

HOW SOUND ARE THE FOUNDATIONS OF THE AGGREGATE PRODUCTION FUNCTION?

Jesus Felipe

Asian Development Bank

and

J. S. L. McCombie

University of Cambridge

INTRODUCTION

It is somewhat paradoxical that one of the concepts most widely used in macroeconomics, namely the aggregate production function, is the one whose theoretical rationale is perhaps the most suspect. We define an aggregate production function as one for which output has to be a value—rather than a physical—measure, regardless of the precise unit of observation, for example, whether it is for the firm or an individual industry. The value measure (such as value added or gross output) has to be used because of the heterogeneity of output. The same is true for the capital stock. These constant price measures are not physical quantities. Hence, the expression “volume of output,” which is commonly used when referring to aggregate output, is a misnomer.¹

The serious problems raised by the Cambridge Capital Theory Controversies dominated the debate in “high theory” in the late 1960s and early 1970s, and eventually led to an agreement that reswitching and capital reversing were theoretically possible [Harcourt, 1972].² This posed serious problems for the justification of the use of the neoclassical one-sector aggregate production function as a “parable,” to use the term coined by Solow [1966]. The revival of interest in growth theory with the development of endogenous growth theory is still squarely in the neoclassical tradition [Pasinetti, 1994].

If this were not enough, a whole series of further problems surrounds the question of whether or not micro-production functions can be aggregated to give a macroeconomic relationship that reflects the underlying technology of the economy in some meaningful way [Walters, 1963; Fisher, 1987; 1992; 2005; Felipe and Fisher, 2003]. A large technical literature has developed over the years on the aggregation problem. Essentially, though, it boils down to the question: does it make sense to sum the inputs and outputs, of, say, oil refining and the textile industry and estimate a function that somehow purports to represent the technological parameters, such as the aggregate elasticity of substitution, of this combined industry? When the aggregate production function for, say, the whole economy or manufacturing is estimated, this,

Jesus Felipe: Asian Development Bank, P. O. Box 789, 0890 Manila, Philippines. E-mail: jfelipe@adb.org.
J. S. L. McCombie: Cambridge Centre for Economic and Public Policy, Department of Land Economy, University of Cambridge, UK. E-mail: js1m2@cam.ac.uk.

of course, consists of an even more disparate collection of industries. If the answer to the question is that it really is nonsensical, then the further question arises as to why the aggregate production function is still so widely used in both theoretical and empirical modeling.

The defense of estimating aggregate production functions has been eloquently put forward by Solow [1966, 1259-60], who can hardly be accused of not being fully aware of the aggregation problem:³ "I have never thought of the macroeconomic production function as a rigorously 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." Wan [1971, 97] reflects this view when he argues that Solow's [1957] influential approach in measuring technical progress may be defended on the grounds that "one may argue that the functional relation between Q and K , L is an 'empirical law' in its own right. In the methodological parlance of Samuelson, this is an operationally meaningful law, since it can be empirically refuted". (Q , K , and L are output, capital, and labor.) From the first studies by Douglas and his various collaborators [Cobb and Douglas, 1928; Douglas, 1948; 1976], it has often been found that the aggregate production function gives a close statistical fit, especially using cross-regional or cross-industry data, with the estimated output elasticities close to the factor shares.^{4,5}

What we shall show, however, is that these studies generally could *not* have failed to find a close correspondence between the output elasticities and the factor shares. This arises from the fact that ideally the production function is a *microeconomic* concept, specifying the relationship between *physical* outputs and inputs (such as numbers of widgets, persons employed, and identical machines). But at the aggregate level, constant price value data are used for capital and output. Yet this is not an innocuous procedure, as it undermines the possibility of empirically testing the aggregate production function. As Simon [1979a, 497] pointed out in his Nobel Memorial lecture, the good fits to the Cobb-Douglas production function "cannot be taken as strong evidence for the classical theory, for the identical results can readily be produced by mistakenly fitting a Cobb-Douglas function to data that were in fact generated by a linear accounting identity (value of goods equals labor cost plus capital cost)."

We shall show that specifications of all aggregate production functions, using value data, are nothing more than approximations to an accounting identity, and hence can convey no information, *per se*, about the underlying technology of the "representative firm."⁶ Moreover, we shall also show that even the use of physical data is not without its problems. Once it is recognized that all that is being estimated is an underlying identity, it can be shown how it is always possible, with a little ingenuity, to obtain a perfect statistical fit to a putative production function, which exhibits constant returns to scale and where the estimated "output elasticities" equal the factor shares. It can also be shown how the results of the estimation of production functions that find increasing returns to scale and externalities are simply due to misspecification of the underlying identity and the estimated biased coefficients may actually be predicted in advance [McCombie, 2000-2001; Felipe, 2001a].

The purpose of this paper is to survey and extend this critique and also to consider some counter-criticisms that have been made. We conclude that the latter leave the central tenet unaffected and thus the concept of the aggregate production function lacks any firm foundations. We do not address here the difficult, but necessary, question of what alternative approaches there are that avoid these problems.

We first begin by considering the accounting identity and, for expositional ease, the Cobb-Douglas production function (although, as we have noted, the critique applies to any production function). The next section shows that the problem stems essentially from the fact that monetary values at constant prices have to be used empirically as a proxy for physical quantities. Nevertheless, as we have noted, the use of data for physical quantities also presents difficulties, which we also discuss. If this critique is so devastating to the notion of the aggregate production function, then what have been the reactions to it and why has it been largely ignored in the literature? We attempt to answer this question in the penultimate section. Finally, we draw some conclusions.

THE ACCOUNTING IDENTITY AND THE COBB-DOUGLAS PRODUCTION FUNCTION

The problem that the accounting identity poses for the interpretation of the aggregate production function may be most easily seen by a consideration of the Cobb-Douglas production function, although it applies to more complex specifications (such as the CES and the translog).

It was first brought to the fore by Phelps Brown [1957] in his seminal paper "The Meaning of the Fitted Cobb-Douglas Function." It is one of the ironies of the history of economic thought that this article, which challenged the whole rationale for estimating aggregate production functions, was published in the same year as Solow's [1957] "Technical Change and the Aggregate Production Function." The latter, of course, was largely responsible for initiating the neoclassical approach to the empirical analysis of economic growth.

Phelps Brown's [1957] critique was largely addressed to the fitting of production functions using cross-sectional data and was specifically directed at Douglas' various studies (see Douglas [1948] for a summary). It is useful to consider this first, as it avoids the complication of technical change. We consider the latter next when we look at time-series data.

Cross-Section Data

The fact that the crucial tenet of Phelps Brown's [1957] argument was presented rather obscurely and was buried in his paper did not help its reception, even though it was published in one of the leading U.S. economics journals. He noted that there always exists an underlying accounting identity for the i^{th} firm given by:

$$(1) \quad V_i \equiv wL_i + rJ_i,$$

where V , L , J , w , and r are output (value added at constant prices), the labor input, the constant price value of the capital stock, the wage rate, and the observed rate of

return.⁷ The last two are the averages across the industries and are assumed to show little variation and thus may be considered approximately uniform. Because of the existence of the identity, Phelps Brown [1957, 557] concluded that the “Cobb-Douglas [α] and the share of earnings in income will be only two sides of the same penny,” where α is the output elasticity of labor.

The argument therefore seems to be this. The elasticity of output with respect to labor of the constant-returns-to-scale Cobb-Douglas production function, $V_i = AL_i^\alpha J_i^{(1-\alpha)}$ is defined as $\alpha = (\partial V_i / \partial L_i)(L_i / V_i)$. Given the assumptions of the neoclassical theory of factor pricing, from the first order conditions, the marginal product of labor equals the wage rate, that is, $\partial V_i / \partial L_i = w$. Consequently, it also follows that $\alpha = a = wL_i / V_i$ the share of labor in value added. The conventional approach sees this as a testable prediction of the neoclassical theory of factor pricing.

An accounting identity given by Equation (1), however, also defines the measure of value added for all units of observation, whether they are at the level of, say, the firm or the whole economy. It should be emphasized that there are no behavioral assumptions underlying this equation, in that it is compatible with any degree of competition, increasing, constant, or decreasing returns to scale and whether or not a well-behaved underlying aggregate production function actually exists. The observed rate of profit, r , is usually measured as $r = (1 - a)V_i / J_i = \Pi_i / J_i$ where $(1 - a)$ is the capital's share in value added and Π is total profits.

Consequently, partially differentiating the accounting identity, $V_i = wL_i + rJ_i$ with respect to L_i gives $\partial V_i / \partial L_i = w$ and it follows that $(\partial V_i / \partial L_i) / (L_i / V_i) = a_i = wL_i / V_i$. This is identical to the result obtained from the aggregate production function and the first order conditions, as shown above. The problem is that as it is derived from an identity, this result is always true, that is, it is impossible to reject it by statistical testing. As the argument stemming from the identity has not made any economic assumptions at all, the finding that the putative output elasticities equal the observed factor shares cannot be taken as a test of whether or not factors of production are paid their marginal products. This is a position, however, that was not accepted by Douglas [1976, 914] himself. “A considerable body of independent work tends to corroborate the original Cobb-Douglas formula, but, more important, the approximate coincidence of the estimated coefficients with the actual shares received also strengthens the competitive theory of distribution and disproves the Marxian.” It is noticeable that in his survey, however, Douglas fails to mention the Phelps Brown [1957] paper.

If the (aggregate) output elasticity of labor and the share of labor's total compensation are merely “two sides of the same penny,” could it be that the Cobb-Douglas is simply an alternative way of expressing the income identity and, as such, has no implications for the underlying technology of the economy? This was the proposition that Simon and Levy [1963] proved some six years later.

Following Simon and Levy [1963] and Intriligator [1978], the isomorphism between the Cobb-Douglas production function and the underlying accounting may be simply shown. The Cobb-Douglas, when estimated using cross-section (firm, industry, or regional) data, is specified either as:⁸

$$(2) \quad V_i = AL_i^\alpha J_i^\beta,$$

where A is a constant, or dividing by $\bar{V} = A\bar{L}^\alpha\bar{J}^{1-\beta}$ and taking logarithms as:

$$(3) \quad \ln(V_i/\bar{V}) = \alpha \ln(L_i/\bar{L}) + \beta \ln(J_i/\bar{J}),$$

where the bar over a variable denotes that it is the value of some reference observation, such as that of the average firm or the base year.

Using the approximation $\ln(X/\bar{X}) \approx (X/\bar{X}) - 1$, Equation (3) may be expressed as

$$(4) \quad V_i = (\alpha \bar{V}/\bar{L})L_i + (\beta \bar{V}/\bar{J})J_i + (1 - \alpha - \beta)\bar{V}.$$

A comparison of Equation (4), which is a linearization of the Cobb-Douglas relationship (viz. Equation (2)), with the income identity Equation (1), namely, $V_i \equiv wL_i + rJ_i$ shows that $w = \alpha \bar{V}/\bar{L}$ and $r = \beta \bar{V}/\bar{J}$. Moreover, it can be further seen from the comparison of these equations that $(1 - \alpha - \beta)\bar{V}$ must equal 0, which implies that $\alpha + \beta = 1$. Consequently, the data will always suggest the existence of “constant returns to scale,” whatever the true technological relationships of the individual production processes.

It is now possible to show conversely that the accounting identity can be transformed into a form that resembles a Cobb-Douglas function. Start from the identity that $V_i = wL_i + rJ_i$. Substitute for w from the definition $w = a\bar{V}/\bar{L}$, where, it will be recalled, a is labor’s share in total income. Likewise substitute for r using $r = (1 - a)\bar{V}/\bar{J}$, where $(1 - a)$ is capital’s share. This yields $V_i/\bar{V} = aL_i/\bar{L} + (1 - a)J_i/\bar{J}$. Now substitute using the approximation that $X/\bar{X} \approx \ln X/\bar{X} + 1$, which gives $\ln V_i/\bar{V} = a \ln L_i/\bar{L} + (1 - a) \ln J_i/\bar{J}$. This may be expressed as:

$$(5) \quad \ln V_i = \ln A + a \ln L_i + (1 - a) \ln J_i - [\ln \bar{V} - \ln A - a \ln \bar{L} - (1 - a) \ln \bar{J}],$$

where, from Equations (2) and (3), the term in square brackets equals zero. Taking antilogarithms gives $V_i = AL_i^a J_i^{(1-a)}$. The argument is that for reasonably small variations of L and J and with w and r constant (the last two conditions are not essential, as we shall see below), a Cobb-Douglas multiplicative power function will give a very good approximation to a linear function. The corollary is that the linear accounting identity will likewise give a good approximation to the Cobb-Douglas relationship. As the linear income identity exists for *any* underlying technology, we cannot be sure that all that the estimates are picking up is not simply the identity. The fact that a good fit to the Cobb-Douglas relationship is found implies nothing, *per se*, about such technological parameters as the elasticity of substitution. Simon [1979b, 466] found that mistakenly fitting a Cobb-Douglas relationship to the accounting identity for a range of capital-labor ratios that greatly exceed those found empirically gave “estimating errors of less than 5 per cent”. (See also Felipe and Holz, [2001].)

The good approximation of the Cobb-Douglas to the accounting identity is also likely to carry through even when we allow w and r to change, provided now the factor

shares do not show very much variation. To see this, assume a continuum of firms and differentiate the accounting identity to give:

$$(6) \quad dV_i = (dw_i)L_i + w_i dL_i + (dr_i)J_i + r_i dJ_i \quad \text{or}$$

$$(7) \quad dV_i/V_i = a_i(dw_i/w_i) + a_i(dL_i/L_i) + (1 - a_i)(dr_i/r_i) + (1 - a_i)(dJ_i/J_i).$$

Let us assume factor shares are constant (and there are many reasons why this should occur other than because there is a Cobb-Douglas production function—for example, firms pursue a constant markup pricing policy or the Kaldorian macroeconomic theory of distribution holds). Equation (7) may be integrated to give:

$$(8) \quad V_i = Bw_i^a r_i^{(1-a)} L_i^a J_i^{(1-a)},$$

where B is the antilogarithm of the constant of integration.

Provided that $w_i^a r_i^{(1-a)}$ either shows very little variation or is orthogonal to the expression $L_i^a J_i^{(1-a)}$, or both, the putative Cobb-Douglas production function will once again give a very good fit to the data. For regions or industries in any one country, w_i and r_i are likely to show relatively small variations compared with those of V_i , L_i , and J_i . It is not surprising, then, that the numerous regressions run by Paul Douglas and his colleagues using this type of data gave exceptionally good statistical fits, with the estimated output elasticities almost exactly equal to the factor shares. (See the summary tables in Douglas [1948; 1976].) Using cross-country data, however, such as that from the Penn World Table that includes both developing and less developing countries, should not give such a good fit because of the international variability in, especially, w . This indeed proves to be the case.⁹ In other words, in this case the Cobb-Douglas is a poor approximation to the linear accounting identity.

It should be noted that this critique of the aggregate production function is not just confined to the Cobb-Douglas production function. Simon [1979b] explicitly considers the CES production function given by $V_i = \psi(\theta L_i^{-\eta} + (1 - \theta)J_i^{-\eta})^{-1/\eta}$, where ψ , η , and θ are parameters. He argues that if the true relationship was given by the accounting identity and we were mistakenly to estimate the CES production function, then if η goes to zero, the function becomes a Cobb-Douglas. He cites Jorgenson [1974] as suggesting that most estimates give η close to zero and so the argument still applies. However, some more recent studies find that the putative aggregate elasticity of substitution is less than unity. But the argument is more general than Simon implies [Felipe and McCombie, 2001].

If we were to express any production function of the form $V_i = f(L_i, J_i)$ (such as the translog) in proportionate rates of change and use the marginal productivity conditions, we would find that $dV_i/V_i = c + a_i dL_i/L_i + (1 - a_i) dJ_i/J_i$, where c is a constant. This is formally exactly equivalent to the accounting identity expressed in proportionate rates of change, provided $a_i dw_i/w_i + (1 - a_i) dr_i/r_i$ is again constant or it will give a good approximation provided that $a_i dw_i/w_i + (1 - a_i) dr_i/r_i$ is orthogonal to $a_i dL_i/L_i + (1 - a_i) dJ_i/J_i$. This may be seen from Equation (7) from which it also follows that $\alpha_i = a_i$ and $(1 - \alpha_i) = (1 - a_i)$. If shares do vary, then we may be able to find an explicit functional form that is more flexible than the Cobb-Douglas that gives a good fit to the

accounting identity; but, of course, this does not mean that the estimated coefficients can now be interpreted as technological parameters.¹⁰ If $a_t dw/w_t + (1 - a_t) dr/r$ does not meet the assumptions noted above, all this means is that the estimated functional form will be misspecified and the goodness of fit will be reduced [McCombie, 2000].

Time-Series Data

The fact that the identity precludes interpreting the cross-section Cobb-Douglas relationship or more flexible functional forms as unambiguously reflecting the underlying technology of the economy implicitly suggests that this is true of estimations using time-series data.¹¹ The arguments, as Shaikh [1974; 1980] has shown, follow through in the case of time-series data. Differentiating the income identity, $V_t \equiv w_t L_t + r_t J_t$, with respect to time, we obtain:

$$(9) \quad \hat{V}_t \equiv a_t \hat{w}_t + (1 - a_t) \hat{r}_t + a_t \hat{L}_t + (1 - a_t) \hat{J}_t,$$

where \hat{V} , \hat{L} , \hat{J} , \hat{w} , and \hat{r} denote exponential growth rates. $a_t \equiv w_t L_t / V_t$ is labor's share in output and $(1 - a_t) \equiv r_t J_t / V_t$ is capital's share. Assuming that factor shares are constant and integrating Equation (9), we obtain:

$$(10) \quad V_t \equiv B w_t^a r_t^{(1-a)} L_t^a J_t^{(1-a)}.$$

Let us assume that the growth of the wage rate occurs at a roughly constant rate ($\hat{w}_t = \hat{w}$) and the rate of profit shows no secular growth ($\hat{r} = 0$), both of which may be regarded as stylized facts. Consequently, $a \hat{w}_t + (1 - a) \hat{r}_t \approx a \hat{w} = \lambda$, a constant, and so Equation (10) becomes the familiar Cobb-Douglas with exogenous technical change, namely, $V_t = A_0 e^{\lambda t} L_t^a J_t^{(1-a)}$.

In fact, while the cross-region or cross-industry studies normally give a very good fit to the Cobb-Douglas (and other) production functions, the time-series estimations sometimes produce implausible estimates with, for example, the coefficient of capital being negative. Table 1 in Sylos Labini [1995, 490] provides a useful summary of some time-series studies that give poor statistical fits. The fact that the results are often so poor may ironically give the impression that the estimated equation is actually a behavioral equation. The failure to obtain plausible estimates will occur, however, if either the factor shares are not sufficiently constant or the approximation $a_t \hat{w}_t + (1 - a_t) \hat{r}_t \approx \lambda$ (a constant) is not sufficiently accurate. In practice, it is the latter that proves to be the case, as estimations of Equation (10) with a variety of data sets produce well-determined estimates of the coefficients, which equal the relevant factor shares. It is found that the rate of profit, especially, has a pronounced cyclical component and so proxying $a \ln w + (1 - a) \ln r$ by a linear time trend (or its growth rate by a constant) biases the estimated coefficients of $\ln L$ and $\ln J$ (or of \hat{L} and \hat{J}).¹² See McCombie [2000-2001], Felipe and Holz [2001], and Felipe and McCombie [2001].

The conventional neoclassical approach, which is based on the maintained hypothesis that an aggregate production function is, in fact, being estimated, usually attributes a poor fit to the failure to adjust the growth of factor inputs for the changes in capacity

utilization. As $a_t \hat{w}_t + (1 - a_t) \hat{r}_t \equiv \lambda_t$ tends to vary procyclically, the inclusion of a capacity utilization variable will tend to improve the goodness of fit and cause the estimated coefficients to approximate more closely the relevant factor shares. While \hat{r} has a pronounced procyclical fluctuation, the growth of the capital stock generally shows little variation. Adjusting \hat{J} for changes in capacity utilization increases its cyclical fluctuation and at the same time reduces that in \hat{r} .¹³ The weighted sum of the logarithms of the wage and profit rates is now more closely proxied by a linear time trend and the regression estimates more closely reflect the identity. As Lucas [1970, 24], commented: "...some investigators have obtained 'improved' empirical production functions (that is, have obtained labor elasticities closer to labor's share) by 'correcting' measured capital stock for variations in utilization rates". (The same argument holds when the labor input is adjusted for changes in its intensity of use over the cycle.)

An alternative procedure would be to introduce a sufficiently complex nonlinear time trend more accurately to capture the variation of λ_t [Shaikh, 1980; Felipe and McCombie, 2003]. With sufficient ingenuity, we should be able eventually to approximate closely the underlying identity, increasing both the R^2 and the values of the t -statistics, and hence find a very good fit for the "production function". Generally, as we have noted above, it is this problem, rather than the change in factor shares, that is of greater empirical importance. As in the case of cross-sectional data, the problem posed by the identity occurs even though the factor shares vary.

THE USE OF QUANTITIES AND VALUE DATA IN ESTIMATING "PRODUCTION FUNCTIONS"

The result that estimates of the supposed output elasticities often approximate the observed relevant factor shares stems from the use of value data as a proxy for "quantity" or "volume" measures. But, as we have noted above, the use of physical data, even if they were available, is not free from problems. We begin by discussing this, before considering the additional problems posed by the use of value data.

The neoclassical model implicitly assumes a one-commodity world in which aggregate output is defined as Q , which is, say, the number of widgets. K is a physical measure of the capital stock, say the number of "leets." This also has an associated accounting identity measured in value terms, namely $pQ \equiv wL + \rho K$, where p is the price (again in, say, dollars per widget). w and ρ are the monetary values of wages and the return to capital. As K is measured in numbers of identical physical units, ρ is the price per machine (in dollars), rather than being a percentage rate of return. It will be readily apparent, however, that it is possible to recover the physical quantities from this identity. In other words, $Q = (w/p)L + (\rho/p)K$, where w/p and ρ/p are the returns to labor and capital measured in terms of widgets per unit of the relevant input (that is, physical magnitudes). Clearly, this is not possible in the case of value added. A possible confusion may arise because the latter may be written as $PV = P \sum p_i Q_i = w^n L + r^n J$, where the superscript n denotes a nominal value. P is not a price but a price deflator and, for example, $w^n/P = w$ is the real wage measured in monetary (and not physical) units.

Under certain circumstances, it is possible to estimate a production function and to test the marginal productivity theory of distribution. To see this, consider the neo-classical approach that uses a micro-production function specified in *physical* terms where there is no technical change, namely $Q = f(L, K)$. (Alternatively, the data could be of a cross-sectional nature—either cross-regional or cross-industry—where there is the same technology.)

Partially differentiating the micro-production function with respect to L and K gives $\partial Q/\partial L = f_L$ and $\partial Q/\partial K = f_K$, where f_L is the marginal product of labor and f_K is the marginal product of capital, both measured in physical units. By Euler's theorem, which, of course, has no economic content, *per se*, we have, assuming constant returns to scale, the following equation: $Q = f_L L + f_K K$ or $pQ = pf_L L + pf_K K$, where p , it will be recalled, is the price per unit output, measured in, say, dollars.¹⁴

Let us assume that the marginal products are roughly constant. Regressing Q on L and K will provide an estimate of f_L and f_K . It does not make any difference if we use the monetary values; we will still obtain the marginal products, although now measured in monetary units. Of course, there is no reason why factors should be paid their marginal products and so a comparison of pf_L and pf_K with the actual factor payments would constitute a test of the marginal productivity theory of factor pricing. This is an important point because, if p is constant, the coefficients of the estimated linear equation $pQ = b_1 L + b_2 K$ are determined by the underlying production function and they will differ from the observed wages and rate of profit if factors are not paid their marginal products. This is because Q is a physical measure and is independent of the distribution of the product. For example, in the case of the Cobb-Douglas production function, we have $f_L = \alpha Q/L$ and $f_K = (1 - \alpha) Q/K$. If the distribution of output between labor and capital is different from that implied by the marginal productivity conditions, so that the factor shares differ from α and $(1 - \alpha)$, estimating the production function using physical data will still produce estimates of the output elasticities of α and of $(1 - \alpha)$. If, however, we use value data, then, as we shall see below, the estimates of the supposed output elasticities will take the values of the shares and not the true output elasticities.

In these circumstances, therefore, the discussions concerning the appropriate estimation procedures of the production function (whether it should be part of a simultaneous equation framework, etc.) become relevant. It is important to note, however, that it is possible to get misleading results estimating a "production function" even with physical data. Define $m \equiv gQ/L$ and $n \equiv (1 - g)Q/K$. The parameter g need not be labor's actual share; it can take any arbitrary value we like, provided that $0 \leq g \leq 1$. If we were to regress $\ln Q$ on $\ln m$, $\ln n$, $\ln L$, and $\ln K$, using either cross-sectional or time-series data, we would find that the coefficients of $\ln m$ and $\ln L$ took a value of g and those of $\ln n$ and $\ln K$ would be $(1 - g)$. Consequently, by an appropriate choice of m and n , we can obtain *any* numerical values of the estimated coefficients g and $(1 - g)$ that we wish and a perfect statistical fit. Alternatively, in the case of time-series data, we may determine a complex time trend that closely approximates $g \ln m + (1 - g) \ln n$ and we would get the same estimates of the output elasticities. It should be noted that although $Q = Bm^g n^{1-g} L^\beta K^{1-\beta}$ is definitionally true and is an identity, it is not necessarily related to either the marginal productivity conditions or to the way output is actually divided up between labor and capital.

In the case under discussion here in which there is no technical change or in the case of cross-sectional data in which there is a common technology, this procedure would give an incorrect specification of the “physical” production function as $g \ln m + (1 - g) \ln n$ is not a constant. Regressing $\ln Q$ on a constant and $\ln L$ and $\ln K$ would be the correct specification and also regressing Q on L and K would always give estimates of f_L and f_K , as noted above.

With time-series data, technical progress may be incorporated into the production function using physical data by including a correctly specified time trend. It should be noted, however, that there are an infinite number of possible linear and nonlinear time trends that could be used and that would give good, or indeed perfect, fits to the data. The estimates of the output elasticities will be statistically significant, but significantly different from the true elasticities, even though physical data is used. The true production function can only be estimated if we know what is the true path of technical change, which of course is not possible.

The neoclassical approach disregards these problems and assumes that the true production function can always be estimated using physical data.¹⁵ It moves seamlessly from this to the use of value data, where it is also assumed that the aggregate production function reflects the underlying technology of the economy. As Ferguson [1971, 250] noted “neoclassical theory deals with macroeconomic aggregates, usually by constructing the aggregate theory by analogy with the corresponding microeconomic concepts. Whether or not this is a useful concept is 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 is thus implicitly arguing that, although the neoclassical production function should be specified in terms of physical quantities, these may be adequately proxied by deflated monetary values where this is necessary for aggregation.

Given the problems with physical data, noted above, the question arises as to why, when using value data, the observed income identity, *per se*, matters, compared with any other hypothetical way value added could be divided between labor and capital. For example, as with the use of physical data, it is possible by including in the regression using value data the additional variables $\ln m'$ and $\ln n'$ (where $m' = gV/L$ and $n' = (1 - g)V/J$ and can now only be expressed in monetary values) to produce estimates of the putative elasticities of L and J that take the values g and $(1 - g)$, respectively. The choice of these values is purely arbitrary, however, and although the statistical fit would be near perfect, the estimated “output elasticities” would not equal the observed factor shares.

To answer the question posed above about what is important about the accounting identity, let us assume that firms pursue a markup pricing policy under which the price is determined by a fixed markup on unit labor costs. (We have assumed for simplicity that no materials are used in production, because, in practice, firms mark up on normal unit costs. See Lee [1999] for a detailed discussion.) Thus $p_i = (1 + \pi_i)w_iL_i/Q_i$, where π is the markup. Value added is $V_i = p_iQ_i = (1 + \pi_i)w_iL_i$ and for industry as a whole $V = \sum p_iQ_i = \sum (1 + \pi_i)w_iL_i$, or, approximately, $V = (1 + \pi)wL$, where π is the average markup and w is the average wage rate. Labor's share is $a = 1/(1 + \pi)$ and will be constant to the extent that the markup does not vary. In practice, it is likely to vary

to the degree that the composition of firms with differing markups alter and there are changes in the individual markups, which may be temporary, as a result of the wage-bargaining process.¹⁶

We also have the identity $V \equiv wL + rJ$ where rJ , total profits, is equal to πwL and capital's share is given by $(1 - a) = \pi/(1 + \pi)$. The identity now poses a major problem for the aggregate production function. Suppose that w and r are the same across the units of observation in cross-sectional data or do not change over time when time-series data is used. If we were to estimate $V = b_3L + b_4J$, then the estimates of b_3 and b_4 will always be w and r , respectively. If factor shares are constant, then, as we have seen above, an approximation to the accounting identity will be given by $V = Bw^\alpha r^{1-\alpha} L^a J^{1-a} = A L^a J^{1-a}$, where A is a constant. But the causation is now from the identity to the multiplicative power function, not the other way around. The values of the putative elasticities are determined by the value of the markup and do not reflect any technological relationship. This explains why the multitude of cross-regional and cross-industry studies that Douglas undertook in the 1920s and 1930s gave such a good fit to the data where the estimates of the "output elasticities" were very close to the observed factor shares. The reason was, as we have noted above, that value added is simply defined as $V = wL + rJ$ and the wage and the profit rate were roughly constant (especially in comparison to the variation in the other variables). The cross-sectional regression will therefore give $\ln V = c + a \ln L + (1 - a) \ln J$, where a and $(1 - a)$ are the observed factor shares and c equals the sum of a constant term and $(a \ln w + (1 - a) \ln r)$ and so is itself a constant.

In the case of time-series data, the fact that the weighted logarithms of the wage rate and the rate of return are strongly trended also means that the statistical estimates of the production function, in which technical progress is proxied by a time trend, will empirically reflect the underlying accounting identity. As we have noted above, these regressions often require the capital stock to be adjusted for capacity utilization, which removes or reduces the cyclical fluctuations in $\ln r$ so that it more closely approximates a constant and $a \ln w + (1 - a) \ln r$ more closely approximates a linear time trend. (See also McCombie [2000-2001] and Felipe and Adams [2005].)

If the markup alters, then V will alter as well through the accounting identity, even though the physical composition of output remains unaltered. This means that if we use the new data to estimate a Cobb-Douglas, the estimates of the "output elasticities" will also change, reflecting the change in the markup. For example, if the markup increases, labor's share will fall and this will result in a fall in labor's "output elasticity" and a corresponding rise in capital's share and its "output elasticity".

The fact that the Cobb-Douglas function gives a good fit to the data does not imply that the aggregate elasticity of substitution (a parameter that likely does not exist) is unity. Shares may vary, in which case a more flexible function than the Cobb-Douglas, such as the translog, will give a better fit; but we still cannot be sure that the data are telling us anything about the underlying technology of the economy, as we have seen.

The problem is that in the case of constant shares, the accounting identity will always give a good fit to the data, even if the technology of the firms in the economy is nothing remotely like a Cobb-Douglas. This is not an identification problem in the econometric sense of the term. In the latter case we have, say, two behavioral equations, but where the reduced form consists of a hybrid of the parameters of the equations. In

TABLE 1
The Relationship between the Accounting Identity and the
Aggregate Production Function Using Time-Series Data

The Accounting Identity	The Neoclassical Production Function
Prices are a markup on unit labor costs:	The micro production function with constant returns to scale is given by:
$p_i = (1 + \pi_p)(w_i L_i / Q_i)$	$Q_i = A e^{\lambda t} L_i^a K_i^{(1-a)}$
A constant markup gives constant shares of labor (a) and capital ($1 - a$) in total value added, regardless of the underlying technology.	Aggregation problems and the Cambridge Capital Theory Controversies show that theoretically the aggregate production function does not exist. Nevertheless, it is assumed that:
$a = 1/(1 + \pi)$ and $(1 - a) = \pi/(1 + \pi)$.	$\Sigma Q_i = Q = A e^{\lambda t} L^a K^{(1-a)}$
The accounting identity is given by:	Assuming (i) perfect competition; and (ii) the aggregate marginal productivity theory of factor pricing gives:
$p_i Q_i \equiv w_i L_i + r_i J_i$, where $r_i = (p_i Q_i - w_i L_i) / J_i$.	$p(\partial Q / \partial L) = p f_L = w$ and $p(\partial Q / \partial K) = p f_K = \rho$
Summing over industries gives:	From Euler's theorem:
$V = \Sigma p_i Q_i = wL + rJ$.	$Q = f_L L + f_K K$
There are no serious aggregation problems. Aggregation may actually reduce the variability of the aggregate factor share compared with the individual factor shares.	and the cost identity is
By definition (and making <i>no</i> assumption about the state of competition or the mechanism by which factors are rewarded) the following conditions hold:	$pQ = wL + \rho K$ or $Q = (w/p)L + (\rho/p)K$,
$\partial V / \partial L \equiv w$ and $\partial V / \partial J \equiv r$	where w/p and ρ/p are physical measures and equal f_L and f_K
Given constant factor shares, the accounting identity at time t may be written as:	It is assumed for empirical analysis that $Q = V$ and $K = J$.
$V_t = B W_t^a r_t^{(1-a)} L_t^a J_t^{(1-a)}$	Using time-series data and estimating $\ln V_t = c + b_5 t + b_6 \ln L_t + b_7 \ln J_t$ provides estimates of b_6 and b_7 , which are the aggregate output elasticities of labor and capital. If a good statistical fit is found, it is inferred that the estimation has not refuted the hypothesis of the existence of the aggregate production function.
or, assuming the stylized fact that	
$a_t \hat{w}_t + (1 - a_t) \hat{r}_t = a \hat{w} = \lambda$, as:	
$V_t = B e^{\lambda t} L_t^a J_t^{(1-a)}$	
Estimating $\ln V_t = c + b_5 t + b_6 \ln L_t + b_7 \ln J_t$ gives estimates of b_6 and b_7 exactly equal to the factor shares for <i>definitional</i> reasons, that is:	The estimates of b_6 and b_7 equal the observed factor shares, that is:
$b_6 = a$ and $b_7 = (1 - a)$.	$b_6 = \alpha = a$ and $b_7 = (1 - \alpha) = (1 - a)$
It is always possible to find an approximation that will give a perfect statistical fit to the data.	if assumptions (i) and (ii) above hold.
	If this is found to occur, it constitutes a failure to refute the theory that markets are competitive and factors are paid their marginal products.

TABLE 1—Continued

The Equifinality Theorem.

Estimating $\ln V = c + b_3 t + b_3 \ln L + b_7 \ln J$ will always give a perfect fit to the data, provided factor shares are constant and the stylized fact $a_t \hat{w}_t + (1 - a_t) \hat{r}_t = a \hat{w} = \lambda$ holds. This is the case regardless of whether there is a “true” underlying aggregate Cobb-Douglas production function (no matter how theoretically implausible this may be) or no aggregate production function exists at all. The data cannot discriminate between these two cases. (The same result holds using growth rates.)

If the conditions of constant factor shares and a constant growth of the weighted wage and profit rates are not met, it is still possible to obtain a perfect fit by a more flexible approximation to the accounting identity than that given by the Cobb-Douglas.

It is therefore not possible empirically to test the existence of the aggregate production function or the aggregate marginal productivity theory of factor pricing.

this case we should be able, in principle, to identify the equations by including variables in one equation that do not appear in the other and that, hence, act to identify the latter. This is not the case with the identity. We know that the estimates will always be picking up an approximation to the identity. Felipe and McCombie [2003] have termed this the “equifinality theorem.” Table 1 provides a summary of the case when factor shares are constant and so the approximation to the underlying identity is a Cobb-Douglas function.

The use of value data also poses difficulties for the neoclassical growth accounting approach. This assumes an aggregate production function, constant returns to scale, and the marginal productivity theory of factor pricing. Under these assumptions, the growth of total factor productivity (the “Solow residual” or, somewhat misleadingly, the rate of technical progress) is given by the weighted growth of the real wage rate and the rate of profit. Hence the growth of total factor productivity is equal to

$$(11) \quad T\hat{F}P_t \equiv \lambda_t \equiv a_t \hat{w}_t + (1 - a_t) \hat{r}_t \equiv \hat{V}_t - a_t \hat{L}_t + (1 - a_t) \hat{J}_t.$$

The problem is that this is formally identical to the accounting identity in value terms and, hence, as we have noted, the neoclassical assumptions underlying it cannot be tested. It should be noted that usually the neoclassical growth accounting approach does not use the observed rate of profit, but instead the rental price of capital. The latter is the cost to the firm of its capital and is derived from an optimization model of the representative firm that exhibits constant returns to scale. The growth of the rental price of capital should differ from that of the observed rate of profit if there are noncompetitive or economic profits. This does not affect our argument, as value added in these circumstances should be defined to exclude any monopoly or economic profits and estimation of a putative production function will, by definition, give a very close statistical fit with the estimate coefficients equal to the cost (rather than revenue) shares.

An Illustration Using a Simulation Study

The implications of the distinction between the use of physical and value data are far reaching as may be seen from simulation studies undertaken by McCombie [2000]. He assumed a two-firm economy and postulated that the firms, although they made

different products, had identical production functions, expressed in physical units, of the form $Q_{it} = A_{i0} L_{it}^a J_{it}^{(1-a)}$, $i = 1, 2$. There was no technical change and labor's output elasticity was taken deliberately to be 0.25 and capital's output elasticity was 0.75. Q , L , and K were allowed to grow over several periods and the resultant pooled time-series data of the two firms were used to estimate the Cobb-Douglas production function. Consequently, the estimation used physical data. The estimates of the output elasticities of labor and capital were the expected one-quarter and three-quarters. When growth rates were used, the coefficient of the intercept was zero and the estimated coefficients once again took the expected magnitudes.

An aggregate production function was next estimated under the following assumptions. Both firms pursued a policy of a markup of 1.33 on unit labor costs. The nominal wages rate was the same for both firms and the measure of aggregate output was given by $V = \Sigma(1 + \pi) wL_i = (1 + \pi)wL$. Both firms used identical machines and so there was no problem in aggregating the capital stock into a single index.

Estimating the aggregate production function using time-series data gave a near perfect fit¹⁷ and a value of the putative output elasticity of labor that was 0.75 (*not* 0.25) and a value of capital's output elasticity of 0.25 (*not* 0.75). The reason is, of course, that the factor shares of labor and capital in value terms were 0.75 and 0.25, as labor's share is equal to $1/(1+\pi)$. The values of these shares were, through the underlying accounting identity, responsible for the values of the estimates of the putative output elasticities.

Clearly, if there was technical change at the firm level, the use of factor shares to calculate the contribution of the growth of the factor inputs to economic growth and hence the growth of aggregate total factor productivity would give a very misleading estimate. This is because the aggregate factor shares for labor and capital differ from the actual output elasticities at the firm level. Of course, the use of aggregate value data cannot show this. Estimating the aggregate Cobb-Douglas production function always gave estimates of the aggregate "output elasticities" that were equal to the factor shares.

A second simulation allowed the two individual firms to exhibit increasing returns to scale of 1.20, but they again had the same markup pricing policy of 1.33. The simulation results produced estimates of labor's and capital's aggregate "output elasticities" of 0.75 and 0.25, respectively. In other words, with the use of value data it is not possible to reject the null hypothesis of constant returns to scale, even though the separate underlying production functions exhibited large increasing returns to scale. This was because, when value data are used, the underlying accounting identity ensures that the Cobb-Douglas transformation will exhibit "constant returns to scale". The effect of the increasing returns in the underlying firm production functions in a dynamic context is captured by a growth rate of factor returns in the accounting identity. It is not possible, therefore, to separate empirically the impact of technical progress and increasing returns to scale.

Finally, to illustrate the general argument more forcefully, the paths of Q_{it} , L_{it} , and K_{it} were all generated as random numbers. In other words there were no true underlying individual production functions. Again, with a constant markup on unit labor costs, a close fit was found for the aggregate Cobb-Douglas production function

with the aggregate output elasticities of labor and capital (which do not exist) once again taking well-determined values of 0.75 and 0.25.

REACTIONS TO THE CRITIQUE

Generally speaking, the critique has been almost totally ignored in the literature. The critique, as it applies to cross-sectional data, is briefly mentioned in three now somewhat dated econometric textbooks, namely, Cramer [1969], Wallis [1979] and Intriligator [1978]. While they generally note the problems posed by the accounting identity, they do not, however, take the argument to its logical conclusions, namely, that it undermines the possibility of testing the aggregate production function. Walters' [1963] early, but influential, study surveying production and cost functions includes a discussion of the Phelps Brown [1957] criticism, but it is buried in a short paragraph on page 37: "The early commentators pointed out that the data may be explained by what Bronfenbrenner called the interfirm function $V = wL + rJ$. Evidence has been adduced by Phelps Brown [1957] to show that the scatter of observations of Australia in 1909 can be explained in terms of this simple linear relationship. Thus, in fitting a Cobb Douglas function (with $\alpha + \beta = 1$), we merely measure the share of wages in the value added. The result does not provide a test of the marginal productivity law."¹⁸ In the very next paragraph, however, he goes on to discuss the CES production function in which "the different ratios of factor prices will generate observations which should trace out the production function" [Walters, 1963, 37].

Passing remarks (literally a few sentences) may be found in Robinson [1970] and Harcourt [1982], and a brief passage in Lavoie [1992]. These authors, however, *did* fully understand the damaging implications of the critique for the concept of the aggregate production function.

The fact that the critique applies to the estimation of time-series data has not permeated down to the textbooks, with the exception of a short reference in Heathfield and Wibe [1987]. They mention Shaikh's [1974] critique, but seemingly dismiss it because of Solow's [1974] one-page rejoinder, which erroneously rejects Shaikh's [1974] paper as simply wrong (see Shaikh [1980] for a full-fledged rejoinder).

Shaikh's [1974] paper contained two themes, however, only one of which was really discussed by Solow [1974]. The first was the one outlined in the section above, demonstrating the equivalence of the accounting identity and the Cobb-Douglas production function. The second was the tautological nature of the procedure involved in Solow's [1957] procedure for estimating various specifications of the production function, and it was this that Solow [1974] addressed. He makes no mention of the problem posed by the underlying identity.

Solow's procedure was first to calculate the growth of total factor productivity or the residual, using annual data, as $\lambda_t \equiv (\hat{V}_t - \hat{L}_t) - (1 - a_t)(\hat{J}_t - \hat{L}_t)$, where $(1 - a_t)$ is capital's share calculated from the national accounts. Thus, Solow assumes the existence of a well-behaved production function, perfect competition, and the neoclassical marginal productivity theory of factor pricing. An index of technology $A(t)$ is then calculated by setting A in the base year equal to unity and then using the annual values of λ_t to calculate the other values of the index. The production function is then

“deflated” by $A(t)$ to remove the effect of technical change. By this procedure, Solow was in fact estimating an identity [Felipe and McCombie, forthcoming]. Moreover, Shaikh [1974; 1980] has shown that any data set (even one in which the plot of productivity on capital per worker traces out the word HUMBUG) will give a good fit to a Cobb-Douglas function by this method so long as factor shares are constant.

Solow’s [1974] rejoinder was that he was not *testing* the neoclassical production function at all, as he had already assumed that it existed. He had, after all, assumed the marginal productivity theory of factor pricing in constructing $A(t)$. Nevertheless, Solow does consider that one can, in principle, “test” the production function. “When someone claims that aggregate production functions work, he means (a) that they give a good fit to input-output data *without* the intervention of data derived from factor shares; and (b) that the function so fitted has partial derivatives that closely mimic observed factor prices” [Solow, 1974, 121, italics in the original]. He then delivers his supposed *coup de grâce* by freely estimating the Cobb-Douglas function in logarithmic form using the Humbug data with a linear time trend to capture technical change and finds no statistically significant relationship. Solow takes this as showing conclusively that Shaikh’s artificial data do not capture even a hypothetical production function. “If this were the typical outcome with real data, we would not now be having this discussion” [Solow, 1974, 121].

Ironically, freely estimating a Cobb-Douglas production function with Solow’s [1957] own data also produces such poor results that if Solow himself had undertaken this, by his own criterion, he would have been forced to concede that the data rejected the existence of the production function [McCombie, 2000-2001]. However, the poor fit in logarithmic form is simply due to the fact that the linear time trend does not provide a good approximation to the weighted logarithm of wages and the rate of profit. There is, of course, nothing to say that “technical change” should be a linear function of time. Shaikh [1980] finds that Humbug data give an excellent fit to the Cobb-Douglas, provided that a complex time trend is used.

As we discussed above and now elaborate, Solow [1987] returns to the critique as outlined by Shaikh, arguing that Shaikh’s analysis would hold even when there is a well-defined micro-production function estimated using physical units. As, he argues, it cannot hold in this case, then the implication is that it equally cannot apply to value data, although this does not logically follow.¹⁹ Let us consider the argument only as it applies to value data. Solow’s contention is that from the accounting identity with constant shares we may derive the equation $V_t = Bw_t^a r_t^{(1-a)} L_t^a J_t^{(1-a)}$, as we have seen above. But we know that $w_t \equiv aV_t/L_t$ and $r_t \equiv (1-a)V_t/J_t$. Following Solow, let us substitute these expressions into the equation for V_t . This gives us:

$$(12) \quad V_t = \left(\frac{Ba^a(1-a)^{(1-a)} V_t}{L_t^a J_t^{(1-a)}} \right) \left(L_t^a J_t^{(1-a)} \right).$$

As $B = a^{-a}(1-a)^{-(1-a)}$, it follows that:

$$(13) \quad V_t = \left(\frac{V_t}{L_t^a J_t^{(1-a)}} \right) \left(L_t^a J_t^{(1-a)} \right).$$

and, assuming no technical change:

$$(14) \quad V_t = \left(\frac{f(L_t, J_t)}{L_t^a J_t^{(1-a)}} \right) \left(L_t^a J_t^{(1-a)} \right) = f(L_t, J_t).$$

“What Shaikh has discovered, in other words, is that any production function can be written as the product of a Cobb-Douglas and something else; and the something else is the production function divided by the Cobb-Douglas” [Solow, 1987, 20]. This argument presupposes that we have data from which the “true” production function can be estimated, however. As we have seen, because of the accounting identity, Equation (14), $V_t = f(L_t, J_t)$, will give a good fit to the data, even though there is no well-defined aggregate production function. So all that Solow has done is to show that V_t equals $f(L_t, J_t)$, which he *assumes* to be the “true technology” and, to this extent, his argument is tautological. As we have shown, because of the underlying accounting identity, Equation (14) may be written as $V_t = f(L_t, J_t) = Bw_t^a r_t^{(1-a)} L_t^a J_t^{(1-a)}$, reflecting merely the identity.

Solow [1987, 20-21] further comments: “What *does* happen if I regress $\ln x$ on $\ln u$ and $\ln v$? [x is output and u and v are the inputs]. That depends on whether $\ln A$ is orthogonal to $a \ln u + (1 - a) \ln v$. But $\ln A$ is $\ln x - [a \ln u + (1 - a) \ln v]$. That is to say, orthogonality will hold precisely when the true production function is Cobb-Douglas and will fail when it is not. In the latter case, nothing prevents us from doing the regression but the estimated elasticities will not coincide with a and $1 - a$, and the regression will show systematic errors” (emphasis in the original). It is possible to interpret Solow along the following lines. As we have seen above, if there is no technical change and we use *physical* data, then, if there exists a true Cobb-Douglas micro-production function, $\ln A$ will be a constant and the estimated coefficients will be equal to the true output elasticities. If factors are paid their marginal products, the output elasticities will equal the appropriate factor shares. What Solow fails to appreciate, however, is the argument made above. Value data, in practice, has to be used because of the heterogeneity of output and the need to aggregate any underlying micro-production functions. If real wages measured in monetary units and the rate of profit are constant (strictly speaking, if $w^a r^{(1-a)}$ is constant), then regressing $\ln V$ on $\ln L$ and $\ln J$ will, with constant factor shares, also give a good fit to the data with the “output elasticities” equaling the factor shares. This will be for purely definitional reasons, however, and reflects the underlying accounting identity. As we have seen, there need be no well-defined aggregate production function at all. It also should be noted that, *pace* Solow, if $a \ln w + (1 - a) \ln r$ (Solow’s $\ln A$) is not constant over time, then it is still possible for the data to give a good fit to the Cobb-Douglas, with the estimated “output elasticities” equaling the factor shares, even though $a \ln w + (1 - a) \ln r$ is not orthogonal to $a \ln L + (1 - a) \ln J$. All that is required is that the included time trend (not necessarily linear) be a good proxy for $a \ln w + (1 - a) \ln r$.

The reason why the critique has had so little impact seems thus to be partly due to the fact that the full implications of it were never appreciated, or else that Solow’s [1974, 1987] mistaken opinion that it is simply wrong has been uncritically accepted. McCombie [1998b] has considered the possible reasons from a methodological point of

view. The critique is not mentioned in any recent textbooks (which, as Kuhn [1962] has demonstrated, play such an important part in the persistence of a paradigm) or the plethora of studies that use the aggregate production function. Thus while, *pace* Solow, the critique has not been refuted, it has nevertheless been almost totally ignored.

CONCLUSIONS

This paper has revisited and extended some problems that, in particular, the use of value data (value added or gross output) pose for the estimation of production functions. It is shown that the estimated coefficients of the supposed production function may be doing no more than capturing an underlying income identity from which data used in the estimation of the production function are drawn. This criticism has its origins in a paper by Phelps Brown [1957], although it was anticipated to a certain extent by, for example, Marshak and Andrews [1944], and was applied originally to the cross-section studies of Douglas. The critique was generalized by Shaikh [1974] and Simon [1979b], however, to the use of time-series data, and further elaborated by the authors. We have shown that the data may give a good fit to the aggregate Cobb-Douglas, even though it is clear that either the aggregation conditions are violated, or the underlying micro-production relations are not Cobb-Douglas, or there is no neo-classical production function at all. All these arguments suggest that the results of regressions purporting to estimate an aggregate production function (whether it is a Cobb-Douglas or a more flexible functional form) must be treated with caution. (Moreover, even the use of physical data in engineering production functions encounters problems.)

It also explains why a number of studies that have either fixed coefficients production functions, Houthakker [1955-56], Shaikh [1987], and Nelson and Winter [1982], the last with a distinctly non-neoclassical evolutionary model with satisficing behavior, give good fits to a Cobb-Douglas production function. It is because factor shares are constant. Fisher [1971], in his well-known simulation experiment, used Cobb-Douglas micro-production functions, but ones in which the conditions for successful aggregation were deliberately violated. To his evident surprise, he found that an aggregate Cobb-Douglas production function gave good predictions of wages. The reason was that factor shares at the aggregate level were stable and this gave rise to the good statistical fits of the Cobb-Douglas. The causation ran from the constant shares to the putative Cobb-Douglas production function, not vice versa. Further problems are shown by Hartley [2000]. He used the production function used in the real business cycle model of Hansen and Sargent [1990]. The production function is not a Cobb-Douglas, although the simulated data have the property that factor shares are roughly constant. He shows that there is no consistent relationship between true technological shocks and the Solow residuals calculated by the standard growth accounting approach.

The conclusions of the paper are thus rather nihilistic. If the critique is accepted, then all work that relies on the aggregate production function, whether empirical or theoretical, must be viewed with the utmost skepticism. This inevitably raises the question as to what should be put in its place. This clearly is an important area for future work, but which is not discussed in this paper.

NOTES

Jesus Felipe is a Senior Economist at the Asian Development Bank (Manila, Philippines) and J. S. L. McCombie is Fellow in Economics, Downing College, Cambridge. We are grateful for comments from participants at seminars at the Departments of Economics at the University of Wellington, New Zealand; University of Otago, New Zealand; the University of Campinas, Brazil; the XXIX Encontro Nacional de Economia, Brazilian Congress of Economists, Salvador, Brazil, 2001; and the 2002 meeting of the Eastern Economic Association, New York. Two anonymous referees provided helpful comments. The 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. The usual disclaimer applies.

1. It should be noted that apart from a very few “engineering” production function studies, all other estimations of production functions have used value data, even at low levels of aggregation, for example, the 3- or 4-digit SIC.
2. Cohen and Harcourt [2003] provide an illuminating retrospective view of the implications of the Cambridge Capital Theory Controversies and Birner [2002] examines the controversies from a methodological standpoint.
3. See, for example, Solow [1955-56].
4. Estimates using time-series data sometimes give poor results. The reasons for this are discussed below.
5. An exception is pooled cross-country data for the advanced and less developed countries. We discuss this further below.
6. This is not a new critique, but first came to prominence in a rudimentary form in Phelps Brown’s [1957] criticism of Douglas’s cross-industry results. This critique was later formalized by Simon and Levy [1963] and elements of it can be found in Marshak and Andrews [1944]. Shaikh [1974; 1980; 1987] generalized it to time-series estimation of production functions and Simon [1979b] also considered the criticism in the context of both cross-section and time-series data. The criticism was re-examined and extended by Felipe and Adams [2005], Felipe and McCombie [2001; 2002a, b; 2003; 2005; forthcoming], Felipe [2001a, b], Felipe and Holz [2001], McCombie [1987; 1998a; 2000; 2000-2001; 2001], McCombie and Dixon [1991], and McCombie and Thirlwall [1994]. The critique as applied to cross-section data was also “rediscovered” by Samuelson [1979].
7. We use V and J to refer to the value measures; Q and K are used below to denote the physical measures of output and capital.
8. It is assumed that there is enough variation in the data to provide well-defined estimates of the parameters.
9. We are grateful to an anonymous referee for pointing this out to us.
10. The CES function may be regarded akin to a Box Cox transformation, which is simply a mathematical transformation that attempts to find the best fit for the identity when w and r vary.
11. Phelps Brown [1957] does consider time-series data, but mainly when there is no time trend included (that is, the data are not de-trended). He shows that, in these circumstances, the estimated coefficients will merely reflect the historical growth rates of the various variables.
12. As we are dealing with an identity, we treat the regressions using either logarithms of the levels or exponential growth rates as equivalent.
13. This is because the two variables are related through capital’s share in output, that is, $r = (1 - a)W/J$.
14. Euler’s theorem does not, of course, require linear homogeneity. It is necessary for the production function to exhibit this if there is to be no “adding-up” problem.
15. Solow [1987] noted the above critique, and erroneously considered it self-evidently wrong that the true production function could not be estimated using physical data. From this he drew the conclusion, which is a *non sequitur*, that the critique also could not apply to value data. This is discussed below.
16. Solow [1958] has shown that aggregation may well *decrease* the variability of the aggregate factor shares compared with the shares of the individual firms/industries.
17. This was because the individual production functions were constructed using data with only a small random term to prevent perfect multicollinearity.

18. The notation has been changed to make it consistent with that in this paper and a footnote omitted.
19. We have shown above, however, that the use of physical data is not problem free.

REFERENCES

- Birner, J.** *The Cambridge Controversies in Capital Theory*. London: Routledge, 2002.
- Cobb, C. W. and Douglas P. H.** A Theory of Production. *American Economic Review, Papers and Proceedings*, March 1928, 139-65.
- Cohen, A. and Harcourt, G. C.** Whatever Happened to the Capital Controversies? *Journal of Economic Perspectives*, Winter 2003, 199-214.
- Cramer, J. S.** *Empirical Econometrics*. Amsterdam: North-Holland, 1969.
- Douglas, P. H.** Are There Laws of Production? *American Economic Review*, March 1948, 1-41.
- _____. The Cobb-Douglas Production Function Once Again: Its History, its Testing, and Some Empirical Values. *Journal of Political Economy*, October 1976, 903-15.
- Felipe, J.** Endogenous Growth, Increasing Returns, and Externalities: An Alternative Interpretation of the Evidence. *Metroeconomica*, November 2001a, 391-427.
- _____. Aggregate Production Functions and the Measurement of Infrastructure Productivity: A Reassessment. *Eastern Economic Journal*, Summer 2001b, 323-44.
- Felipe, J. and Adams, G. F.** 'A Theory of Production'. The Estimation of the Cobb-Douglas Function: A Retrospective View. *Eastern Economic Journal*, Summer 2005, 427-45.
- Felipe, J. and Fisher, F. M.** Aggregation in Production Functions: What Applied Economists Should Know. *Metroeconomica*, May-September 2003, 208-62.
- Felipe, J. and Holz, C.** Why do Production Functions Work? Fisher's Simulations, Shaikh's Identity and Some New Results. *International Review of Applied Economics*, July 2001, 261-85.
- Felipe, J. and McCombie, J. S. L.** The CES Production Function, the Accounting Identity and Occam's Razor. *Applied Economics*, August 2001, 1221-32.
- _____. A Problem with Some Recent Estimations and Interpretations of the Mark-Up in Manufacturing Industry. *International Review of Applied Economics*, April 2002a, 187- 215.
- _____. Are Estimates of Labour Demand Functions Merely Statistical Artefacts? Paper presented at the 2002 meeting of the Eastern Economic Association, Boston, 2002b.
- _____. Some Methodological Problems with the Neoclassical Analysis of the East Asian Economic Miracle. *Cambridge Journal of Economics*, August 2003, 695-721.
- _____. Why Are Some Countries Richer than Others? A Skeptical View of Mankiw-Romer-Weil's Test of the Neoclassical Growth Model. *Metroeconomica*, July 2005, 360-92.
- _____. Is a Theory of Total Factor Productivity Really Needed? *Metroeconomica*, forthcoming.
- Ferguson, C. E.** Capital Theory up to Date: A Comment on Mrs Robinson's Article. *Canadian Journal of Economics*, May 1971, 250-54.
- Fisher, F. M.** Aggregate Production Functions and the Explanation of Wages: A Simulation Experiment. *The Review of Economics and Statistics*, November 1971, 305-25.
- _____. Aggregation Problems, in *The New Palgrave. A Dictionary of Economics*, edited by J. Eatwell, M. Milgate, and P. Newman. Basingstoke: Macmillan, 1987, 53-55.
- _____. *Aggregation: Aggregate Production Functions and Related Topics. The Collected Papers of Franklin M. Fisher*, edited by J. Monz. London: Harvester Wheatsheaf, 1992.
- _____. Aggregate Production Functions—A Pervasive, but Unpersuasive, Fairytale. *Eastern Economic Journal*, Summer 2005, 489-91.
- Hansen, L. P. and Sargent, T. J.** Recursive Linear Models of Dynamic Economies. *NBER Working Paper Series*, no. 3479, 1990.
- Harcourt, G. C.** *Some Cambridge Controversies in the Theory of Capital*. Cambridge: Cambridge University Press, 1972.
- _____. *The Social Science Imperialists*, edited by P. Kerr. London: Routledge and Kegan Paul, 1982.
- Hartley, J. E.** Does the Solow Residual Actually Measure Changes in Technology? *Review of Political Economy*, January 2000, 27-44.
- Heathfield, D. F. and Wibe, S.** *An Introduction to Cost and Production Functions*. Basingstoke: Macmillan, 1987.

- Houthakker, H. S.** The Pareto Distribution and the Cobb-Douglas Production Function in Activity Analysis. *Review of Economic Studies*, 1955-56, 27-31.
- Intriligator, M. D.** *Econometric Models, Techniques and Applications*. Englewood Cliffs, N.J.: Prentice Hall, 1978.
- Jorgenson, D. W.** Investment and Production: A Review, in *Frontiers of Quantitative Economics*, vol. II, edited by M. D. Intriligator and D. A. Kendrick. Amsterdam: North Holland, 1974.
- Kuhn, T. S.** *The Structure of Scientific Revolutions*. Chicago: Chicago University Press, 1962 (2d ed. 1970).
- Lavoie, M.** *Foundations of Post-Keynesian Economic Analysis*. Aldershot: Edward Elgar, 1992.
- Lee, F. S.** *Post Keynesian Price Theory*. Cambridge: Cambridge University Press, 1999.
- Lucas, R. E.** Capacity, Overtime, and Empirical Production Functions. *American Economic Review, Papers and Proceedings*, May 1970, 23-27.
- Marshak, J. and Andrews, W. H.** Random Simultaneous Equations and the Theory of Production. *Econometrica*, July-October 1944, 143-205.
- McCombie, J. S. L.** Does the Aggregate Production Function Imply Anything about the Laws of Production? A Note on the Simon and Shaikh Critiques. *Applied Economics*, August 1987, 1121-36.
- _____. 'Are There Laws of Production?: An Assessment of the Early Criticisms of the Cobb-Douglas Production Function. *Review of Political Economy*, April 1998a, 141-73.
- _____. Rhetoric, Paradigms, and the Relevance of the Aggregate Production Function. *Method, Theory and Policy in Keynes. Essays in Honour of Paul Davidson, Vol. III*, edited by P. Arestis and M.C. Sawyer. Aldershot: Edward Elgar, 1998b.
- _____. The Regional Production and the Accounting Identity: A Problem of Interpretation. *Australasian Journal of Regional Studies*, 2000, 133-55.
- _____. The Solow Residual, Technical Change and Aggregate Production Functions. *Journal of Post Keynesian Economics*, Winter 2000-2001, 267-97 (errata *ibid.* Spring 2001, 544).
- _____. What does the Aggregate Production Function Tell Us? Second Thoughts on Solow's 'Second Thoughts on Growth Theory.' *Journal of Post Keynesian Economics*, Summer 2001, 589-615.
- McCombie, J. S. L. and Dixon, R.** Estimating Technical Change in Aggregate Production Functions: A Critique. *International Review of Applied Economics*, January 1991, 24-46.
- McCombie, J. S. L. and Thirlwall, A. P.** *Economic Growth and the Balance-of-Payments Constraint*. Basingstoke: Macmillan, 1994.
- Nelson, R. R. and Winter, S. G.** *An Evolutionary Theory of Economic Change*. Cambridge, Mass.: Harvard University Press, 1982.
- Pasinetti, L. L.** The Structure of Long-Term Development: Concluding Comments, in *Economic Growth and the Structure of Long-Term Development*, edited by L. L. Pasinetti and R. M. Solow. Basingstoke: Macmillan, 1994.
- Phelps Brown, E. H.** The Meaning of the Fitted Cobb-Douglas Function. *Quarterly Journal of Economics*, November 1957, 546-60.
- Robinson, J. V.** Capital Theory up to Date. *Canadian Journal of Economics*, May 1970, 309-17.
- Samuelson, P. A.** Paul Douglas's Measurement of Production Functions and Marginal Productivities. *Journal of Political Economy*, June 1979, 923-39.
- Shaikh, A.** Laws of Production and Laws of Algebra: The Humbug Production Function. *Review of Economics and Statistics*, February 1974, 115-20.
- _____. Laws of Production and Laws of Algebra: Humbug II, in *Growth, Profits and Property, Essays in the Revival of Political Economy*, edited by E. J. Nell. Cambridge: Cambridge University Press, 1980, 80-95.
- _____. Humbug Production Function, in *The New Palgrave: A Dictionary of Economic Theory and Doctrine*, edited by J. Eatwell, M. Milgate, and P. Newman. London: Macmillan, 1987.
- Simon, H. A.** Rational Decision Making in Business Organizations. *American Economic Review*, September 1979a, 493-513.
- _____. On Parsimonious Explanation of Production Relations. *Scandinavian Journal of Economics*, December 1979b, 459-74.
- Simon, H. A. and Levy, F. K.** A Note on the Cobb-Douglas Function. *Review of Economic Studies*, June 1963, 93-94.
- Solow, R. M.** The Production Function and the Theory of Capital. *Review of Economic Studies*, 1955-56, 101-108.

- _____. Technical Change and the Aggregate Production Function. *Review of Economics and Statistics*, August 1957, 312-20.
- _____. A Skeptical Note on the Constancy of Relative Shares. *American Economic Review*, September 1958, 618-631.
- _____. Review of *Capital and Growth*. *American Economic Review*, December 1966, 1257-60.
- _____. Laws of Production and Laws of Algebra: The Humbug Production Function: A Comment. *Review of Economics and Statistics*. February 1974, 121.
- _____. Second Thoughts on Growth Theory, in *Employment and Growth: Issues for the 1980s*, edited by A. Steinherr and D. Weiserbs. Dordrecht: Martinus Nijhoff Publishers, 1987.
- Sylos Labini, P.** Why the Interpretation of the Cobb-Douglas Production Function Must be Radically Changed. *Structural Change and Economic Dynamics*, December 1995, 485-504.
- Wallis, K. F.** *Topics in Applied Econometrics*. London: Gray-Mills Publishing, 1979.
- Walters, A. A.** Production and Cost Functions: An Econometric Survey. *Econometrica*, January-April 1963, 1-66.
- Wan, H. Y.** *Economic Growth*. New York: Harcourt Brace Jovanovich, 1971.