimization over the assignment of production to firms makes sense in constructing an aggregate production function. Competitive factor markets will do this. These considerations lead to an altogether different set of aggregation conditions.

This way of looking at the problem stressed that the aggregation problem does not only affect the aggregation of capital. It was pointed out that there exist equally important labor and output aggregation problems. Indeed, there would be aggregation problems even if each type of capital were physically homogeneous and the same in all firms. Indeed, even if there were only one type of capital, labor and output aggregation problems would continue to exist. Secondly, from the point of view of the aggregation literature the problem is whether an economy-wide (or a sector or indeed a firm) production function can be constructed that exhibits the properties needed to establish factor demand functions. Therefore, the aggregation problem emerges out of the need to justify the use of the neoclassical aggregate production function in building theoretical models, and in empirical testing.

4.1. Capital Aggregation

In the simplest case of capital is physically homogeneous, where total capital can be written as $K = \sum_{v} K(v)$, efficient production requires that aggregate output Y be maximized given aggregate labor (L) and aggregate capital (K). Under these simplified circumstances, it follows that $Y^M = F(K, L)$ where Y^M is maximized output, since, as was pointed out by May (1946, 1947), individual allocations of labor and capital to firms would be determined in the course of the maximization problem (note that without optimal allocation even factor homogeneity does not help). This holds even if all firms have different production functions and whether or not there are constant returns.

In the (somewhat) more realistic case where only labor is homogeneous and technology is embodied in capital, Fisher proposed to treat the problem as one of labor being allocated to firms so as to maximize output, with capital being firm-specific.

As argued above, when a particular factor is homogeneous across firms and allocated optimally across firms, aggregation does not pose a special problem as regards that factor. But when capital is not homogeneous, i.e., firms use different techniques, one cannot add up heterogeneous quantities meaningfully unless there is some formula that converts heterogeneous items into homogeneous units.

Fisher (1965) pointed out that it is important to remark that the assumption that technology is embodied in capital (i.e., capital is firm-specific) induces difficulties whether or not a capital aggregate exists for each firm. However, no such difficulties exist as to aggregate labor *if* there is only one type of labor. The reason is that labor is assumed to be assigned to firms efficiently. Now, given that output is maximized with respect to the allocation of labor to firms, and denoting such value by Y^*, the question is: under what circumstances is it possible to write total output as $Y^* = F(J, L)$ where $J = J\{K(1),..., K(n)\}$, where $K(\nu)$, ν=1…n, represents the stock of capital of each firm (i.e., one kind of capital per firm)? Since the values of $L(\nu)$ are determined in the optimization process there is no labor aggregation problem. The entire problem in this case lies in the existence of a capital aggregate. Recalling that Leontief's condition is both necessary and sufficient for the existence of a group capital index, the previous expression for Y^* is equivalent to $Y^* = G\{K(1),..., K(n), L\}$ if and only if the marginal rate of substitution between any pair of $K(\nu)$ is independent of L.

Fisher then proceeded to draw the implications of this condition for the form of the original firm production function. He found that under the assumption of strictly diminishing returns to labor (i.e., $f^{\nu}_{LL} < 0$), a necessary and sufficient condition for capital aggregation is that if any one firm has an additively separable production function (i.e., $f^{\nu}_{KL} \equiv 0$), then *every* firm must have such a production function.[12] This means that capital aggregation is not possible if there is both a firm which uses labor and capital in the same production process, and another one which has a fully automated plant.[13] More important, assuming constant returns to scale, the case of capital-augmenting technical differences (i.e., embodiment of new technology can be written as the product of the amount of capital times a coefficient) turns out to be *the only case* in which a capital aggregate exists. This means that each firm's production function must be writeable as $F(b_{\nu}K_{\nu}, L_{\nu})$, where the function $F(\cdot,\cdot)$ is common

[12] Here and later, such subscripts denote partial differentiation in the obvious manner.

[13] Strictly speaking, Fisher found that a necessary and sufficient condition for capital aggregation is that every firm's production function satisfy a partial differential equation in the form

$$\frac{f^{\nu}_{KL}}{f^{\nu}_{K}\, f^{\nu}_{LL}} = g(f^{\nu}_{L})$$, where g is the same function for all firms.

to all firms, but the parameter b_v can differ. Under these circumstances, a unit of one type of new capital equipment is the exact duplicate of a fixed number of units of old capital equipment ("better" is equivalent to "more"). As we would expect, given constant returns to scale, the aggregate stock of capital can be constructed with capital measured in efficiency units.[14] Summing up: aggregate production functions exist if and only if all micro production functions are identical except for the capital efficiency coefficient. Certainly this conclusion represents a step beyond Nataf's answer to the problem. But certainly it also continues to require an extremely restrictive aggregation condition, one that actual economies do not satisfy.

To see that this condition permits aggregation, consider two firms, and define $J = b_1K(1) + b_2K(2)$, with $L = L(1) + L(2)$. The sum of the outputs of the two firms is $Y = F(b_1K(1), L(1)) + F(b_2K(2), L(2))$. Since efficient allocation of labor requires that labor have the same marginal product in both uses, it is clear that when Y is maximized with respect to labor allocation, the ratio of the second argument to the first must be the same in each of the two firms. Thus, $\frac{L(1)}{b_1K(1)} = \frac{L(2)}{b_2K(2)} = \frac{L}{J}$ when labor is optimally allocated. If we let $\lambda = \frac{b_1K(1)}{J} = \frac{L(1)}{L}$ (this second equality holds when labor is optimally allocated). It then follows that $Y^* = F(\lambda J, \lambda L) + F((1-\lambda)J, (1-\lambda)L) = F(J, L)$ because of constant returns.[15]

But the bite of the theorem is that the capital-augmentation condition is necessary (as well as sufficient) for capital aggregation under constant returns. Thus, an implication of Fisher's work is the importance of the aggregation level. On the one hand the aggregation problem appears both with two firms and with one thousand. One the other hand, it is fair to say that the higher the number of firms, the *more likely* it is that they will differ in ways that prevent aggregation, for the more likely it will be

[14] Fisher (1965) indicates that he could not come up with a closed-form characterization of the class of cases in which an aggregate stock of capital exists when the assumption of constant returns is dropped. Nevertheless, as he shows, there do exist classes of non-constant returns production functions which do allow construction of an aggregate capital stock. Capital aggregation is possible under the restrictive assumption that the individual firm's production function can be made to yield constant returns after suitable "stretching of the capital axis" and there are other cases as well. On the other hand, if constant returns are not assumed there is no reason why perfectly well behaved production functions cannot fail to satisfy the partial differential equation given in the preceding footnote. Capital aggregation is then impossible if any firm has one of these "bad apple" production functions.

[15] This proof holds for any constant returns to scale production function. Of course this construction is only for the case in which (only) labor is optimally assigned.

that (at least) one of them will fail to satisfy the partial differential equation condition mentioned in a footnote above.

It is important to mention that, working with the profit function rather than with the production function, Gorman (1968) reached similar conclusions to those of Fisher on capital aggregation. Gorman used the restricted profit function to derive aggregation conditions. Gorman also set out to find what the technologies of the individual firms should be so that aggregates of fixed factors (e.g., buildings) would exist. These quantities are required to depend only on the amounts of the various types of equipment used in individual firms. Gorman showed that if the micro (labor optimized) variable profit functions Π^{*m} can be written as

$\Pi^{*m}(p,w,z^m) = b(p,w)\,h^m(z^m) + c^m(p,w)$ for m=1,2...M (sectors of the economy), then capital aggregation is possible (p is a vector of output and intermediate input prices; z^m is a vector of fixed capital input; and w is a vector of labor prices); that is, the macro (labor optimized) variable profit function Π^* can be written as

$$\Pi^*(p,w,z^1,...,z^M) = \sum_{m=1}^{M} \Pi^{*m}(p,w,z^m) = b(p,w)\left[\sum_{m=1}^{M} h^m(z^m)\right] + \sum_{m=1}^{M} c^m(p,w) \equiv$$

$$\equiv \Pi^*\left[p,w,\sum_{m=1}^{M} h^m(z^m)\right].$$ Therefore, the separability restriction on the micro production possibility sets is sufficient to imply the existence of the aggregate.

Is the restriction on the micro variable profit functions a stringent aggregation condition? Perhaps it is not very restrictive if every z^m is a scalar, that is, if there is only one fixed capital good for each sector. However, the restriction becomes more unrealistic from an empirical point of view as the number of fixed capital goods in each sector increases.

As extensions of his original work, Fisher considered the following cases:

(i) Fisher (1965) analyzed the case where each firm produces a single output with a single type of labor, but two capital goods, i.e., $Y(v) = f^v(K_1, K_2, L)$. Here Fisher distinguished between two different cases. First, aggregation across firms over one type of capital (e.g., plant, equipment). Fisher concluded that the construction of a sub-aggregate of capital goods requires even less reasonable conditions than for the construction of a single aggregate.[16] For example, if there are constant returns in K_1, K_2, and L, there will not be constant returns in K_1 and L, so that the difficulties of the

[16] The conditions turn out to be twofold: (i) $\dfrac{f^v_{K_1L}}{f^v_{K_1}\, f^v_L} = g(f^v_L)$; (ii) $f^v_{K_1K_2} - \dfrac{f^v_{K_1L}}{f^v_{K_1}\, f^v_L} = 0$

2-factor non-constant returns case appear. Further, if the ν–th firm has a production function with all three factors as complements, then no K_1 aggregate can exist. Thus, for example, if any firm has a generalized Cobb-Douglas production function (omitting the ν argument) in plant, equipment, and labor $Y = AK_1^{\alpha} K_2^{\beta} L^{1-\alpha-\beta}$, one cannot construct a separate plant or separate equipment aggregate for the economy as a whole (although this does not prevent the construction of a full capital aggregate).

The other case Fisher (1965) considered was that of the construction of a complete capital aggregate. In this case, a necessary condition is that it be possible to construct such a capital aggregate for each firm taken separately; and a necessary and sufficient condition (with constant returns), given the existence of individual firm aggregates, is that all firms differ by at most a capital augmenting technical difference. That is, they can differ *only* in the way in which their individual capital aggregate is constructed.

(ii) Fisher (1982) asked whether the crux of the aggregation problem derives from the fact that capital is considered to be an immobile factor. He argued the aggregation problem only seems to be due to the fact that capital is fixed and is not allocated efficiently. This is true in the context of a two-factor production function. However, if one works in terms of many factors, all mobile over firms, and asks when it is possible to aggregate them into macro groups, it turns out that the mobility of capital has little bearing on the issue. In fact, where there are several factors, each of which is homogeneous, optimal allocation across *firms* does not guarantee aggregation across *factors*. The conditions for the existence of such aggregates are still very stringent, but this has as much to do with the necessity of aggregating over firms as with the immobility of capital. A possible way of interpreting the existence of aggregates at the firm level is that each firm could be regarded as having a two-stage production process. In the first one, the factors to be aggregated, $X_i(\nu)$, are combined together to produce an intermediate output, $\phi^{\nu}(X(\nu))$. This intermediate output is then combined with the other factor, L(ν), to produce the final output. Aggregation of X can be done if and only if firms are either all alike as regards the first stage of production, or all alike as regards the second stage. If they are all alike as regards the first stage, then the fact that L is mobile plays no role. If, on the other hand, they are all alike as regards the second stage, then the fact that the X_i are mobile plays no role. These conditions imply that mobility of capital permits instant aggregation over firms of any one capital type across firms. However, the fact that aggregation over firms is involved, whether or not capital is fixed, restricts aggregation to the cases described above.[17]

[17] When there are more than two firms, aggregation over the entire set of firms requires aggregation over every pair (the two-firm case). This implies that an aggregate over n firms exists if and only if at least one of the following two holds: (i) All the $F^{\nu}(\bullet,\bullet)$ can be taken to be the same; (ii) All the $\phi^{\nu}(\bullet)$ can be taken to be the same.

(iii) Finally, Fisher (1983) is another extension of the original problem to study the conditions under which full and partial capital aggregates, such as "plant" or "equipment" would exist simultaneously. Not surprisingly, the results are as restrictive as those above. Fisher showed that the simultaneous existence of a full and a partial capital aggregate (e.g., plant) implies the existence of a complementary partial capital aggregate (e.g., equipment), and that the two partial capital aggregates must be perfect substitutes.[18]

4.2. Labor and Output Aggregation

Fisher (1968) extended his work on capital aggregation to the study of problems involved in labor and output aggregation, thus pointing out that the aggregation problem is not restricted to capital. Output aggregation and labor aggregation are also necessary only if one wants to use a sector-wide or economy-wide aggregate production function.

The problem that Fisher studied is in the context of cross-firm aggregation that arises because labors or outputs are shifted over firms, given the capital stocks and production functions, to achieve efficient production. That is, now there is a vector of labors $L_1(v),...,L_S(v)$ and a vector of outputs $Y_1(v),...,Y_S(v)$ (it does not matter whether there is one or more types of capital).[19] In the simplest case of constant returns, a labor aggregate will exist if and only if a given set of relative wages induces all firms to employ different labors in the same proportion. Similarly, where there are many outputs, an output aggregate will exist if and only if a given set of relative output prices induces all firms to produce all outputs in the same proportion.

The implication of these conditions is that the existence of a labor aggregate requires the absence of specialization in employment; and the existence of an output aggregate requires the absence of specialization in production, indeed all firms must produce the same market-basket of outputs differing only in their scale.[20]

---

[18] Blackorby and Schworm (1984) is an extension of Fisher (1983). By presenting an alternative formulation of the problem in which one can have both a full and a partial capital aggregate without the restrictive substitution implications derived by Fisher. They show that there need be only one partial aggregate and that if there are two partial aggregates, they need not be perfect substitutes. The conditions nevertheless remain very restrictive.

[19] An interesting issue in this context is that the aggregates of labor and output might exist for each firm separately, but not for all firms together. However, since this would imply some strange things about aggregation, Fisher assumed that an aggregate at the firm level exists. No similar problem arises in the case of capital, where aggregation over all firms requires the existence of an aggregate for each firm separately.

[20] The "same market basket" condition for output aggregation and the similar condition for labor aggregation are cases of the "common aggregator" condition in Fisher (1982) (see above). Blackorby and Schworm (1988) is an extension of Fisher (1968).

4.3. Fisher's Simulations

Fisher (1969a, pp.572-574) posed an interesting conundrum, namely, that despite the stringency of the aggregation conditions, the fact is that when one fits aggregate data on output to aggregate data on inputs, the results tend to be "good," meaning that the fit tends to be relatively high, and that in the case of the Cobb-Douglas, the elasticities are close to the factor shares in output. Furthermore, the wage rate is well explained by the marginal product. Fisher sketched several possible reasons for this paradox, of which he favored the following: for unspecified reasons, firms always invest in proportion (i.e., fixed ratios) to a particular index. In such case the index would be an approximate aggregate. And likewise, if outputs were always produced and labor hired in approximately fixed proportions, then an approximate output and labor aggregates would exist.

Fisher (1971) and Fisher et al. (1977) are two attempts at providing an answer to the question of why, despite the stringent aggregation conditions, aggregate production functions seem to work when estimated econometrically, and why the marginal product of labor appears to give a reasonable good explanation of wages. To answer the question, Fisher undertook a series of simulation analyses. The important aspect of the simulations is that the series were aggregated even though the aggregation conditions were violated. Under these circumstances, if the aggregate production function yields "good results," one cannot take it as evidence that the aggregate production function summarizes the true technology.

In the first of these papers, Fisher (1971) set up an economy consisting of N firms or industries (N=2, 4, or 8 in the simulations), each hiring the same kind of labor and producing the same kind of output. Each firm, however, has a different kind of capital stock, and its technology is embodied in that stock. This implies that capital canno not be reallocated to other firms. In the aggregation process, the conditions for successful aggregation were violated. The micro-production functions were Cobb-Douglas, and labor was allocated optimally to ensure that output was maximized. This economy was simulated over 20 periods. The total labor force, the firms' technology and their capital stocks were assumed to grow at a constant rate (with a small random term to reduce multicollinearity in the subsequent regression analysis). In certain of the experiments, some of these growth rates were set equal to zero and the growth of the capital stock was allowed to vary between firms.

Fisher observed that, in all his experiments (a total of 1010 runs each covering a 20-year period), the fit was around 0.99, although he pointed out that this "reflects the fact that with everything moving in trends of one sort of another, an excellent fit is obtained regardless of misspecifications of different sorts" (Fisher (1971, p.312).

The most important conclusion Fisher drew from his results was that as long as the share of labor happened to be roughly constant, the aggregate production function would yield good results, even though the underlying technical relationships are not consistent with the existence of any aggregate production function. And this conclusion remained even in cases where the underlying variables showed a great deal of relative movement. This suggests that the (standard) view that constancy of the labor share is due to the presence of an aggregate Cobb-Douglas production function is wrong. The argument runs the other way around, that is, the aggregate Cobb-Douglas works well *because* labor's share is roughly constant.

In a subsequent paper, Fisher et al. (1977) extended the simulation analysis to the case of the CES production function developed by Arrow et al. (1961). The simulations were similar in spirit to those in Fisher (1971), with the corresponding complications introduced by the fact that the micro production functions were CES and have more coefficients to parameterize (elasticity of substitution and distribution parameter). The objective was the same, that is, to learn when the CES, despite the aggregation problems, would perform well in empirical work. The aggregate series of output, labor, and capital were also generated following procedures similar to those in Fisher (1971). And the aggregation conditions for capital were violated as in Fisher (1971). Thus the authors stated that "the elasticity of substitution in these production functions is an "estimate" of nothing; there is no true aggregate parameter to which it corresponds" (Fisher et al. 1977, p.312). Each firm had a different elasticity of substitution, ranging between 0.25 and 2.495. For each choice of the elasticities of substitution, the distribution parameters were chosen in two sets, half the runs having distribution parameters and substitution elasticities positively correlated, and half of them negatively correlated (ranging between 0.15 and 0.35). The objective was to generate a labor share approximately of 0.75. It must be mentioned that in this paper, besides the aggregate CES, Fisher et al. (1977) also estimated the Cobb-Douglas, and the log-linear relationship implied by the CES with constant returns to scale, namely, $\ln(Y^* / L) = H + \sigma \log w$, where σ is the elasticity of substitution. They called the latter the "wage equation." This is used in what they refer to as the "hybrid estimate" of the wage equation and the production function. This was obtained imposing the elasticity of substitution estimated from the wage equation on the production function; and then they used the latter to estimate the distribution and efficiency parameters in the production function.

What conclusions did Fisher et al. (1977) reach? The fit in all cases was very good. They also established that the hybrid wage predictions were the best, and that the wage equation estimates of the elasticity of substitution are better than those given by the production function. Likewise, Fisher's earlier findings with Cobb-Douglas were confirmed in these simulations, i.e., the Cobb-Douglas works well when the observed factor share is fairly stable. But the authors failed to find any similar organizing principle with which to explain when the aggregate CES production function does or does not give good wage predictions. In other words, while in Fisher (1971) the organizing principle was that the aggregate Cobb-Douglas would work when factor shares were constant, in the case of the CES, they could not establish any similar "rule."

5. HOUTHAKKER – SATO AGGREGATION CONDITIONS

Whereas Fisher sought to develop conditions where aggregate production functions would always work, Houthakker and Sato considered cases in which the problem was restricted by assuming that the distribution of capital over firms remains constant. In such cases it is obvious that one can aggregate over capital. Houthakker and Sato's contributions (see also Levhari 1968) showed the relationships between the distribution of capital and the form of the aggregate production function.

Kazuo Sato's (1975) approach to the aggregation problem was based on the procedure that Hendrik Houthakker had developed years before. Houthakker (1955-56) postulated that factor proportions are distributed in a certain way among the firms over which the aggregation is to take place. He then showed that if individual production functions are of the fixed-coefficients type (not necessarily the same in each firm), and if the input-output ratios (the capacity density function) are distributed according to a Pareto distribution $Q = C\left(\frac{L}{Q}\right)^{\alpha_1 - 1}\left(\frac{K}{Q}\right)^{\alpha_2 - 1}$ with $\alpha_1 > 1$ and $\alpha_2 > 1$, then the aggregate production function is the Cobb-Douglas with decreasing returns to scale $Q = AL^{\alpha_1/(\alpha_1+\alpha_2+1)}K^{\alpha_2/(\alpha_1+\alpha_2+1)}$. The conclusion of Houthakker's model is that if all individual firms operate according to Leontief production functions, and if efficiencies are distributed according to a Pareto distribution, then the aggregate production function will be Cobb-Douglas.[21] In other words, while the aggregate production function has the appearance of a technology with an elasticity of substitution of unity, at the micro level there is no possibility of substitution between inputs.

Sato (1975) developed and extended the procedure introduced by Houthakker with a view to investigating how the macro behavior in production relates to the macro behaviors via the distribution of input coefficients. He allowed for elasticities of substitution to exceed zero, and the distribution function needed no longer be Pareto. Sato proceeded by splitting the aggregation problem into two sequential questions. First, suppose one has the production function $Q = Q(K_1,...,K_n, L)$. Then he asked: can this form be compressed into a form like $Q = F(K, L)$ by aggregating the vector of K's? In this step one must find both the capital aggregate K and the macro function F. Sato called this the *existence* problem. This must be done for each distribution. This will give rise to a series of F's. The second step was to ask for the conditions that the distributions must satisfy if they are to generate the same F. This is the *invariance* problem. As a corollary, Sato asked whether two entirely different distributions

[21] Levhari (1968) reversed Houthakker's procedure and derived the distribution of factor proportions for a CES production function.

could yield macro production functions $Q = F(K, L)$ identical in every respect. Sato showed that if the efficiency distribution is stable, the resulting estimates should reflect a production function. Thus, the key of this approach lies in the stability of the distribution function.

6. IMPLICATIONS FOR GROWTH ANALYSES

One implication of the aggregation results is that intuitions based on micro variables and micro production functions will often be false when applied to aggregates (Felipe and Fisher 2003, p.250). In this section we note three areas in growth analysis for which the aggregation results have serious implications. The first one is at the (theoretical) modeling level; the second one is at the policy level; and the third one is at the empirical level.[22]

Howitt (2002) offers one of the few instances in the modern literature where there is an acknowledgement that there is something wrong with the concept of an aggregate production function. Nevertheless, he argued that under simplifying assumptions, some endogenous growth models allow one to represent the economy's macro behavior with a Cobb-Douglas aggregate production function. One such case is Howitt (2000).[23] However, we believe this is not the case. The reason is that Howitt seems to be thinking of the aggregate production function as a *parable* (see section 7) and without considering the implications of the aggregation problems. This is so because the model he uses is conceived in the following terms:

"Consider a single country in a world economy with m different countries. There is one final good, produced under perfect competition by labor and a continuum of intermediate products, according to the production function

$$Y_t = \int_0^{N_t} A_i(i)F[x_t(i), L_t / N_t]di,$$

[22] It must be noted that we do mean to question each and every single growth regression. Our point is that the disregard for the aggregation literature poses very serious questions for the interpretation of many theoretical and empirical results.

[23] See Sylos-Labini (2001) and Howitt (2002). Silos-Labini's (2001) laid some criticisms against the new endogenous growth models for using aggregate production functions. He referred to the Cambridge, U.K. criticisms. In his review of Sylos-Labini's (2001) book, Howitt (2002) acknowledged the relevance of the Cambridge debates, but argued that the "criticisms laid out in this book are wide of the mark" and went on to argue that, under certain simplifying assumptions, some Schumpeterian endogenous growth models, allow one to represent the economy's macro behavior with a Cobb-Douglas aggregate production function. In private correspondence with Peter Howitt, he indicated that one such case is Howitt (2000). We are grateful to him. This does not necessarily mean that Peter Howitt endorses our views.

where Y_t *is the country's gross output at date* t*,* L_t *is the flow of labor used in production,* N_t *measures the number of different intermediate products produced and used in the country,* $x_t(i)$ *is the flow of output of intermediate product* $i \in [0, N]$*,* $A_t(i)$ *is the productivity parameter attached to the latest version of intermediate product* i*, and* $F(\bullet)$ *is a smooth, concave, constant returns production function. For simplicity attention is restricted to the Cobb-Douglas case:* $F(x, \ell) \equiv x^{\alpha}\ell^{1-\alpha}$, $0 < \alpha < 1$"

(Howitt 2000, p.831).

In our opinion, Howitt is thinking in terms of Samuelson's (1961-1962) parable, which was developed in the context of the Cambridge-Cambridge debates, and discussed in the next section. If this is the case, then Garegnani (1970) showed why Samuelson's parable was unconvincing. And from the point of view of the aggregation problem, Howitt's attempt to represent the economy's macro behavior with an aggregate Cobb-Douglas production function is utterly unconvincing as a case of proper aggregation.[24]

We must add to the above that since the 1980s, economists have been developing the so-called endogenous growth models. These models posit production functions with increasing returns to scale, an elasticity of capital of unity, external effects, or some combination of these. In order to assess the importance of these assumptions economists are estimating aggregate production functions for entire economies, for the manufacturing sector, or for more narrowly defined industry aggregates with a view to providing some evidence (or lack of it) of these assumptions. Romer (1987), Hall (1990), Caballero and Lyons (1992), Mankiw et al. (1992), Backus et al. (1992), Basu and Fernald (1995, 1997), Burnside (1996), *inter alios*, have attempted to document the empirical importance of the phenomena of increasing returns and externalities hypothesized by the new growth models. This work, however, faces the problem that, as noted above, except under constant returns, aggregate production functions are unlikely to exist at all. Thus, it is difficult to understand both at the theoretical and the empirical levels what this work ultimately does.[25]

At the policy level, Rodrik (2005) has argued that to the well-known problems of parameter heterogeneity, outliers, omitted variables, model uncertainty, measurement errors and endogeneity, discussed in the growth literature, one has to add the problem of the interpretation of the coefficients of estimated growth models "when policies are not random but are used systematically by governments to achieve certain ends –whether good or bad." Rodrik refers to the many regressions that have appeared during the last decade where researchers have regressed a country's growth rate (or it may be

[24] See the recent paper by Aghion and Howitt (2005) linking growth theory and policy analysis.
[25] See Felipe (2001) for a review of this work.

a cross-country regression) on a number of policies (e.g., fiscal policy, government consumption, inflation, black market premium on foreign exchange, overvaluation of the exchange rate, financial liberalization, trade policy, state ownership in industry and banking). Algebraically, these regressions take the form $g_i = \alpha h \; y_{i0} + Z_i'\beta + \gamma s_i + \varepsilon_i$, where g is the country's growth rate; y_0 is the level of initial income; s is a policy variable of country; and Z denotes a vector of other covariates. The objective of these regressions is to obtain an estimate of γ, the impact of policy intervention on growth. Often these regressions are specified ad-hoc, while other times they are obscurely related to a growth model.

Note that Rodrik is not writing from the view of the aggregation problem. However, the aggregation results imply that the policy implications of neoclassical growth models are very dubious, to say the least. Our view is that the profession should pause before continuing to do work without sound foundations. In particular, we should consider the possibility of modeling growth without resorting to a production function (e.g., McCombie and Thirlwall 1994). And second, we should dedicate some time to studying other approaches to estimating the impact of national economic policies in order to understand which questions can legitimately be posed to the empirical aggregate data.

Empirically, the non-existence of the aggregate production function poses a conundrum. If indeed aggregate production functions do not exist because they cannot be derived theoretically, there must be a reason (unrelated to the existence of the aggregate production function) why they seem to work empirically.[26] The answer has been in the literature for a long time (Simon and Levy 1963, Simon 1979, Shaikh 1974, 1980), and more recently Felipe (2001), Felipe and Adams (2005) and Felipe and McCombie (2001, 2002, 2003, 2005a, 2005b, 2005c) have elaborated upon it.[27] However, like the theoretical arguments underlying the non-existence of the aggregate production function, these arguments have also been ignored. The argument is that because at the aggregate level the data used in empirical applications are not physical quantities but values, the accounting identity that relates definitionally the value of total output to the sum of the value of total inputs can be rewritten as a form that resembles a production function. The implication is simply that the neoclassical growth model is no more than a series of non-refutable propositions. Since the argument is relatively

[26] A recent survey on growth econometrics by Durlauf et al. (2005) contains no single reference to the aggregation problem.

[27] See also Samuelson (1979), who seems to discover the same argument without any reference to previous discussions. Based on the argument, he raised very fundament questions about the original work by Douglas on the so-called Cobb-Douglas (aggregate) production function. See the symposium in the Eastern Economic Journal (2005), in particular the paper by Felipe and Adams (2005).

simple, it is worth summarizing it. The National Income and Products Account (NIPA) indicates that value added equals the wage bill plus total profits, that is,

$$V_t \equiv W_t + \Pi_t \equiv w_t L_t + r_t J_t \tag{1}$$

where V is value added, W is the total wage bill, Π denotes total profits (operating surplus in the NIPA terminology), w is the average real wage rate, L is employment, r is the average ex-post profit rate, and J is the deflated value of the stock of capital, with all variables expressed in real terms.[28] Expressing equation (1) in growth rates yields:

$$\hat{V}_t \equiv a_t \hat{w}_t + (1-a_t)\hat{r}_t + a_t \hat{L}_t + (1-a_t)\hat{J}_t \equiv \lambda_t + a_t \hat{L}_t + (1-a_t)\hat{J}_t \tag{2}$$

where ^ denotes a proportional growth rate, $a_t \equiv w_t L_t / V_t$ is the share of labor in output, $1-a_t \equiv r_t J_t / V_t$ is the share of capital and $\lambda_t \equiv a_t \hat{w}_t + (1-a_t)\hat{r}_t$. Expressions (1) and (2) are accounting identities. Suppose that in this economy factor shares are relatively stable. This could be due, for example, to the fact that firms set prices according to a mark-up on unit labor costs. Assume also that in this economy wage and profit rates grow at constant rates. This implies that expression (2) can be written as

$$\hat{V}_t \equiv \lambda + a\hat{L}_t + (1-a)\hat{J}_t \tag{3}$$

where $\lambda \equiv a\hat{w} + (1-a)\hat{r}$. If we now integrate (3) and take antilogarithms we obtain

$$V_t \equiv A_0 \exp(\lambda t) L_t^a J_t^{1-a} \tag{4}$$

Expression (4) is simply the income accounting identity, expression (1) rewritten under the two assumptions mentioned above. It is certainly not a Cobb-Douglas production function, as it does not exist.

The devastating implications of this simple derivation are elaborated upon in the works by Felipe and McCombie mentioned above.[29] For example, our arguments are very important to understand discussions about total factor productivity growth. The aggregation problem *matters* because "without proper aggregation we cannot interpret the properties of an aggregate production function, which rules the behavior of total factor productivity" (Nadiri 1970, p.1144). It is worth quoting Nadiri's views on the issue:

"The conclusion to be drawn from this brief discussion is that aggregation is a serious problem affecting the magnitude, the stability, and the dynamic changes of total factor productivity. We need to be cautious in interpreting the results that depend on the existence and specification of the aggregate production function...That the

---

[28] It must be stressed that equation (1) is not a behavioral equation and it is not derived from Euler's theorem. It is simply an accounting identity that holds always.

[29] One such implication is, for example, that most if not all, econometric problems discussed in the literature (e.g., presence of unit roots, simultaneity bias) have to be dismissed as irrelevant. All researchers do in most growth regressions is to approximate, one way or another, the accounting identity.

use of the aggregate production function gives reasonably good estimates of factor productivity is due mainly to the narrow range of movement of aggregate data rather than the solid foundation of the function. In fact, the aggregate production function does not have a conceptual reality of its own" (Nadiri 1970, p.1145-1146).

In agreement with this view, Felipe and McCombie (2005c) have recently shown using simulations that the true rate of technical progress, computed using firm-level data, is very different from that obtained using aggregate data. The results show that the two measures of productivity are so different from each other that it is concluded that total factor productivity growth calculated with aggregate data is, in no way, a proxy for the true rate of technological progress.

Let's see with three examples some the implications of these arguments. Felipe (2000) delved into the question of whether total factor productivity growth in Singapore was zero, and discussed what this could mean. For this, he calculated $\lambda_t \equiv a_t\hat{w}_t + (1-a_t)\hat{r}_t \equiv \hat{V}_t - a_t\hat{L}_t - (1-a_t)\hat{J}_t$ (see equation (2)); but unlike Young (1992, 1995), who drew his conclusions from $\lambda_t \equiv \hat{V}_t - a_t\hat{L}_t - (1-a_t)\hat{J}_t$, Felipe calculated and analyzed $\lambda_t \equiv a_t\hat{w}_t + (1-a_t)\hat{r}_t$ and confirmed that indeed $\lambda_t \equiv 0$. Following Young, Singapore's labor share had been around 0.5 for the period under consideration. These observations imply that $\hat{w}_t \cong -\hat{r}_t$. In other words, the reason why Singapore's TFP growth rate was zero was simply that the average profit rate had declined at a rate approximately equal to that of the increase in the average wage rate. This result follows solely from the accounting identity.

The second example questions Young's (1994) analysis. Young estimated a cross-country production function using 118 countries for the period 1970-85. Regressing the growth of output per worker on a constant and the growth of the capital-labor ratio $(\hat{J}_i - \hat{L}_i)$ gave the result:

$$(\hat{V}_i - \hat{L}_i) = -0.21 + 0.45(\hat{J}_i - \hat{L}_i) + \varepsilon_i \tag{5}$$

The residual ε_i measures the growth of country *i's* total factor productivity less the world average (no statistical diagnostics were reported). Young noted that the residuals for the East Asian countries were very close in value to his much more detailed analysis using the growth accounting methodology. Singapore, for example, had an annual growth rate of TFP of –0.4 per cent per annum from the growth accounting exercise, whereas the value obtained from the regression was 0.1 per cent per annum.

But what does regression (5) tell us? We know that from the accounting identity:

$$(\hat{V}_i - \hat{L}_i) \equiv a_i\hat{w}_i + (1-a_i)\hat{r}_i + (1-a_i)(\hat{J}_i - \hat{L}_i) \equiv$$

$$\equiv \text{TFPG}_i + (1-a_i)(\hat{J}_i - \hat{L}_i) \tag{6}$$

where the subscript i denotes the i-th country. Consequently, if we estimate $(\hat{V}_i - \hat{L}_i) = b_1 + b_2(\hat{J}_i - \hat{L}_i) + \varepsilon_i$, it is apparent that the estimate of b_2 will be the average value of the share of capital, and the sum of the constant (b_1) and ε_i will provide, by definition, an estimate of λ_i (the estimates may be subject to some bias if λ_i is not orthogonal to $(\hat{J}_i - \hat{L}_i)$. Thus, it is hardly surprising that the estimates from this "back-of-the envelope" calculation do not differ from the more detailed studies. The fact that the estimated slope coefficient is reasonably close to the average factor share has no implications for whether or not perfect competition is a reasonable first approximation for analyzing the growth rates of these countries.

Finally, in one of the most recent applications of standard growth accounting, Rodrik and Subramanian (2004) project India's future "potential output growth rate" through 2025 and conclude that it is close to 7% for aggregate output, or 5.5 percent for output per capita. Rodrik and Subramanian (2004) use equation (2) above as follows. They set the share of capital in output $(1 - a_t)$ at 0.35 during the whole forecast period. The growth rate of capital ($\hat{J}_t$) is assumed to be 8.3% (currently it is about 6%). The authors believe that the dependency ratio will decline and this will release resources (savings), so that the savings ratio will increase from 25% to around 39%. This will allow an equal increase in domestic investment. This translates into a growth rate of the capital stock of around 8.3%. The growth rate of employment ($\hat{L}_t$) was set at 1.9% (this is the projected growth rate of the working age population). Finally, the rate of total factor productivity growth is estimated to be of 2.5%. This is the same as in the last two decades. Plugging these figures into equation (2) gives a growth rate of 6.64%. The authors argue that this is a lower bound estimate and, even so, would be significantly greater than the per capita growth rate of 3.6% percent achieved in the 1980s and 1990s.

However, in light of the discussion above we know that the estimate of *TFPG* equals $a_t\hat{w}_t + (1 - a_t)\hat{r}_t$ from the national accounts. While Rodrik and Subramanian's arguments about the evolution of factor shares, growth rate of employment and the growth rate of the capital stock might be valid, the ones regarding the evolution of total factor productivity growth are, in our view, arguable. The authors claim that *TFPG* of 2.5% a year was achieved with relatively modest reforms and that, therefore, there is still unexploited potential, and argue that empirical work seems to suggest that India's level of TFP is between 1/3 and 40 percent of what it should be, thus creating the scope for productivity improvements simply based on catching up.

7. WHY DO ECONOMISTS CONTINUE USING AGGREGATE PRODUCTION FUNCTIONS?

As argued in the Introduction, many economists today are not aware of the aggregation results and the implications. They simply think of the aggregate production function as part of the basic toolkit that they have been taught as both undergraduate and graduate students.

Some other economists, on the other hand, are aware of the aggregation results and yet continue using aggregate production functions. The reason for doing it falls under three broad justifications. First, and following Samuelson (1961-62), aggregate production functions are seen as useful parables. Second, based on the methodological position known as instrumentalism, it is argued that as long as aggregate production functions appear to give empirically reasonable results, why shouldn't they be used? Neoclassical macro theory deals with macroeconomic aggregates derived by analogy with the micro concepts (Ferguson 1971). The usefulness of this approach is strictly an empirical issue. In other words, as aggregate production functions tend to work empirically (i.e., when estimated econometrically, the statistical fit tends to be relatively high, the estimated output elasticities are relatively close to the factor shares and the marginal productivity of labor is close to the wage), it must mean that the estimates indeed reflect the true technological parameters. Finally, it is argued that for the applications where aggregate production functions are used, there is no other choice. However, in the light of the aggregation results, none of these reasons seems valid.

The first argument sometimes given to justify the use of aggregate production functions is that the aggregate production function is to be thought of as a *parable*, following the arguments in Samuelson (1961-1962) work. Samuelson's arguments, however, were stated in the context of the so-called Cambridge Capital theory debates that we have briefly referred to in the Introduction.[30] Samuelson showed that even in cases with heterogeneous capital goods, some rationalization could be provided for the validity of the neoclassical parable, which assumes that there is a single homogenous factor referred to as capital, and whose marginal product equals the interest rate. Samuelson worked with a one-commodity model assuming a well-behaved, constant returns-to-scale production function referred to as the surrogate production function. His surrogate production function relies on the crucial assumption that the same

[30] It should be not be thought, however, that the issues we have discussed (i.e., aggregation problems) have no bearing on the Cambridge-Cambridge debates. The discovery that aggregate production functions can violate properties that one expects of production functions (so-called reswitching and reverse capital-deepening) was at bottom a discovery that the aggregate concept used is not a production function at all. The aggregation problem literature shows that this was to be expected.

proportion of inputs is used in the consumption-goods and capital-goods industries; that is, the machines required for different techniques on the surrogate production function are different with respect to engineering specifications, but, with each technique, the ratio of labor to machines required to produce its machines is the same as that required to produce homogeneous consumption goods. This means that the cost of capital is determined solely by labor embodied in the machines required for each technique and the time pattern of all techniques is the same.[31] Then, Samuelson showed that the relation between the wage rate and the profit rate would be the same as that obtained from an appropriately defined surrogate production function with surrogate capital as a single factor of production. In competitive equilibrium, the wage rate is determined by the marginal productivity of labor. The latter is a ratio of two physical quantities, independent of prices (i.e., independent of distribution). And the same for the rate of profit: it is determined by the marginal productivity of capital. It is also measured in physical quantities. Under these circumstances, since there is a well-behaved production function, there is a unique inverse relation between the intensity of the factors and the relative price, and thus, as a resource becomes more scarce, its price increases. Thus, Samuelson turned the real economy with heterogeneous goods into an imaginary economy with a homogeneous output.

However, in the light of the aggregation literature, Samuelson's parable loses its power. Furthermore, the results of the one-commodity model do not hold in heterogeneous commodity models, and Samuelson's results depend crucially on the assumption of equal proportions, as shown by Garegnani (1970). For the surrogate function to yield the correct total product, the "surrogate capital" would have to coincide with the value in terms of consumption of the capital in use. The surrogate production function cannot be generally defined.

A variation of the parable argument is that the aggregate production function should be understood as an *approximation*. It is evident that Fisher's (exact) aggregation conditions are so stringent that one can hardly believe that actual economies will satisfy them. Fisher (1969b), therefore, asked: What about the possibility of a *satisfactory approximation*? The motivation behind the question is very simple. In practice, what one cares about is whether aggregate production functions provide an adequate approximation to reality over the values of the variables that occur in practice. Thus suppose the values of capitals and labors in the economy lie in a bounded set. And suppose further that the requirement is that an aggregate production function exists within some specified distance of the true production function for all points

[31] Samuelson, apart from working with a model where there was only one consumption good, and where input coefficients were fixed at the micro level, also assumed constant returns to scale, perfect competition, that only the n-th capital good was used to produce the n-th capital good, and that depreciation of a capital good is independent of its age.

in the bounded set. Does this new restriction help the conditions for aggregation? One possible way to answer this question is by requiring that the exact conditions hold only *approximately* (e.g., for approximate capital aggregation it suffices that all technical differences among firms be approximately capital augmenting). Is this a useful solution? Fisher showed it is not. The reason is that in reality there will be differences that are not approximately capital augmenting. Therefore, the interesting question is whether there are cases where the exact aggregation conditions are not approximately satisfied but in which an aggregate production function gives a satisfactory approximation for all points in the bounded set. Fisher (1969b) proved that *the only way* for approximate aggregation to hold without approximate satisfaction of the Leontief conditions is for the derivatives of the functions involved to wiggle violently up and down, an unnatural property not exhibited by the aggregate production functions used in practice.[32]

The second argument is that despite the aggregation results, neoclassical macroeconomic theory deals with macroeconomic aggregates derived by analogy with the micro concepts. Then, the argument goes, why not continue using them? This position is the one espoused by Ferguson (1971) in his reply to Joan Robinson (1970): "Neoclassical theory deals with macroeconomic aggregates, usually by constructing the aggregate theory by analogy with the corresponding microeconomic concepts. Whether or not this is useful in an empirical question to which I believe an empirical answer can be given. This is the "faith" I have, but which is not shared by Mrs. Robinson. Perhaps it would be better to say that the aggregate analogies provide working hypotheses for econometricians" (Ferguson 1971, pp.251-252). This is also Solow's (1966) position quoted at the start of the paper. This argument, however, is based on pure instrumentalism, a methodological position today rejected by most philosophers of science in general, and by experts on the methodology of economics in particular (e.g., Blaug 1992).

Naturally, the aggregation problem appears in all areas of economics, including consumption theory, where a well-defined micro consumption theory exists. The neoclassical aggregate production function is also built by analogy. This is Ferguson's (1971) argument. The aggregation problem is therefore viewed as being merely a nihilistic position. However, in the light of the discussion in this paper, this argument is untenable. Employing macroeconomic production functions on the unverified premise that inference by analogy is correct appears to be inadmissible, and the concept of "representative firm" à la Marshall is, in general, inapplicable. Furthermore, the difference with the case of the consumption function is that the conditions for successful aggregation in this case, while strong, do

[32] This paper contains a minor error, latter corrected. The correction is incorporated in the version found in Fisher (1993).

not seem to be so outlandish as those in the case of the production function. The aggregate consumption function can be shown to exist so long as either individual marginal propensities to consume are constant and about equal; or so long as the distribution of income remains relatively fixed. These seem relatively plausible. See Green (1964, chapter 5).[33]

The third and final argument given for the use of aggregate production functions is that there is no other option if one is to answer the questions for which the aggregate production function is used, e.g., to discuss productivity differences across nations. This argument acquires shape with the remark that it is hoped that the results be more or less qualitatively correct, and that they provide "some guide to orders of magnitude" (Solow 1988, p.314). This reasoning is a by-product of the instrumentalist position and it also clashes with the results of the aggregation work. Of course, if one insists on a research program whose goal is, for example, to split overall growth at the country level into the alleged contribution of technical progress and factor accumulation (i.e., growth accounting, as done by Solow (1957) and more recently by Young 1995)), surely one needs an aggregate production function in order to allegedly relate aggregate output to aggregate inputs (and thus to speak of a country's multi-factor or total factor productivity). But if one realizes that that the whole meaning of aggregates such as investment, GDP, labor, and capital is questionable, as Fisher (1987) pointed out, the legitimacy of the research program collapses. And even at the conceptual level, the objective behind a growth accounting exercise for purposes of estimating total factor productivity growth is by no means universally shared (e.g., Kaldor 1957; Pasinetti 1959; Nelson 1973, 1981; Nelson and Winter 1982; Scott 1989).[34]

8. CONCLUSIONS

This paper has reviewed the literature on the aggregation problem in production functions and its implications for theoretical and applied work. This work was pioneered by Lawrence Klein, although it flourished years after his seminal contribution in the 1940s. The theoretical results on aggregation are not a mere intellectual curiosity. If one considers economics a field where research is guided by the scientific method,

---

[33] Interestingly, Solow indicated that "the aggregate production function is only a *little less* legitimate concept than, say, the aggregate consumption function..." (Solow 1957, p.349; italics added). Certainly we disagree. Fisher (1969, p.575) compares the two sets of conditions, for production and consumption functions, and concludes that the former are substantially more stringent.

[34] Fisher (1993) indicates that as far back as 1970 he had already called "into question the use of aggregate production functions in macroeconomic applications such as Solow's famous 1957 paper" (Fisher (1993, p.xiii)).

the disregard for the implications of the aggregation results should be an issue of serious concern.

The use of aggregate production functions in theoretical and applied work reminds us of the long process it took mankind to understand the motions of the planets, delayed for over two thousand years. It resulted in the perpetuation of the Ptolomaic geocentric model of the movement of the planets long after Aristarchus, circa 250 B.C., first questioned this system and put forward the alternative heliocentric theory. This was partly due to the fact that by the use of epicycle after epicycle to describe the motions of the planets (although the epicycle is a totally artificial construct), the Ptolemaic system produced very good predictions. But prediction is not the same as explanation.

Briefly, an examination of the conditions required for aggregation yields results such as:

- Even under constant returns, the conditions for aggregation are so very stringent as to make the existence of aggregate production functions in real economies a non-event. This is true not only for the existence of an aggregate capital stock but also for the existence of such constructs as aggregate labor or even aggregate output.
- One cannot escape the force of these results by arguing that aggregate production functions are only approximations. While, over some restricted range of the data, approximations may appear to fit, good approximations to the true underlying technical relations require close approximation to the stringent aggregation conditions, and this is not a sensible thing to suppose.

To these, we add the fundamental point that when a researcher works – as one must at an aggregate level – with quantities measured in value terms, the appearance of a well-behaved aggregate production function tells one nothing at all about whether there really is one. Such an appearance stems from the accounting identity that relates the value of outputs to the value of inputs – *nothing more*.

Yet the implications of the points we have listed are not merely theoretical. They include:

- The specification and estimation of the aggregate demand curve for labor;
- The measurement of productivity and, especially, the interpretation (or, perhaps more properly, the *mis*interpretation) of the Solow residual;
- The use of aggregate production functions to validate the neoclassical theory of distribution; and
- The testing and potential refutation of the neoclassical model.

To these we would add generally the interpretation of such concepts as "investment", "capital", "labor", or "gross domestic product" in public policy discussions or any other context that supposes them to be related as inputs or outputs in a true production relationship. We are talking here of the very foundations of neoclassical macroeconomics.[35]

BIBLIOGRAPHY

AGHION, P. AND HOWITT, P. (1998) *Endogenous Growth Theory*. Cambridge: MIT Press.

AGHION, P. AND HOWITT, P. (2005) Appropriate Growth Policy: A Unifying Framework. Manuscript. 2005 Joseph Schumpeter Lecture, delivered to the 20th Annual Congress of the European Economic Association, Amsterdam, August 25, 2005.

BACKUS, D.; KEHOE, P.J. AND KEHOE, T.J. (1992) In Search of scale Effects in Trade and Growth, *Journal of Economic Theory*, 58, 377-409.

BARRO, R. AND SALA-I-MARTIN, X. (1995) *Economic Growth*. McGraw-Hill International Editions.

BASU, S. AND FERNALD, J.G. (1995) Are Apparent Spillovers a Figment of Specification Error?, *Journal of Monetary Economics*, 36, 165-188.

________ (1997) Returns to Scale in U.S. Production: Estimates and Implications, *Journal of Political Economy*, 105, 249-283.

BLACKORBY, C. AND SCHWORM, W. (1984) The Structure of Economies with Aggregate Measures of Capital: A Complete Characterization. *Review of Economic Studies*, 51, 633-50.

________ (1988) The Existence of Input and Output Aggregates in Aggregate Production Functions. *Econometrica*, 56(3), 613-43.

BLAUG, M. (1992) *The Methodology of Economics*. Cambridge University Press.

BROWN, M. (1980) The Measurement of Capital Aggregates: A Post-reswitching Problem. In Dan Usher (ed.): *The Measurement of Capital*, 377-420. Chicago and London: The University of Chicago Press.

BURNSIDE, C. (1996) Production Function Regressions, Returns to Scale, and Externalities, *Journal of Monetary Economics*, 37, 177-201.

[35] Neoclassical *microeconomics* is not affected by these results.

CASS, D. (1965) Optimum Growth in an Aggregative Model of Capital Accumulation. *The Review of Economic Studies*, 32(3), 233-240.

CABALLERO, R.J. AND LYONS, R.K. (1992) External Effects in U.S. Procyclical Productivity, *Journal of Monetary Economics*, 29, 209-225.

COHEN, A.J. AND HARCOURT, G.C. (2003) Whatever Happened to the Cambridge Capital Theory Controversies? *Journal of Economic Perspectives*, 17(1), 199-214.

DENISON, E.F. (1961) Measurement of Labor Input: Some Questions of Definition and the Adequacy of Data. In *Conference on Research in Income and Wealth, Output, Input and Productivity Measurement, Studies in Income and Wealth*, 25, 347-72. Princeton: Princeton University Press.

________ (1972a) Some Major Issues in Productivity Analysis: An Examination of Estimates by Jorgenson and Griliches. *Survey of Current Business,* 52(II), 37-63.

________ (1972b) Final Comments. *Survey of Current Business,* 52(II), 95-110.

DURLAUF, S.N., JOHNSON, P.A. AND TEMPLE, J.R.W. (2005) Growth Econometrics. In P. Aghion and S. N. Durlauf (eds.) *Handbook of Economic Growth* North-Holland.

EASTERN ECONOMIC JOURNAL (2005) Symposium on the Aggregate Production Function, Vol.31, N° 3 (Summer).

FELIPE, J. (2000) On the Myth and Mystery of Singapore's Zero TFP. *Asian Economic Journal*, 14, 187-209.

FELIPE, J. (2001) Endogenous Growth, Increasing Returns, and Externalities: An Alternative Interpretation of the Evidence. *Metroeconomica*, 52, 4, 391-427.

FELIPE, J. AND ADAMS, F.G. (2005) 'A THEORY OF PRODUCTION'. The Estimation of the Cobb-Douglas Function: A Retrospective View. *Eastern Economic Journal*, 31, 3 (Summer):

FELIPE, J. AND FISHER, F.M. (2003) Aggregation in Production Functions: What Applied Economists Should Know. *Metroeconomica*, 54(2-3), 208-62.

FELIPE, J. AND McCOMBIE, J.S.L. (2001) The CES Production Function, the Accounting Identity, and Occam's Razor. *Applied Economics*, 33, 10, 1221-32.

FELIPE, J. AND McCOMBIE, J.S.L. (2002) A Problem with Some Recent Estimations and Interpretations of the Mark-up in Manufacturing Industry. *International Review of Applied Economics*, 16, 2, 187-215.

FELIPE, J. AND McCOMBIE, J.S.L. (2003) Some Methodological Problems with the Neoclassical Analysis of the East Asian Miracle. *Cambridge Journal of Economics*, 54(5), 695-721.

FELIPE, J. AND McCOMBIE, J.S.L. (2005a) Why are Some Countries Richer than Others? A Skeptical View of Mankiw-Romer-Weil's Test of the Neoclassical Growth Model. *Metroeconomica*, 56, 3, 360-392

FELIPE, J. AND McCOMBIE, J.S.L. (2005b) Is a Theory of Total factor Productivity Really Needed? *Metroeconomica*. Forthcoming.

FELIPE, J. AND McCOMBIE, J.S.L. (2005c) The Tyranny of the Identity: Growth Accounting Revisited. *International Review of Applied Economics*. Forthcoming.

FERGUSON, C.E. (1971) Capital Theory up to Date: A Comment on Mrs. Robinson's Article. *Canadian Journal of Economics*, 4,2, 250-254.

FISHER, F.M. (1965) Embodied Technical Change and the Existence of an Aggregate Capital Stock. *Review of Economic Studies*, 32, 263-88.

FISHER, F.M. (1968) Embodied Technology and the Existence of Labor and Output Aggregates. *Review of Economic Studies*, 35, 391-412.

________ (1969a) The Existence of Aggregate Production Functions. *Econometrica*, 37(4), 553-77.

________ (1969b) Approximate Aggregation and the Leontief Conditions. *Econometrica*, 37(4), 457-69.

________ (1971) Aggregate Production Functions and the Explanation of Wages: A Simulation Experiment. *The Review of Economics and Statistics,* 53(4), 305-25.

________ (1982) Aggregate Production Functions Revisited: The Mobility of Capital and the Rigidity of Thought. *Review of Economic Studies*, 49(4), 615-26.

________ (1983) On the Simultaneous Existence of Full and Partial Capital Aggregates. *Review of Economic Studies*, 50(1), 197-208.

________ (1987) Aggregation Problem. In Eatwell, Millgate and Newman (eds.): *The New Palgrave. A Dictionary of Economics*, 53-5. London and Basingstoke: Macmillan.

________ (1993) *Aggregation. Aggregate Production Functions and Related Topics*. Cambridge, MA: The MIT Press.

FISHER, F.M., SOLOW, R. AND KEARL, J.M. (1977) Aggregate Production Functions: Some CES Experiments. *Review of Economic Studies*, 44(2), 305-20.

GAREGNANI, P. (1970) Heterogeneous Capital, the Production Function and the Theory of Distribution. *Review of Economic Studies*, 37, pp.407-436.

GORMAN, W.M. (1959) Separable Utility and Aggregation. *Econometrica*, 27, 469-81.

________ (1968) Measuring the Quantities of Fixed Factors. In J.N. Wolfe (ed.): *Value, Capital, and Growth. Papers in Honour of Sir John Hicks*, 141-72, Edinburgh University Press.

GREEN, J. (1964) *Aggregation in Economic Analysis. An Introductory Survey*. Princeton: Princeton University Press.

HALL, R.E. (1990) "Invariance Properties of Solow's Productivity Residual, In Peter Diamond (ed.), *Growth/Productivity/Employment*, The MIT Press, Cambridge.

HOUTHAKKER, H.S. (1955-56) The Pareto Distribution and the Cobb-Douglas Production Function in Activity Analysis. *Review of Economic Studies*, 23, 27-31.

HOWITT, P. (2000) Endogenous Growth and Cross-Country Income Differences. *The American Economic Review*, 90(4), 829-46.

HOWITT, P. (2002) Review of Underdevelopment: A Strategy for Reform by Paolo Sylos-Labini. *Journal of Economic Literature*, 41, 617-18.

JORGENSON, D. (1972) Issues in Growth Accounting: A Reply to Edward F. Denison. *Survey of Current Business*, 52(II), 65-94.

JORGENSON, D. AND GRILICHES, Z. (1967) The Explanation of Productivity Change. *Review of Economic Studies*, 34(3), 249-83.

KALDOR, N. (1957) A Model of Economic Growth. *The Economic Journal*, Vol.67, No.268 (December), pp.591-624.

KLEIN, L. (1946a) Macroeconomics and the Theory of Rational Behavior, *Econometrica*, 14(2), 93-108.

________ (1946b) Remarks on the Theory of Aggregation, *Econometrica*, 14(4), 303-12.

LEONTIEF, W.W. (1947a) Introduction to a Theory of the Internal Structure of Functional Relationships. *Econometrica*, 15(4), 361-73.

________ (1947b) A Note on the Interrelationship of Subsets of Independent Variables of a Continuous Function with Continuous First Derivatives. *Bulletin of the American Mathematical Society*, 53, 343-50.

LEVHARI, D. (1968) A Note on Houthakker's Aggregate Production Function in a Multifirm Industry. *Econometrica*, 36(1), 151-54.

MANKIW, G., ROMER D. AND WEIL, D. (1992) A Contribution to the Empirics of Economic Growth. *Quarterly Journal of Economics*, 107(2), 407-37.

MAY, K. (1946) The Aggregation Problem for a One-Industry Model. *Econometrica*, 14(4), 285-98.

________ (1947) Technological Change and Aggregation. *Econometrica*, 15(1), 51-63.

MCCOMBIE, J.S.L. AND THIRLWALL, A.P. (1994). *Economic Growth and the Balance of Payments Constraint*. St. Martin's Press.

NADIRI, I. (1970) Some Approaches to the Theory and Measurement of Total Factor Productivity: A Survey. *Journal of Economic Literature*, 8, 1137-77.

NATAF, A. (1948) Sur la Possibilité de Construction de Certains Macromodèles. *Econometrica*, 16(3), 232-44.

NELSON, R. (1964) Aggregate Production Functions and Medium Range Growth Projections. *The American Economic Review*, 54(5), 575-606.

NELSON, R. (1973) Recent Exercises in Growth Accounting: New Understanding or Dead End?. *American Economic Review*, 73(2), 462-68.

________ (1981) Research on Productivity Growth and Productivity Differences: Dead Ends and New Departures. *Journal of Economic Literature*, 19, 1029-64.

NELSON, R. AND WINTER, S.G. (1982) *An Evolutionary Theory of Economic Change*. Cambridge, MA: Harvard University Press.

PASINETI, L. (1959) On Concepts and Measures of Changes in Productivity. *The Review of Economics and Statistics*, 41, 270-86.

PU, S.S. (1946) A Note on Macroeconomics. *Econometrica,* 14(4), 299-302.

RODRIK, D. (2005) *Why We Learn Nothing from Regressing Economic Growth on Policies.* Harvard University. Manuscript (March).

RODRIK, D. AND SUBRAMANIAN, A. (2004) Why India can grow at 7 percent a year or more: Projections and Reflections. *IMF Working Paper 04/118.*.

ROMER, P. (1987) Crazy Explanations for the Productivity Slowdown, *NBER Macroeconomics Annual 1987*, pp. 163-202. The MIT Press, Cambridge, MA.

SAMUELSON, P. (1961-62) Parable and Realism in Capital Theory: The Surrogate Production Function. *Review of Economic Studies*, 29, 193-206.

________ (1966) A Summing Up. *Quarterly Journal of Economics*, 80(4), 568-83.

________ (1979) Paul Douglas's Measurement of Production Functions and Marginal Productivities. *Journal of Political Economy*, 87, 923-39.

SATO, K. (1975) *Production Functions and Aggregation*. North-Holland.

SHAIKH, A. (1974) Laws of Production and Laws of Algebra: The Humbug Production Function. *Review of Economics and Statistics*, LVI, 115-120.

SHAIKH, A. (1980) Laws of Production and Laws of Algebra: Humbug II, in Edward J. Nell (ed.), *Growth, Profits and Property, Essays in the Revival of Political Economy.* pp. 80-95, Cambridge: Cambridge University Press.

SIMON, H. (1979) On Parsimonious Explanations of Production Relations. *Scandinavian Journal of Economics*, 81, 459-74.

SIMON, H. AND LEVY, F.K. (1963) A Note on the Cobb-Douglas Function. *Review of Economic Studies,* 93-4.

SOLOW, R. (1955-56) The Production Function and the Theory of Capital. *Review of Economic Studies*, 23(2), 101-8.

SOLOW, R. (1957) Technical Change and the Aggregate Production Function. *Review of Economics and Statistics*, 39 3, 312-20.

SOLOW, R. (1966) Review of Capital and Growth. *American Economic Review*, 56(5), 1257-60.

SOLOW, R. (1988) Growth Theory and After. *The American Economic Review*, 78, 3, 307-17.

SYLOS-LABINI, P. (2001) *Underdevelopment: A Strategy for Reform*. Cambridge, New York and Melbourne: Cambridge University Press.

TEMPLE, J. (1999) The New Growth Evidence. *Journal of Economic Literature*, 37, 112-56.

TEMPLE, J. (2003) The Long-Run Implications of Growth Theories. *Journal of Economic Surveys*, 17(3), 497-510.

YOUNG, A. (1994) Lessons from the NICs: a Contrarian View. *European Economic Review*, 38, 964-73.

YOUNG, A. (1992) A Tale of Two Cities: Factor Accumulation and Technical Change in Hong Kong Singapore. *National Bureau of Economic Research Macroeconomics Annual 1992*, 13-63.

YOUNG, A. (1995) The Tyranny of the Numbers: Confronting the Statistical Realities of the East Asian Growth Experience. *Quarterly Journal of Economics*, 110, 641-80.