THE AGGREGATE PRODUCTION FUNCTION: A CONSIDERATION OF SOME EXISTENTIAL PROBLEMS

Jesus Felipe
De La Salle University (Phillipines)
John McCombie
University of Cambridge (United Kingdom)
Corresponding author: jslm2@cam.ac.uk

Manuscript received 4 November 2024; accepted 12 June 2025.

ABSTRACT

The concept of the aggregate production function, together with it supposedly being a test of the neoclassical marginal product theory of distribution, is now central to much of neoclassical macroeconomic theory. Severe criticisms of this concept that include the Cambridge Capital Theory Controversies and the aggregation problem are now simply ignored. The reason is that estimates of the Cobb-Douglas and the Constant Elasticity of Substitution (CES) production functions, *inter alia*, give good statistical fits with the estimated output elasticities being often close to their factor shares. However, the reason for this is that in both cases all that is being estimated is merely the mathematical transformation of an accounting identity. As such, it has no economic implications and, in particular, says nothing about "the laws of production". This analysis considers this criticism of the aggregate production function. In particular, it revisits and extends the criticisms made of Solow's (1957) model of technical change and the use of the aggregate production function. It is shown that Solow's (1974; 1987) defence of the aggregate production function is not convincing. The CES production function is also shown to be

http://dx.doi.org/10.22201/fe.01851667p.2025.333.91816

simply a transformation of the accounting identity and, consequently, also has no economic implications.
Keywords: Aggregate production functions, accounting identity.
JEL **Classification:** B5, O4.

LA FUNCIÓN DE PRODUCCIÓN AGREGADA:
UNA CONSIDERACIÓN DE ALGUNOS PROBLEMAS EXISTENCIALES

RESUMEN

El concepto de la función de producción agregada, junto con el supuesto de que se trata de una prueba de la teoría de la distribución del producto marginal neoclásica, ahora es fundamental para gran parte de la teoría macroeconómica neoclásica. Las severas críticas a este concepto, que incluyen las controversias de la teoría del capital de Cambridge y el problema de agregación, hoy día simplemente son ignoradas. La razón es que las estimaciones de las funciones de producción Cobb-Douglas y la elasticidad constante de sustitución (ECS), *inter alia*, dan buenos resultados estadísticos en los que las elasticidades producto estimadas con frecuencia se aproximan a las participaciones de los factores. Sin embargo, la razón de esto es que en ambos casos todo lo que se está estimando es meramente la transformación matemática de una identidad contable. En tal sentido, no tiene ninguna implicación económica y, en particular, no dice nada acerca de "las leyes de la producción". El presente análisis considera esta crítica de la función de producción agregada. En particular, revisa de nuevo y extiende las críticas hechas al modelo de Solow (1957) del cambio técnico y el uso de la función de producción. Se demuestra que la defensa de Solow (1974; 1987) de la función de producción agregada no es convincente. Se demuestra también que la función de producción ECS es una simple transformación de la identidad contable y, en consecuencia, no tiene ninguna implicación económica.
Palabras clave: funciones de producción agregadas, identidad contable.
Clasificación JEL: B5, O4.

1. INTRODUCTION

The aggregate production function is one of the most widely used neoclassical macroeconomic concepts and has a history of its statistical estimation dating back over nearly a century to Cobb and Douglas's famous paper of 1928. Its theoretical origin may be traced back even further to Wicksell in 1896 (see Samuelson, 1979, p. 296, fn. 2). However, its foundation as the neoclassical explanation of economic growth largely stems from Solow's (1956) use of the aggregate production function in his well-known theoretical growth model. Also extremely influential was Solow's applied use of his model in his 1957 paper, "Technical Change and the Aggregate Production Function". This gave rise to the "growth accounting approach", which attempts to quantify the contributions of various economic factors to overall economic growth. See, for example, Denison's (1967) pioneering study.

Although now widely and uncritically used, it is not an understatement to say that the concept of the aggregate production function has been a source of controversy ever since Cobb and Douglas's (1928) original publication (McCombie, 1998). Nevertheless, Solow (1957, p. 312) in his pioneering paper on economic growth began with the famous introduction:

> In this day of rationally designed econometric studies and super-input-output tables, it takes something more than the usual "willing suspension of disbelief" to talk seriously of the aggregate production function. But the aggregate production function is only a little less legitimate a concept than, say, the aggregate consumption function, and for some kinds of long-run macro-models it is almost as indispensable as the latter is for the short-run. As long as we insist on practicing macro-economics we shall need aggregate relationships.

However, even a few moments consideration is sufficient to demonstrate how implausible is this view, even though it is now widely accepted. The aggregate production function, including the ubiquitous Cobb-Douglas and the more flexible Constant of Elasticity of Substitution (CES) [Arrow *et al.,* 1961] ideally should describe an aggregate relationship between output and the corresponding inputs measured in *physical* terms. This is a relationship that is essentially what may be described as an "engineering production function".

But the production function is universally specified in terms of constant-price *monetary* terms for both output (value added) and the capital stock. However, as we shall show, this undermines the whole concept of the aggregate production function, both theoretically and empirically. In many ways this argument can been seen as rather nihilistic, as it shows how tenuous (and indeed some would say meaningless) are the foundations of the supposed aggregate production function.

The problem of the generally uncritical acceptance of the aggregate production function, perhaps not surprisingly, starts with both the macroeconomics and the microeconomics textbooks. Students are introduced to a simple and uncritical example of a micro-production function in, for example, the introductory textbook of Mankiw and Taylor (2008, p. 64). Here, the authors give a specific example of a bakery. The output of the bakery is the number of loaves produced, the kitchen and the equipment are its capital and the workers are its labour. All these are physical concepts. A legerdemain occurs later in the book on page 72, where the *aggregate* production function for the whole economy is suddenly introduced. But, crucially, here the output and the capital input are all measured in *value* and not in *physical* terms, as in the bakery example. This change in measurement, important though it is, is not commented on in the textbook.

Clearly, a problem is that it is not possible to sum either the various outputs produced in the economy in physical terms or the physical capital inputs used in the production process. The same argument occurs for labour of different abilities and skills, but this is also generally ignored in the literature. However, the student is not usually made aware of the importance of this distinction between value and physical terms. The question also arises as to where are the intermediate inputs in the production function? In the above example, this includes the flour for the loaves. Using value data, it is conceptually a relatively straightforward matter to deduct the value of materials from the value of output to give value added. However, this procedure is not possible with physical data. But let us ignore this question as do most economics textbooks. Nevertheless, see Moseley (2015) for an incisive discussion of this shortcoming, especially with respect to some intermediate microeconomics textbooks, and its implications.

This slight of hand of using value data for physical data becomes important in our subsequent analysis and critique of the aggregate production function. Robinson (1970), in her review of Ferguson's *The Neoclassical Theory of Production and Distribution* (1969) and Ferguson (1971), in his reply, agreed (and this is about the only thing they did agree on) that theoretically production functions should be specified in heterogeneous *physical* terms. Cobb and Douglas (1928) make the same point at the beginning of their seminal paper. However, constant-price *value data* have to be used in any empirical analysis for measuring the output and the capital stock. Cobb and Douglas (1928) wrote that "The progressive refinement during the recent years in the measurement of the volume of *physical* production in manufacturing suggests the possibility of attempting (1) to measure the changes in the amount of labor and capital which have been used to turn out this volume of goods, and determine what relationships existed between the three factors of labor, capital, and product" (Cobb and Douglas, 1928, p. 139, italics added)[1].

The reason for the use of value data is straightforward. How do you aggregate, say, the capital inputs of factory floor space and forklift trucks in physical terms? As Robinson (1953-4) repeatedly pointed out, how would it be possible even start to include individual capital goods as separate physical measures (pints, tons, numbers) in a production function. It is thus a self-evident necessity to use value terms, *faute de mieux*, but it is this more than anything that undermines the concept of the aggregate (or indeed the firm) production function, as will be seen. Even if there were detailed figures in physical terms for the blueprints of the various outputs of an office or a factory, how would one even conceptually measure the degree of substitution between a very small increase in, say, the number of square meters of factory or office floor space with respect to a small change of employees?

In fact, what would probably be found is that with the increase in floor space, the number of employees would increase *pari passu* as the

---

1 Cobb and Douglas (1928) repeatedly refer to output being measured as a "physical volume of production" using Day and Person's (1920) figures. However, while Day used physical measures for individual manufacturing categories, these indices were aggregated using 1909 prices. Consequently, the aggregate index of production used is a constant-price value measure.

output increased. In other words, the physical relationship between the workers and the physical measure of office or factory space is likely to be one of fixed coefficients. However, the elasticity of substitution between information technology and, say, accountants is likely to be high. There have been only a very small number of studies of engineering production functions, but these have been for very narrowly defined processes, such as ploughing, wood working and oil transportation. See, for example, the survey of Wibe (1984).

If the whole economy is considered, what about the approximate 80 percent of Gross Domestic Product (GDP) where there is no *physical* measure of output at all? This is true of valuations of, for example, the output of government expenditure where the value of output is usually defined as the value of the inputs (wages and the cost of capital) with either no, or an arbitrary, allowance for productivity gains. There are also various methods employed to deflate the current price values, the procedures of which differ to a certain extent between countries. For a detailed consideration of measuring UK government output and its limitations, see Atkinson (2005).

Haldane, Brennan, and Maduros (2010) point to the insuperable problems of measuring the output of the finance sector. As they point out "According to the National Accounts, the nominal gross value-added (GVA) of the financial sector in the UK grew at the fastest pace on record in 2008Q4 (…) At a time when people believed banks were contributing the least to the economy since the 1930s, the National Accounts indicated that the financial sector was contributing the most since the mid-1980s" (Haldane, Brennan, and Maduros, 2010, p. 88).

The durability of the Cobb-Douglas (and the CES) aggregate production function as an empirical relationship is in itself surprising. The structures and methods of production have changed out of all recognition over the last hundred years or so. When Cobb and Douglas were writing in 1928, manufacturing, and indeed the economy itself, was dominated by heavy industry and industrial factories producing, for example, steel, motor vehicles, railways, and such consumer products as tobacco. These had all greatly expanded in the US in the so-called "Roaring Twenties". This period saw the shift from highly-skilled artisans to the factory system with increased mechanization. Today, production is dominated by such corporations as Netflix, Amazon, and Apple. There has been a revolution in the production process, with the rise of the importance of intangibles, artificial intelligence, the internet and the development of what has been termed *Capitalism Without Capital* (Haskel and Westlake, 2017). See also Corrado *et al.* (2022). Yet, the functional forms of the Cobb-Douglas and the CES are still used and paradoxically, and perhaps surprisingly to some, still give good statistical fits to the data. The reasons for this are discussed below.

The aggregate production function, while generally accepted by neoclassical macroeconomists, has had its serious critics in the past. These include those who have shown that the conditions for successful aggregation of micro-production functions are so restrictive as to make the aggregate production function a non-event. Then there were also the Cambridge Capital Theory Controversies between Cambridge, UK, and Cambridge, Massachusetts, that started in the early 1950s and more or less petered out in the mid-1970s. While Cambridge, UK, raised severe and largely unanswered criticisms about the aggregate production function, these are now subsequently ignored and largely forgotten by neoclassical macroeconomists (Birner, 2002; Cohen and Harcourt, 2003).

While these criticisms will be touched on, we shall outline what we see as a more fundamental critique that shows that the foundations of the aggregate production function are so tenuous as to make its use literally a nonsense. Given the widespread use, and indeed acceptability, of the aggregate production function in macroeconomics, it is appreciated that this may take some doing. But the argument is deceptively simple, and for reasons that will become quickly apparent, it is termed the "accounting identity critique" or simply the "accounting critique".

2. THE AGGREGATION PROBLEM AND THE CAMBRIDGE CAPITAL THEORY CONTROVERSIES

The major problem, as has been argued above, with the aggregate production function, is the implausibility of regarding a physical measure of aggregate output, at the industry or whole economy level, as being adequately proxied by a constant-price series of value added. The same goes for the proxying of the stock of physical capital by its constant price value calculated by using the perpetual inventory method. However, even if one does make this heroic assumption, there is still the problem as to whether or not production functions of identical measures of output, capital and labour can be aggregated to give a single aggregate production function. There is now a considerable literature on the aggregation problems that stretch back to the 1940s. (See Walters, 1963; Fisher, 1969; Felipe and Fisher, 2003; 2006; 2008). There is not space to discuss this technical literature here. The damaging implications of the aggregation problem have been usefully summarised by Fisher (2005, p. 490) as follows:

- Except under constant returns, aggregate production functions are unlikely to exist at all.
- Even under constant returns, the conditions for aggregation are so very stringent as to make the existence of aggregate production functions a non-event. This is true not only for the existence of the capital stock but also for 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.

It should be emphasised that these conclusions apply even when the outputs and inputs are assumed to be homogeneous and physical entities. Blaug (1974, pp. 5-18) also provides a succinct summary of the problems facing the aggregate production function in his commentary on the Cambridge Capital Theory Controversies. Blaug (1974, p. 17, omitting footnotes) sums up the position as follows: "The concept of an economically meaningful aggregate production function requires much stronger and much less plausible conditions than the concept of an aggregate consumption function[2]. And yet, undisturbed by Walter's [1963] conclusions or Fisher's [1992] findings, economists have gone on happily in increasing numbers estimating aggregate production functions of even more complexity, barely halting to justify their procedures or to

[2] Here Blaug differs from Solow (1957), as evidenced by Solow's quotation above.

explain the economic significance of their results." As long ago as 1963, Walters (p. 11) concluded that "After surveying the problems of aggregation one may easily doubt whether there is much point in employing such a concept as an aggregate production function".

A notable critic of the aggregate production function was Joan Robinson, but the accounting critique has nothing at all to do with the Cambridge Capital Theory Controversies. In fact, Robinson herself seemed later to play down the Controversies in, for example, "The Unimportance of Reswitching" (Robinson, 1975). Yet she saw the importance of an argument by Phelps Brown (1957) that formed one of earlier post-war expositions of the accounting critique and which is discussed below. Nevertheless, it is fair to say that Phelps Brown's paper, apart from leading to Robinson's comment, did not have any great immediate effect on the profession[3]. Likewise, Samuelson (1979, p. 929 and pp. 932-934), in a paper commemorating Douglas's work, also touched on the accounting critique, but did not take it to its logical conclusion. But probably the most notable critic at the time was Simon who thought it of such importance as worthy of mentioning it in his Nobel Prize lecture (1979a, p. 497).

Simon was, of course, a distinguished polymath and not a product of the neoclassical macroeconomics paradigm. He had read the early criticisms of the Cobb-Douglas production function including, especially, Reder (1943), Bronfenbrenner (1944), Marschak and Andrews (1944), Phelps Brown (1957) and Arrow *et al.,* (1961). In 1963 he was one of the first, with Levy, to prove mathematically that the Cobb-Douglas production function was nothing more than a transformation of a linear accounting identity (Simon and Levy, 1963). As such the production function says nothing about the "laws of production". Simon's (1979b) paper is still one of the best short introductions as to why neoclassical production functions, both the Cobb-Douglas and the CES, are spurious.

---

3 The entry on Phelps Brown in the *New Palgrave Dictionary of Economics* did not even cite the paper (Routh, 2008). Moreover, Phelps Brown (1996), himself, in an autobiographical note did not mention it. See also Falkus (1996) who also followed suit. An exception is Riach (2019, abstract) in his entry on Phelps Brown in the *Palgrave Companion to LSE Economics* which recognised that Phelps Brown "provided a devastating critique of the Cobb-Douglas production function".

Also of particular importance was a note by Shaikh (1974) criticising Solow's famous (1957) paper on applied economic growth and the aggregate production function. This paper of Solow is very much the empirical counterpart to his famous 1956 theoretical paper on neoclassical growth theory. Shaikh's note was somewhat polemical in style, with the title "Laws of Production and Laws of Algebra: The Humbug Production Function". Solow's (1974) response was to dismiss the arguments out of hand. This was a view reiterated in Solow's (1987) "Second Thoughts on Growth Theory". However, evidence is presented below that, perhaps remarkably, the arguments in neither of Solow's papers are particularly convincing nor stand up to carefully scrutiny. Shaikh was fundamentally correct in his criticisms of Solow's (1957) famous paper. This debate is discussed below.

3. THE ACCOUNTING IDENTITY CRITIQUE AND THE AGGREGATE PRODUCTION FUNCTION

The crucial flaw with the neoclassical aggregate production function is the necessity of using constant-price value data together with an accounting identity obtained from the National Income and Product Accounts (NIPA) or, say, the Census of Manufactures. This identity is given by the equation:

$$Q_{it} = w_{it} L_{it} + r_{it} K_{it} \quad [1]$$

Q is value added, w is the average wage rate, L is the labour input, r is the average rate of profit and K is the value of the capital stock. These data are readily available for GDP and industry and services at the 2-digit, and often, at the 3-digit and 4-digit Standard Industrial Classification level. Equation [1] is available for both different industries (denoted by i) and different years (t).

It is worth emphasising again that equation [1] is an *identity*, that is to say it is true by definition and it is not a behavioural relationship, *i.e.*, one that can be refuted or proved wrong. The fundamental problem is that a mathematical transformation of the accounting identity can be used, which makes no economic assumptions, to derive something that looks like a neoclassical aggregate production function, but which is clearly not.

For example, if this accounting equation for an industry is differentiated and then integrated, a relationship is obtained that is exactly the same as a Cobb-Douglas production function or, under certain well-known assumptions, resembles a CES production function. The latter is discussed below.

To see this, differentiating equation [1] for an industry with respect to time at a specific time, $t = t'$, gives the accounting identity in growth rates as:

$$\widehat{Q_t} \equiv a\widehat{w_t} + (1-a)\widehat{r_t} + a\widehat{L_t} + (1-a)\widehat{K_t} \qquad [2]$$

where a circumflex denotes an exponential growth rate. The parameters a and $(1 - a)$ are labour's and capital's share of output respectively, measured at t' and the growth rates are also evaluated at this time.

Equation [2] may be integrated and expressed equivalently as:

$$Q_t \equiv c_0 w_t^a r_t^{(1-a)} L_t^a K_t^{(1-a)} \qquad [3]$$

where $c_0 = a^{-a}(1-a)^{-(1-a)}$ and is the constant of integration. It should be noted that equation [3] is *not* an approximation for the accounting identity. It is an exact relationship, or isomorphism, that holds precisely for the base year when the differentiation is undertaken, *i.e.*, at $t = t'$. See McCombie (2011, pp. 176-177), who explicitly confirms this using UK total manufacturing statistical data. It is simply a "distribution function" rather than a "production function". As will be seen, equation [3] gives a very close approximation for other years. If the growth of the wage rate and the rate of profit, each weighted by its factor share (*i.e.*, $a\hat{w} + (1-a)\hat{r}$) is constant and equal to λ, then equation [3] may be written as:

$$Q_t \equiv Ae^{\lambda t} L_t^a K_t^{(1-a)} \qquad [4]$$

When the weighted growth rates of wages and the rate of profit are not constant, this gives:

$$Q_t \equiv Ae^{\int \lambda_t d_t} L_t^a K_t^{(1-a)} \qquad [5]$$

Consequently, what is termed by Solow as technological change is nothing more than, *by definition,* the growth of the wage rate and the rate of profit each weighted by its factor share. It does not necessarily have anything to do with the pace of technical change and to term it so, as Solow (1957) *inter alios* does, has led to a fundamental misunderstanding as to its interpretation.

The above argument follows through when cross-sectional data (either for individual firms, states or regions). The accounting identity is given for a firm by:

$$Q_i \equiv c_0 w_i^a r_i^{(1-a)} L_i^a K_i^{(1-a)} \qquad [6]$$

where c_0 is again the constant of integration. If the weighted wage rate and rate of profit do not vary greatly across firms, or at least compared with the variation in L_i and K_i, this gives:

$$Q_i \equiv A L_i^a K_i^{(1-a)} \qquad [7]$$

In other words, equation [7] is merely another way of writing the accounting identity.

More generally there is the more flexible CES and translog production function, which equally suffer from the same problem. These relationships tell us *nothing* about the underlying technology of the economy. This bold assertion will be justified below, but for the moment, the history of how this fundamental criticism developed and was ignored by nearly all neoclassical macroeconomists will be discussed.

As already mentioned, an early criticism of the Cobb-Douglas function was that of Phelps Brown (1957). Our argument extends Phelps Brown's approach. Nevertheless, Robinson (1970, p. 317) saw the potential importance of the argument. "It must have needed an even tougher hide to survive Phelps Brown's article entitled 'The Meaning of the Fitted Cobb-Douglas Function' than to ward off the Cambridge Criticism of the marginal productivity theory of distribution" (omitting a footnote). The most important point to note again is that all empirical specifications of the aggregate production function have, in practice, output measured in constant-price value added as well as the stock of capital measured in constant-price terms.

Phelps Brown's (1957) argument is relatively straightforward. He noted in a critical assessment of Cobb and Douglas's (1928) paper, constant-price value added is given by an accounting identity for a specific firm or industry. Alternatively, as we have noted above, the data may be given as a time series.

As has been seen, the data are summed arithmetically to give: $Q_i \equiv w_i L_i + r_i K_i$, where Q is value added, w is the average wage rate, L is the labour input, r is the rate of profit (a pure number) and K is the constant-price value of the capital stock. The subscript i denotes the ith firm. The Cobb-Douglas production function needs no introduction. For cross-sectional data for firms it is given by equation [7]. Increasing or decreasing returns putatively can be included by specifying $(1 - a)$ as b and determined by the values of $(a + b)$. But the underlying accounting identity given by equation [1], and hence as equation [3], shows the coefficients must by definition sum to unity.

If equation [1], the identity, is differentiated with respect to, say L, the following is obtained:

$$\frac{\partial Q_i}{\partial L_i} = w_i \qquad [8]$$

This, of course, is *not* the marginal product of labour, as it is derived from an identity.

Similarly, if equation [7], the putative Cobb-Douglas "production function" is partially differentiated with respect to L, the following is obtained:

$$\frac{\partial Q_i}{\partial L_i} = a_i \frac{Q_i}{L_i} \qquad [9]$$

which, from equation [8] equals w_i and it follows that the following must hold:

$$a_i = \left(\frac{w_i L_i}{Q_i} \right) \qquad [10]$$

It should be noted that the second-order derivatives of equations [1] and [7] are not the same, but as will be seen, this does not affect the argument.

As Phelps Brown (1957, p. 557) put it, “The Cobb-Douglas [a_i], and the share of earnings, will be only two sides of the same penny”.

The implicit argument here is that it is solely the accounting identity that is responsible for the apparent condition that the marginal productivity of labour equals the wage rate. The argument also applies equally when time-series data is used. Samuelson (1979, p. 929), in his tribute to Douglas mentioned that when, as a postgraduate student, he estimated the “inadmissible form” of $Q = b_1 L + b_2 K$ using Cobb and Douglas's data[4], he, not surprisingly, got the same value of the multiple correlation coefficient, namely $R = 0.97$, that Douglas got for the Cobb-Douglas form. Samuelson also commented that when the cross-industry specification that is given by the multiplicative form is estimated to determine whether the output elasticities sum to unity “that result tends to follow purely as a cross-sectional *tautology* base on the residual computation of the nonwage share” (Samuelson, 1979, p. 932, emphasis in the original). Surprisingly, Samuelson (1979) makes no criticism of Solow's (1957) paper (or others using the Cobb-Douglas production function) along these lines.

As has been mentioned above, there is a short note by Simon and Levy (1963) who considered the cross-industry production function given by $Q = AL^a K^{(1-a)}$, but supposed the true underlying production function was the simple accounting identity $Q = wL + rK$. Expanding the Cobb-Douglas function using a Taylor-series expansion, they show that the value of a in the multiplicative function given by the Cobb-Douglas function is given by $a = wL/Q$ when evaluated at the means of the linear function. They conclude, similarly to Phelps Brown, that: “Thus, the existence of a fitted Cobb-Douglas function with a value of [a] in agreement with the actual [a] does *not* imply that the underlying production function is truly Cobb-Douglas. In fact, we expect this agreement when the true function is given by [the linear accounting identity]” (Simon and Levy (1963, p. 94, emphasis in the original)[5].

The equivalence between the accounting identity and the Cobb-Doug-

---

4 Henceforth, for notational simplicity we omit the time (*t*) and industry (*i*) subscripts.

5 An alternative method is to use the approximation $\ln\left(X_i/\overline{X}\right) \approx \left(X_i/\overline{X}\right) - 1$ where *X* is the small deviation from its mean $\overline{X}$.

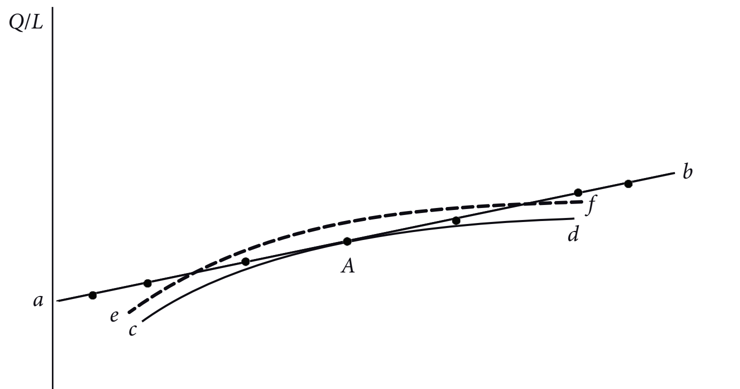
las may be best illustrated by Figure 1, where the linear accounting identity is depicted by the line *ab*. It is assumed, for expositional purposes, that the capital-labour ratio varies, but the wage rate and the rate of profit are constant. The corresponding equivalent Cobb-Douglas function is given by the curve *cd* and the point of tangency with the accounting identity is given by point *A*. Let us assume that the true relationship when the capital-labour ratio varies is given by the points along the linear accounting line *ab*. If we were to estimate a Cobb-Douglas function using these data, the best statistical fit would be given by the dotted line *ef*, which minimises the sum of the squared residuals.

An interesting question is that given the true relationship is the accounting identity, what is the size of the error in mistakenly fitting a Cobb-Douglas function? Following Simon (1979b), the degree of error in approximating the inter-firm Cobb-Douglas given by the curve *cd* to the linear data is given by:

$$\frac{Q'}{Q} = \frac{(AL^{a} K^{(1-a)})}{(wL + rK)} \qquad [11]$$

where Q' is the approximate calculated value of output given by the Cobb-Douglas function and Q is given by the linear accounting identity. Simon calculated the ratio of Q'/Q over a range where the capital-labour

Figure 1. The Cobb-Douglas approximation to the linear accounting identity



ratio varies from 5 to 0.2. Simon finds that even with this wide range of the capital/labour ratio, which is "wider than any encountered in the literature", the maximum error is less that 15 per cent, and for realistic values of the K/L ratio "we can expect average estimating errors of less than 5 percent" (Simon, 1979b, p. 466).

To summarise: The value of the output of the Cobb-Douglas and the (correct) linear accounting identity are identical at their point of tangency (*A* in Figure 1). But even though when we depart from this, the divergence in the actual and predicted value is extremely small. This is true when the erroneous Cobb-Douglas relationship is estimated using these data which give the best statistical fit (the dotted line given by *ef*).

4. ON SOLOW'S (1957) "TECHNICAL CHANGE AND THE AGGREGATE PRODUCTION FUNCTION"

As has been noted, an important paper that purported to show the empirical relevance of the Cobb-Douglas aggregate production function was that of Solow (1957). However, it will be shown that the analysis is far from convincing. Solow assumes output is given by an aggregate production function, with neutral technical change, specified as $Q = A(t)f(L,K)$, with the usual notation.

The two variables, Q and K, measured in constant prices, are seen implicitly as good proxies for their physical magnitudes, although, in view of the above discussion and the argument below, this presents major problems for Solow's argument. Solow's analysis begins by expressing the above equation in annual growth rates and he crucially assumes that the marginal productivity conditions hold (and, hence, there are constant returns to scale). This gives the Cobb-Douglas production function expressed in terms of growth rates as:

$$\widehat{Q} = \widehat{A} + a\widehat{L} + (1-a)\widehat{K} \qquad [12]^{6}$$

or, alternatively,

[6] We use exponential growth rates while, Solow uses $\Delta A/A = (A_t - A_{t-1})/A_{t-1}$. This makes little difference to the analysis.

$$\hat{q} = \hat{A} + (1-a)\hat{k} \quad [13]$$

where $\hat{q}$ is the growth of labour productivity (Q/L or q) and $\hat{k}$ is the growth of the capital-labour ratio (K/L or k).

The coefficient $(1 - a)$ is taken to be the value of the annual share of capital in output. Solow's data are the various components of US private non-farm GDP from 1909 to 1942, taken from a variety of sources.

From equation [13], the annual rate of supposed technical progress is calculated as:

$$\hat{A} = \hat{q} - (1-a)\hat{k} \quad [14]$$

Given data for Q, $(1 - a)$, K and L, Solow calculates the annual growth rate $\hat{A}$ for the U.S. from 1909 to 1949.[7] The annual rate of putative technical change $\hat{A}$ is shown in Solow's Chart 2 (1957, p. 314). Solow notes from this that the annual growth rate of technical change $\hat{A}$, is "essentially constant over time, exhibiting more or less random fluctuation about a fixed mean" (p. 316). Solow also calculates the series for the levels of A using $A_t = A_{t-1} + \Delta A_t / A_{t-1}$ and setting A_0 equal to unity.

Hence, Solow "corrects" Q/L (or q) for the impact of technical change to give q/A. Using these time-series data, Solow estimates a number of different functional forms, with the Cobb-Douglas and the semi-logarithmic specification, performing the best. The estimates of the Cobb-Douglas are:

$$\ln\frac{q}{A} = \underset{(-261.33)}{-0.78} + \underset{(122.70)}{0.353}\ln k \qquad \overline{R^2} = 0.977$$

This is the estimate of Solow's (1957, p. 318) equation [4d], using Hogan's corrected data. It goes without saying that the goodness of fit is remarkably close, especially as this is supposed to be such a complex concept as an aggregate production function.

---

7 There is an error in Solow's data for the years 1943-1949, which Solow excludes from his statistical analysis. (See Hogan, 1958, on this). Hogan's corrected values were included for this period in the data set. The rate of unemployment is used to adjust the capital-labour ratio for changes in capacity utilisation, but this makes little difference to the argument.

The mean value of capital's annual shares, (1 – a), is 0.341 with a standard deviation of 0.0174, but as Solow admits, the closeness of this to the estimated coefficient of ln$k$ cannot be taken to confirm, or rather not to refute, the marginal productivity theory of factor pricing. The usual well-known neoclassical argument is that if the freely estimated output elasticities are not statistically significantly different from the factor shares, this demonstrates that factors are paid their marginal products. But in this case, the values of capital's share have been used in the calculation of equation [14].

5. HOGAN (1958) AND SHAIKH (1974) RAISE SOME IMPORTANT QUESTIONS

An early important, but largely overlooked comment, was that of Hogan (1958). Hogan started from Solow's premise that the theory underlying Solow's analysis was an aggregate production function.

However, Hogan noted that as the annual values of $\widehat{A}$ are *calculated* by Solow as $\hat{q}-(1-a)\hat{k}$, the equation will give an exact result for any particular year.

If we integrate the growth rate equation $\hat{q}=\widehat{A}+(1-a)\hat{k}$, we get $q=c_0Ak^{(1-a)}$ where c_0 is again the constant of integration. As Hogan (1958, p. 410) points out "There is no room for error so long as (1 – a) is constant."

The implications are that:

> In the first place the estimate of "technical progress" seems to be of little value as it incorporates the effects of any occurrence, whether chance or otherwise, upsetting a relationship between capital and output. And this relationship is determined solely by the share of income accruing to the capital stock. Secondly, the close fit of the data stems from the nature of the model itself (Hogan, 1958, p. 410, excluding a footnote).

Hogan (1958, p. 410) considers that "this production function does not appear to have any real meaning". The reason is that it is simply an identity. Solow (1958, p. 411) in his reply merely states that "I don't want to argue here the case for and against aggregate production functions (with or without marginal productivity)", but it is clear Solow's exercise assumes an aggregate production function, as he admits. Solow concedes that his method is tautological "but not all tautology is bad". This seemingly oxymoron may be explained as follows. Equation [14] is, by definition, a tautology for a single value of a growth rate of *A* calculated between two different points in time. (This is annual data in Solow's case, but the initial and terminal years could be further apart.)

Consequently, it is possible to *calculate* the rate of "technical progress" using any pair of observations, no matter how far apart in time. This is because the (average) share of capital is by definition constant over the period of calculation. But Solow points out that the *annual* values of the share of profits are not precisely constant over time, although they are almost so, ranging from 0.312 to 0.397.

Thus, the data *must* give a very close fit to Solow's supposed Cobb-Douglas production function. Nevertheless, it is theoretically possible for the estimate of the capital's share to fluctuate annually to such an extent that the regression gives a very poor statistical fit. Hence, although a tautology for each year, it may not be when several years are used in a regression. According to this argument, the latter does tell us something empirical about the rate of technical progress. In Solow's case, as has been seen, the share of profits is almost a constant so the regression is presumably of a near, or a "good", identity.

Shaikh (1974), as was noted above, made an important contribution to this critique by confirming empirically, and somewhat dramatically, that the estimates of the Cobb-Douglas production function are nothing more than those of the transformation of an accounting identity, as has been seen above.

The accounting identity may be expressed in growth rate form as:

$$\hat{q} \equiv a\hat{w} + (1-a)\hat{r} + (1-a)\hat{k} \quad [15]$$

A comparison with equation [13] shows that the rate of "technical progress" is definitionally equal to the sum of the growth of the wage rate and the rate of profit. Equations [12] and [13] are simply the accounting, or distributional, identity expressed in growth rates.

Thus, the analysis can tell us nothing about the underlying technological structure of the economy. Moreover, this explains why estimates of the productions function give what may be seen as close statistical fits even though aggregation problems suggest that they do not actually exist. See Fisher (1992), *inter alios.*

Shaikh (1974) also provides an important illustration about the problems with Solow's approach. He constructs a hypothetical data set where the output-input data (the relationship between q and k) when plotted on a graph spells out the word HUMBUG.

The value of capital's share is the same as in Solow's analysis and, as has been seen, nearly constant. Following Solow's procedure, Shaikh (1974, p. 118) finds that estimating a putative Cobb-Douglas production function gives, not surprisingly in the light of the accounting critique, a very close statistical fit to his HUMBUG data[8]:

$$\ln\frac{q}{A} = -0.453 + 0.34\ln k \qquad R = 0.9964$$

As Shaikh puts it: "Humbug data can be extremely well represented by a Cobb-Douglas production function having constant returns to scale, neutral technical progress, and marginal products equal to factor rewards" (Shaikh, 1974, p. 118).

6. SOLOW (1974) STRIKES BACK

Solow's one-page rejoinder is similar to his riposte to Hogan (1958), but far more polemical. He starts: "Mr Shaikh's article is based on misconceptions pure and simple" (Solow, 1974, p.121). Reflecting his reply to Hogan, Solow argues that his analysis was in no sense a *test* of the aggregate production function or of the marginal productivity theory. According to Solow, the only purpose of his paper was to confirm that if the share of capital was roughly constant over time, the result must be a well-determined Cobb-Douglas aggregate production function. It was also a method to estimate the rate of technical change. But whether or not this is the case can be seen from the variation in the annual data for capital's share. There is little point in estimating the tautological equation

[8] Shaikh does not report the usual diagnostic statistics such as the t-values of the estimated coefficients.

used by Solow. Although it has been argued that the calculated path of "technical progress" over the period is not without interest, Solow nowhere mentions that this simply must equal weighted growth of the wage rate and the rate of profit.

Solow (1974) cites Fisher's (1971) simulation paper (although this hardly supports Solow's argument), as ironically does Shaikh (1974, p. 116). Fisher constructed a simulation exercise where the conditions for successful aggregation were violated, but found that the Cobb-Douglas production function gave a good fit to the data. He concluded that this was because the factor shares were constant and later conceded the explanation was due the underlying accounting identity. In other words, it is the constancy of the factor shares that gave rise to the Cobb-Douglas production function and not *vice versa*.

Solow's putative *coup de grâce* was that when Shaikh's data are used to estimate the Cobb-Douglas production function with a *linear* time trend instead of $\widehat{A}$, although, as already noted, the latter fluctuates cyclically around a trend, there is no statistically significant relationship. The results reported by Solow (1974) using Shaikh's HUMBUG data are:

$$\ln q = -0.14090 + 0.00532t - 0.3307 \ln k \qquad R^2 = 0.0052$$
$$(0.52072) \quad (0.01246) \quad (0.76098)$$

where the figures in brackets are the standard errors.

Consequently, there is no statistically significant relationship using the Humbug data. According to Solow (1974, p. 121), "If this were the typical outcome with real data we would not now be having this discussion. The humbug seems to be on the other foot". He thus dismisses Shaikh's critique out of hand.

However, Solow surprisingly omitted to estimate (or report) the same specification using the US "real world" data that he used in Solow (1957). When this is done, the results are (where the t-values are in parentheses):

$$\ln q = -0.435 + 0.019t - 0.081 \ln k \qquad \bar{R}^2 = 0.971$$
$$(-9.01) \quad (32.07) \quad (-1.46)$$

According to this regression result, the use of Solow's data comprehensively refutes the hypothesis that the equation gives a good statistical fit to the relationship between ln$q$ and ln$k$. But this also raises the question, so why are we getting these seemingly perverse results for estimating a putative identity?

Of course, according to Solow there is nothing in neoclassical theory that stipulates technical progress (*i.e.*, the growth of the wage rate and rate of profit weighted by their factor share) occurs at a constant rate, as Solow's (1957, Chart 2, p. 314) confirms. So both the above regressions of the identity are misspecified. As already noted, studies suggest that the weighted growth of the wage rate and the rate of profit varies procyclically with output [as does Solow's (1957) estimate of $\widehat{A}$]. Consequently, using a linear, rather than a more complex, time trend explains why the identity results break down.

As Shaikh (1980) and others have shown, it is always possible to find a sufficiently complex time trend that closely proxies the weighted growth of wages and the rate of profit and ensures that the data will give a perfect fit to the production function, or rather the identity. On whose foot the humbug is on is very much a moot point.

7. SOLOW HAS SECOND (OR FURTHER) THOUGHTS

Solow (1987) returned to a consideration of the accounting critique, and especially Shaikh's (1974; 1980) expositions of it. Solow presented some new arguments that Shaikh's analysis was flawed (McCombie, 2001). The reason was that although Shaikh's argument was directed at the supposed "aggregate production function" estimated using value data, according to Solow, the analysis could be applied "word for word" to the estimation of a micro-production function. This latter is assumed to have a single physical output and well-defined inputs of different types of labour and individual physical capital goods. We also are assumed to know separately the individual prices for the physical outputs and inputs. As was noted above, these assumptions are rather implausible, even for the simplest of micro-production functions, which would, in practice, normally use value data.

Solow (1987, p. 19) remarks that "In this case, the Shaikh argument asserts something truly remarkable", so much that it turns out to be "simply wrong".

> The first thing I want to point out is that it is not an argument about aggregate production functions, but about [micro] production functions. It is, of course, the habit of estimating and interpreting aggregate production functions that Shaikh is attacking; but the argument would apply word for word if [Q^*] were a single physically well defined output and [L^*] and [K^*] were an exhaustive list of physically well-defined inputs with [P_L] and [P_K] their prices in output units. In this case the Shaikh arguments assert something truly remarkable (Solow, 1987, p. 19)[9].

In order to demonstrate this, Solow first repeats the accounting critique using value data. As discussed above, this is that the identity, where, with the usual notation, $Q \equiv wL + rK$, is formally equivalent to $\ln Q \equiv [a \ln w + (1-a)\ln r] + a \ln L + (1-a)\ln K \equiv A(t)L^a K^{(1-a)}$.[10] Or to put this another way, the use of value data must give a near perfect fit to the Cobb-Douglas function, not because it reflects the underlying technology, but because it is merely estimating a transformation of the linear accounting identity.

It will be recalled that the linear and the power (Cobb-Douglas) functional forms of the identity are exactly equivalent at the point of evaluation. If we are using time-series data and factor shares change with time, as will be seen below, a more flexible function form for the linear identity will need to be used. Nevertheless, the central tenet of Shaikh's critique, as has been demonstrated, is that with value data, Solow's specification of the Cobb-Douglas function is merely estimating an underlying identity and says nothing about the "laws of production" and the technology.

Let us consider next Solow's analysis when physical data are used in the case of a micro-production function. In this case, Solow is correct in that the estimation is of the underlying technology and this is not subject to the accounting identity critique. To demonstrate this, as noted above, the strong assumption is necessarily made that there is only

9 Solow uses the notation of *p*, *u*, *q*, *v*, *x* instead of *w*, *L*, *r*, *K* and *Q*. The reason is that Solow (1987, p. 18) does "not want you to be worrying about things like the meaning of 'capital' or 'labour' for that matter, because the argument does not involve such issues in any way". We find, however, that his notation merely obscures the argument.

10 For expositional ease, the constant of integration is ignored.

one homogeneous physical output, Q^*, and this is measured as, say, the number of "widgets". The physical capital good, K^*, is measured as the number of "leets" and L^* is the number of identical workers. The star on the variable indicates that it is measured in physical terms. (L is equal to L^* as it is, say, the number of workers.)

A micro-production function can be estimated using only these *physical* measures, *i.e.*, $Q^* = f(L^*, K^*)$. This is immune from the accounting identity critique, which only holds for the use of value data[11]. But to be consistent, the specification of the physical production function should only include physical magnitudes.

However, we have no *a priori* idea as to what are the values of the coefficients of the physical micro-production function or, indeed, whether or not they are statistically significant. We need to statistically estimate the micro-production function. However, *pace* Solow, this provides no support for the argument that only using value, and not physical, data is merely picking up an accounting identity is erroneous.

So how is it that Solow, on the contrary, states that Shaikh's argument holds also for the use of physical data and therefore cannot be meaningful?

The argument is that the accounting identity using the physical terms is now given by $p_Q Q^* \equiv p_L L^* + p_K K^*$. p_Q, p_L and p_K are the monetary price of a widget, the wage rate and the rental price of capital. Solow takes p_Q as the numeraire and so it is equal to unity. Consequently, the accounting identity becomes $Q^* \equiv P_L L^* + P_K K^*$ where P_L and P_K are the price of labour and the price of homogeneous physical capital (leets), both measured in terms of units of physical output.

Let us next consider what Solow considers to be Shaikh's mistake. Solow's argument is couched in terms of physical data, but, as noted above, where we also have the individual prices.

Differentiating and then integrating $Q^* \equiv P_L L^* + P_K K^*$ gives:

$$Q^* \equiv P_L^{\theta} P_K^{(1-\theta)} L^{*\theta} K^{*(1-\theta)} \qquad [16]^{12}$$

[11] However, as Solow notes, a physical production function should be specified as having a list of different physical capital goods, as these cannot be aggregated into a single magnitude. This means it is problematic as to whether a micro-production function can ever be specified in physical terms. However, we shall ignore this problem, important though it is.

[12] The constant of integration is again ignored for expositional ease.

where θ equals labour's share of the physical output and (1 – θ) is capital's share.

From the above, the expression $P_L^{\theta}P_K^{(1-\theta)}$ equals $Q^* L^{*-\theta}K^{*-(1-\theta)}$. Substituting these expressions into the equation [16], Solow (1987, p. 20) obtains:

$$Q^* \equiv \left(\frac{Q^*}{L^{*\Theta}K^{*(1-\Theta)}}\right)L^{*\Theta}K^{*(1-\Theta)} \equiv f(L^*, K^*) \qquad [17]$$

Solow argues that this is essentially Shaikh's argument, but using physical values for output and capital rather than constant price values. Estimating $Q^* \equiv P_L^{\theta}P_K^{(1-\theta)}L^{*\theta}K^{*(1-\theta)}$ "with any reasonable flexible function of time is bound to give back a Cobb-Douglas function whose elasticities mimic the observed relative shares" (Solow, 1987, p. 19).

Solow (1987, p.20) argues from equation [17] that "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; the something else is the production function divided by the Cobb-Douglas". The implication of Solow's argument seems to be that "the something else" is as yet some undefined unique production function specified in *physical* terms.

But Shaikh's (1974; 1980) argument concerning the identity is couched in terms of the use of value data, which is used universally in the estimation of "production functions". This makes a great deal of difference to the interpretation of the argument.

The estimation of the micro-production function uses a unique set of physical data, namely, Q^*, L^* and K^*. The accounting identity of the micro-production function can be expressed as $p_Q Q^* \equiv p_L L^* + p_K K^*$ or equivalently as $Q^* \equiv P_L L^* + P_K K^*$ and is simply a *distributional relationship*. There are numerous combinations of prices in terms of output that satisfy this equation, but the economist in practice only knows the data with one set of prices.

In other words, Q equals $p_Q Q^*$ or output measured in constant-price monetary terms, K equals the cumulation of net gross fixed capital over time, or the constant-price of the value of capital stock. Our economist knows only the constant price value of output and capital, not their physical magnitudes.

Following Solow's reasoning above, but using value data, the comparable equation obtained is:

$$Q \equiv \left(\frac{Q}{L^a K^{(1-a)}} \right) L^a K^{(1-a)} \equiv f(w,r,L,K) \qquad [18]$$

The interpretation of equation [18] is that the accounting identity expressed in value terms as a Cobb Douglas is tautologically equal to itself, multiplied by $L^a K^{(1-a)}$ and divided by $L^a K^{(1-a)}$. In other words, with this procedure of Solow, Shaikh's argument has merely shown that the unique identity in value terms simply equals the identity given by the accounting identity in value terms. It follows that equation [18] $Q = f(w,r,L,K)$ is simply the accounting identity when value data are used. If a supposed aggregate production function is estimated using constant-price monetary values, as do all empirical estimations, Shaikh's analysis is correct. All that Solow's (1957) specification of the aggregate production function is of an accounting identity measured in monetary terms.

But Shaikh's argument does not apply when only physical data are used in estimating a micro-production function. Consequently, Solow's argument that Shaikh's critique does apply is erroneous. Shaikh's analysis would only apply if $p_Q Q^*$, $p_L L^*$ and $p_K K^*$ were included in the regression instead of Q^*, L^* and K^*. But the estimation of physical production function should only use Q^*, L^* and K^* as the relevant variables, which Solow assumes to be known. Prices should not be included in the estimation of a physical micro-production function. Consequently, Shaikh's criticism does not apply to the estimation of micro-production functions using only physical data, *pace* Solow[13].

8. SOLOW AND GROWTH ACCOUNTING

From his data and estimations, Solow (1957, p. 320) also argued that "Gross output per man hour doubled over the interval, with 87½ per cent of the increase attributable to technical change and the remaining 12½ per cent to increased use of capital". In his Nobel Prize lecture, Solow

[13] There should be no measure of "technical progress" when physical data are used, due to the law of the conservation of matter. It merely represents measurement error of one or more of the physical variables or a reduction in the lack of efficiency in the so-called micro-production function.

(1988, p. 313) found these results "startling". He continued; "I think I had expected to find a larger role for straightforward capital accumulation, than I actually found."

However, If Kaldor's stylised fact that the rate of growth of output is approximately equal to the growth of capital is used, the accounting identity may be expressed in growth rates as $\hat{A} \equiv a\hat{w} + (1-a)\hat{r} \equiv a\hat{q}$. Given that labour's share is about 0.75, this implies that that the rate of technical progress (or more accurately the weighted growth of the real wage rate and the rate of profit) accounts for about three-quarters of productivity growth. This result is simply due to the accounting identity. This also explains Solow's similar startling results mentioned above; but the only surprising thing is that anybody should find them surprising. Moreover, Solow must have been fully aware from the interchanges with Hogan (1958) and Shaikh (1974) that the rate of technical change was definitionally equal to the weighted growth of the wage rate and the rate of profit. However, this was never prominently mentioned by Solow.

9. ON THE CES "AGGREGATE PRODUCTION FUNCTION"

If time-series data are used and factor shares are changing, then the Cobb-Douglas will not necessarily give the best statistical fit. Arrow *et al.* (1961) found that using time-series country data and regressing lnQ/L on ln$w$, the estimated coefficient was significantly less than unity. As a result of this, Arrow *et al.* (1961), by using the marginal productivity condition, derived a more flexible function, namely the Constant Elasticity of Substitution (CES) production function. This is also subject to the accounting critique (Felipe and McCombie, 2001).

The CES is now generally specified with biased technical change as:

$$Q = C(a_0(A_L L)^{-\rho} + (1-a_0)(A_K K)^{-\rho})^{-\frac{1}{\rho}} \qquad [19]$$

where a_0 and $(1 - a_0)$ are the shares of labour and capital in the base, or some other particular year. In this specification A_L and A_K are (misleadingly) interpretated as labour-saving and capital-saving technical change. C is a constant. The "elasticity of substitution" is given by $\sigma = 1/(1+\rho)$.

The question arises as to how is this merely a transformation of the accounting identity? At time $t = t'$, $Q \equiv wL + rK$ not only definitionally equals equation [18], but the latter is derived from it under certain assumptions as t varies.

If the orthodox approach is followed, but it is assumed that $A_L = w$ and $A_K = r$ and L'Hôpital's rule is used for equation [19], the following result is obtained that when $\rho \to 0$:

$$Q_t = Cw_t^{a_t} r_t^{(1-a_t)} L_t^{a_t} K_t^{(1-a_t)} \qquad [20]$$

In other words, over time the factor shares are changing and so the CES as a single equation may give a better approximation in determining the level of output than the Cobb-Douglas. But at any particular time t, the Cobb-Douglas also gives a good approximation. And as has been seen, at any time the correct relationship is the linear accounting identity with factor shares equal to a_t and $(1 - a_t)$, or at the risk of repetition, given by $Q_t \equiv w_t L_t + r_t K_t$.

One way of demonstrating the correspondence between these three different functional forms without using the marginal productivity condition, or assuming that an underlying production function exists, is simply to use a Box-Cox mathematical transformation (Felipe and McCombie, 2011, pp. 285-286.) It should be emphasised that this makes no economic assumptions.

The Box-Cox is given by the following transformations:

$$X^{(\eta)} = \frac{X^{\eta} - 1}{\eta} \text{ when } \eta \neq 0 \qquad [21a]$$

If the equation does not have an intercept then this becomes $X^{(\eta)} = X^{\eta} / \eta$. This is sometimes termed the Tukey transformation.

$$X^{(\eta)} = \ln(X) \text{ when } \eta \to 0 \qquad [21b]$$

$$X^{(\eta)} = X \text{ when } \eta = 1 \qquad [21c]$$

The fundamental equation is given by the accounting identity, namely, $Q \equiv wL + rK$.

Alternatively, $Q^{\eta} = c + b_1 L^{\eta} + b_2 K^{\eta}$ will give the linear accounting identity when $\eta = 1$ and there is no intercept.

The specification using the transformation [21b] is the Cobb-Douglas expressed as $Q = c(wL)^a (rK)^{(1-a)}$ at the point where it is tangential to the linear accounting identity. This is given by the transformation [21b] (see also Simon and Levy, 1963).

The specification using the transformation [21a] is:

$$Q^{\eta} = c(b_1(wL)^{\eta} + b_2(rK)^{\eta}) \qquad [22]$$

which, when η equals –ρ is equal to the CES specification of the accounting identity, namely:

$$Q = c(a_0(A_L L)^{-\rho} + (1-a_0)(A_L L)^{-\rho})^{-1/\rho} \qquad [23]$$

where $b_1 = a_0$ and $b_2 = (1 - a_0)$. It can be seen that λ_L and λ_K, interpreted in equation [19] as the rate of labour and capital-augmenting technical change, are simply the wage rate and the rate of profit.

And using L'Hôpital's rule again, it is found that:

$$Q = c(wL)^{a}(rK)^{(1-a)} \qquad [24]$$

For an analysis that gives similar results see Simon (1979b, Section V, pp. 466-467) which refers to the interfirm CES. This uses a different method to show how the CES is nothing more than a special case of the accounting identity. Simon shows that as ρ goes to zero, the CES function goes to the Cobb-Douglas. Simon (1979b, p. 467) concludes:

> Thus, for the [CES] function as for the Cobb-Douglas function, the fact that the observed labor and capital shares are close to the values required by the theory is readily explained as an artifact, without the need for marginalist assumptions; the alternative assumption is simply that the data being fitted really represent the relation [given by the accounting identity].

In Section VI of his paper Simon (1979b) considers the case of time-series data, although he introduces the assumption that there is a constant capital-output ratio. He derives a less general analysis to the one in this paper, but nevertheless confirms that all that the Cobb-Douglas and the CES function are merely reflecting the underlying accounting identity.

McCombie and Dixon (1991) present some empirical results confirming this. They considered the estimates of Whiteman (1988), who used the dual of translog production function to estimate the supposed rate of growth of labour and capital efficiency, namely, $\hat{A}_L$ and $\hat{A}_K$. The accounting identity implies that $\hat{A}_L$ and $\hat{A}_K$ should equal $\hat{w}$ and $\hat{r}$, respectively, and this is what McCombie and Dixon found empirically. The results showed that growth of labour efficiency was 3.78% per annum (with a t-statistic of 11.29) while the growth of the real wage from the estimation of a linear time trend was 3.76 % per annum with a t-statistic of 55.26. The rate of growth of capital efficiency was 0.78 % per annum (t-statistic 2.30) and the growth of the rate of profit was 0.78 % per annum (t-statistic of 8.59). These results suggest that the estimates of the translog production function were merely capturing the transformation of the accounting identity.

Whiteman also reported similar results for 34 manufacturing industries. In all cases there was no significant difference between the growth of labour efficiency, $\hat{A}_L$, and the growth of the real wage. The growth of capital efficiency, $\hat{A}_K$, and the growth of the rate of profit were also very similar. But the accounting critique would lead us to expect this result, *a priori.*

10. SOME FURTHER SIMULATION RESULTS SHOWING WHY THE COBB-DOUGLAS GIVES SUCH GOOD STATISTICAL FITS

McCombie (2001) and in a related paper, Felipe and McCombie (2006), used some simple simulated data and regression analysis to demonstrate the crucial difference between using physical and value data. The example used here is taken from Felipe and McCombie (2006, pp. 290-292).

The simulated *physical* data was given for 10 firms and used in the *cross-industry* Cobb-Douglas production function:

$$Q \quad AL^{a}K^{(1 \quad a)} \qquad [25]$$

where A for each firm was taken to be unity, and $K = Q^{1/(1-\alpha)}L^{\alpha/(1-\alpha)}$.[14] The parameters α and $(1 - \alpha)$ took the values 0.25 and 0.75, so the actual

[14] Strictly speaking, there are problems with this physical specification noted above. Nevertheless, it is useful for expositional purposes.

production function was $Q = AL^{0.25}K^{0.75}$. The values 0.25 and 0.75 were deliberately chosen to be in marked contradistinction to the usual values of 0.75 and 0.25 that are found empirically in, for example, the National Income and Product Accounts.

For value data, a simple mark-up model was used given by:

$$p = (1+\pi)\left(\frac{wL}{Q}\right) \qquad [26]^{15}$$

where p is the price of a unit of physical output and π is the markup and takes a value of 1/3. This gives labour's share as $a = \left(wL/pQ\right) = 0.75$ and capital's share as 0.25. The value of capital is given by $K = (pQ - wL)/r$, where r was taken to be 0.10. The markup theory of pricing has received considerable empirical support ever since the pioneering study of Hall and Hitch (1939). This is summarized in Wilson and Andrews (1951). See also Coutts and Norman (2013) for a survey of the literature.

The data for the physical magnitudes given by Q and L were used in the construction of the value data which was used in the estimation. It should be noted that a small random error term was added to the data where it is necessary to prevent per multicollinearity. Using the constant-price value data to estimate the cross-firm production function gives the results:

$$\begin{array}{llll} \ln Q = & 2.867 & +0.750 \ln L + & 0.25 \ln K \qquad \overline{R}^2 = 0.999 \\ & (478.77) & (136.40) & (45.41) \end{array}$$

This gives a remarkably close fit to the Cobb-Douglas production function, which is not surprising given the method used to construct the data. The fact that the estimated coefficients equal the factor shares does not mean factors are paid their marginal products. The "true" output elasticities given by the use of physical data are 0.25 (and not 0.75) for labour and 0.75 (and not 0.25) for capital. However, when value data are used, and the data constructed using mark-up pricing, the values are 0.75 and 0.25.

---

[15] More generally, it is given by $p = (1+\pi)(wL + mM)/Q$ where m is the price of materials and M the volume of materials.

If there are true increasing returns to scale and physical data, when the value measures are calculated using a constant mark-up as above, we still obtain estimates of the output elasticities equal to their factor shares. It is the size of the mark-up that is solely responsible for generating the spurious results of the output elasticities when the Cobb-Douglas is estimated using value data.

A subsequent exercise generated the production function using the physical series that were random variables. In other words, there was no well-defined underlying production function in physical terms. The values for the output and capital stock in monetary terms were generated as before using the markup. The estimation using the value data again gave a very good fit to the Cobb-Douglas with the resulting values of the "output elasticities" of labour and capital as 0.75 and 0.25. The use of the "true" physical data, *per se*, gave a statistically insignificant relationship. This does not mean to say that output is actually a random function of the inputs. It may simply be due to the fact that any production function with many different types of physical inputs are much more complex than the simple Cobb-Douglas. Thus, the seeming randomness of the inputs may simply be a reflection of the severe error inherent in the specification of the micro-production function as a Cobb-Douglas.

However, it is important to note that even when there is no well-defined firm production function, as in this example, the use of value added data and a constant price value for capital will give the impression that a well-behaved aggregate Cobb-Douglas production exists, which could not be further from the truth.

McCombie (2001) and Felipe and McCombie (2006) extend the analysis to consider time-series estimations of the Cobb-Douglas with technical change, and increasing returns to scale. This made no difference to the estimated outcomes. As we have noted, Fisher (1971, p. 325), at the conclusion of his simulation study where the Cobb-Douglas production theoretically should not have existed but nevertheless gave a good fit to his simulation data, stated that it was caused by factor shares being constant. It was not that the production function gave rise to the constant shares.

The results here suggest that one answer lies in the constancy of the firm's mar-kup. Likewise, if the mark-up and hence labour's share changes over time, because of deregulation of labour markets, globalization and increased competition from abroad, this may be the reason why the CES provides a better approximation to the data than the Cobb-Douglas. But this has absolutely nothing to do with the existence of an aggregate production function.

Finally, Felipe, McCombie, and Mehta (2024) have updated and extended this work. Using simulated data from a cross-sectional Cobb-Douglas production function in physical terms from which they generate the corresponding series in monetary values, they show that the coefficients of labour and capital derived from the monetary series are: (*a*) biased relative to the elasticities by simultaneity and by the error that results from proxying physical output and capital with their monetary values; and (*b*) biased relative to the factor shares in value added as a result of a peculiar form of omitted variables bias. They show what these biases are and conclude that estimates of production functions obtained using monetary values are likely to be closer to the factor shares than to the factor elasticities. An alternative simulation that does not assume the existence of a physical production function confirms that estimates from the value data series will converge to the factor shares when cross-sectional variation in the factor prices is small. This is, again, the result of the fact that the estimated relationship is an approximation to the distributional accounting identity.

11. CONCLUSIONS

The argument here is that for well over a century a theory that is fundamentally flawed has been at the core of one branch of economics, namely neoclassical macroeconomics.

The aggregate production function which can only be estimated using value data, in fact reveals nothing about the "laws of production". The remarkable statistical fits of the aggregate production function are the result of using value data. The aggregate production function is nothing more than a mathematical transformation of an accounting identity. It is this that gives it good statistical fits when value data are used for output and the capital stock and which has led to the uncritical use of the aggregate production function for nearly a hundred years.

Douglas (1976, p. 914) summed up the implications of estimating the aggregate production function as follows. "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". However, no such conclusions can be drawn from the statistical estimates of putative production functions using constant-price value data, even when used as a proxy for physical data. The critique of the estimation of the aggregate production functions provides a clear-cut refutation of Friedman's (1953) instrumental methodology. This is that the realism or otherwise of the assumptions of a theory is irrelevant. All that matters is the statistical goodness of fit and the predictive power of the model. The estimation of the aggregate production suggests otherwise (see Felipe and McCombie, 2013, for a general discussion.)

But let us reiterate a word of warning from Fisher (2005) in his paper, "Aggregate Production Functions - A Pervasive but Unpersuasive, Fairytale": "I am informed (by Jesus Felipe) that attempts to explain the impossibility of using aggregate production functions in practice are often met with great hostility, even outright anger. To that I say (…) that the moral is: Don't interfere with fairytales if you want to live happily ever after". ◄

REFERENCES

Arrow, K.J., Chenery, H.B., Minhas, B.H., and Solow, R.M. (1961). Capital-Labor Substitution and Economic Efficiency. *Review of Economics and Statistics*, 243(3), 225-250. http://dx.doi.org/10.2307/1927286

Atkinson, A.B. (2005). *The Atkinson Review: Final Report. Measurement of Government Output and Productivity for the National Accounts*. Basingstoke, England: Palgrave Macmillan.

Birner, J. (2002). *Cambridge Controversies in Capital Theory. A Methodological Analysis*. London and New York: Routledge. http://dx.doi.org/10.4324/9780203416686

Blaug, M. (1974). *The Cambridge Revolution, Success or Failure? A Critical Analysis of Cambridge Theories of Value and Distribution*. London: Institute of Economic Affairs.

Bronfenbrenner, M. (1944). Production Functions: Cobb-Douglas, Interfirm, Intrafirm. *Econometrica,* 12(1), 35-44. http://dx.doi.org/10.2307/1905566

Cobb, C.W., and Douglas, P.H. (1928). A Theory of Production. *American Economic Review* (Supplement), 18(1), 139-165. https://www.jstor.org/stable/1811556

Cohen, A.J., and Harcourt, G.C. (2003). Retrospectives: Whatever Happened to the Cambridge Capital Theory Controversies? *Journal of Economic Perspectives*, 17(1), 199-214. http://dx.doi.org/10.1257/089533003321165010

Corrado, C., Haskel, J., Jona-Lasinio, C., and Iommi, M., (2022). Intangible Capital and Modern Economies. *Journal of Economic Perspectives*, 36(3), 3-28. http://dx.doi.org/10.1257/jep.36.3.3

Coutts, K., and Norman, N. (2013). Post-Keynesian Approaches to Industrial Pricing. A Survey and Critique. In: G.C. Harcourt and P. Kriesler (eds.), *The Oxford Handbook of Post-Keynesian Economics*. Oxford: Oxford University Press. https://doi.org/10.1093/

Day, E.E, and Persons, W.M. (1920). An Index of the Physical Volume of Production III. Manufacture, 1899-1919. *Review of Economics and Statistics*, 2(11), 309-337. http://dx.doi.org/10.2307/1928576

Denison, E.F. (1967). *Why Growth Rates Differ: Postwar Experience in Nine Western Countries*. (With the assistance of Jean-Pierre Poulliet). Washington, DC: The Brookings Institution.

Douglas, P.H. (1976). The Cobb-Douglas Production Function Once Again: Its History, Its Testing, and Some New Empirical Values. *Journal of Political Economy*, 84(5), 903-916. http://dx.doi.org/10.1086/260489

Falkus, M. (1996). Henry Phelps Brown as Economic Historian. *Review of Political Economy*, 8(2), 157-166. http://dx.doi.org/10.1080/09538259600000050

Felipe, J., and Fisher, F.M. (2003). Aggregation in Production Functions, What Applied Economists Should Know. *Metroeconomica*, 54(2-3), 208-262. http://dx.doi.org/10.1111/1467-999X.00166

Felipe, J., and Fisher, F.M. (2006). Aggregate Production Functions, Neoclassical Growth Models and the Aggregation Problem. *Estudios de Economía Aplicada*, 24(1), 27-16.

Felipe, J., and Fisher, F.M. (2008). Aggregation (Production). In: S.N. Durlauf and L.E. Blume (eds.), *The New Palgrave Dictionary of Economics*. Second Edition. London: Palgrave Macmillan. http://dx.doi.org/10.1057/978-1-349-95189-5_2552

Felipe, J., and McCombie, J.S.L. (2001). The CES Production Function, the Accounting Identity, and Occam's Razor. *Applied Economics*, 33(10), 1221-1232. http://dx.doi.org/10.1080/00036840122836

Felipe, J., and McCombie, J.S.L. (2006). The Tyranny of the Identity, Growth Accounting Revisited. *International Review of Applied Economics*, 20(3), 283-299. http://dx.doi.org/10.1080/02692170600735963

Felipe, J., and McCombie, J.S.L (2011). On Herbert Simon's Criticisms of the Cobb-Douglas and the CES Production Functions. *Journal of Post Keynesian Economics*, 34(2), 275-294. https://doi.org/10.2753/PKE0160-3477340205

Felipe, J., and McCombie, J.S.L. (2013). *The Aggregate Production Function and the Measurement of Technical Change: 'Not Even Wrong'*. Cheltenham: Edward Elgar. http://dx.doi.org/10.4337/9781782549680

Felipe, J., McCombie, J.S.L, and Mehta, A. (2024). *The Estimation of Production Functions with Monetary Values* [Working Paper no. 1036]. Levy Economics Institute of Bard College, New York.

Ferguson, C.E. (1969). *The Neoclassical Theory of Production and Distribution*. 2nd edition (reprinted with corrections, 1971). Cambridge: Cambridge University Press. http://dx.doi.org/10.1017/CBO9780511896255

Ferguson, C.E. (1971). Capital Theory up to Date: A Comment on Mrs Robinson's Article.*The Canadian Journal of Economics/Revue Canadienne d'Economique*, 4(2), 250-254. http://dx.doi.org/10.2307/133530

Fisher, F.M. (1969). The Existence of Aggregate Production Functions. *Econometrica*, 37(4), 553-577. http://dx.doi.org/10.2307/1910434

Fisher, F.M. (1971). Aggregate Production Functions and the Explanation of Wages: A Simulation Experiment. *Review of Economics and Statistics*, 53(4), 305-325. http://dx.doi.org/10.2307/1928732

Fisher, F.M. (1992). *Aggregation. Aggregate Production Functions and Related Topics.* (J. Monz, ed.). London: Harvester Wheatsheaf.

Fisher, F.M. (2005). Aggregate Production Functions – A Pervasive, but Unpersuasive, Fairytale. *Eastern Economic Journal,* 31(3), 489-491. https://www.jstor.org/stable/40326426

Friedman, M. (1953). The Methodology of Positive Economics. In: M. Friedman (ed.), *Essays in Positive Economics*. Chicago: University of Chicago Press. http://dx.doi.org/10.1017/CBO9780511819025.010

Haldane, A., Brennan, S., and Maduros, V. (2010). What is the Contribution of the Financial Sector: Miracle or Mirage? In: A. Turner (ed.), *The Future of Finance. The LSE Report*. London: LSE.

Hall, R., and Hitch, C. (1939). Price Theory and Business Behaviour. *Oxford Economic Papers,* 2(1), 12-45. https://doi.org/10.1093/oxepap/os-2.1.12

Haskel, J., and Westlake, S. (2017). *Capitalism without Capital. The Rise of the Intangible Economy*. Princeton: Princeton University Press. http://dx.doi.org/10.2307/j.ctvc77hhj

Hogan, W.P. (1958). Technical Progress and Production Functions. *Review of Economics and Statistics*, 40(4), 407-411. http://dx.doi.org/10.2307/1926345

McCombie, J.S.L (1998). "Are there Laws of Production?": An Assessment of the Early Criticisms of the Cobb-Douglas Function. *Review of Political Economy,* 10(2), 146-173. http://dx.doi.org/10.1080/09538259800000023

McCombie, J.S.L. (2001). What does the Aggregate Production Show? Further thoughts on Solow's "Second Thoughts on Growth Theory". *Journal of Post Keynesian Economics*, 23(4), 589-615. http://dx.doi.org/10.1080/01603477.2001.11490301

McCombie, J.S.L. (2011). "Cantabrigian Economics" and the Aggregate Production Function. *European Journal of Economics and Economic Policies: Intervention*, 8(1), 165-182. https://creativecommons.org/licenses/by/4.0/

McCombie, J.S.L., and Dixon, R. (1991). Estimating Technical Change in Aggregate Production Functions: A Critique. *International Review of Applied Economics*, 5(1), 24-46. http://dx.doi.org/10.1080/758524204

Mankiw, N.G., and Taylor, J.P. (2008). *Macroeconomics*. European Edition. New York: Worth Publishers.

Marschak, J., and Andrews, W. (1944). Random Simultaneous Equations and the Theory of Production. *Econometrica*, 12(3/4), 143-120 http://dx.doi.org/10.2307/1905432

Moseley, F. (2015). The Marginal Productivity Theory of Capital in Intermediate Microeconomic Textbooks: A Critique. *Review of Radical Political Economics*, 47(2), 292-308. http://dx.doi.org/10.1177/0486613414557919

Phelps Brown, E.H. (1957). The Meaning of the Fitted Cobb-Douglas Function. *Quarterly Journal of Economics*, 71(4), 546-560. http://dx.doi.org/10.2307/1885710

Phelps Brown, E.H. (1996). Autobiographical Notes. *Review of Political Economy*, 8(2), 129-140. http://dx.doi.org/10.1080/09538259600000045

Reder, M.W. (1943). An Alternative Interpretation of the Cobb-Douglas Function. *Econometrica*, 11(3 /4), 259-264. http://dx.doi.org/10.2307/1905678

Riach, P.A. (2019). Henry Phelps Brown (1906-1994). In: R.A. Cord (ed.), *The Palgrave Companion to LSE Economics* (p. 463-485). London: Palgrave Macmillan. http://dx.doi.org/10.1057/978-1-137-58274-4_18

Robinson, J.V. (1953-4). The Production Function and the Theory of Capital. *Review of Economic Studies,* 21(2), 81-106. http://dx.doi.org/10.1016/B978-0-12-590550-3.50012-4

Robinson, J.V. (1970). Capital Theory Up to Date. *The Canadian Journal of Economics/Revue Canadienne d'Economique*, 3(2), 309-317. http://dx.doi.org/10.2307/133680

Robinson, J.V. (1975). The Unimportance of Reswitching. *Quarterly Journal of Economics*, 89(1), 32-39. http://dx.doi.org/10.2307/1881707

Routh, G. (2008). Phelps Brown, (Ernest), Henry, (1906-1994). In: S.N. Durlauf and L.E. Blume (eds.), *The New Palgrave Dictionary of Economics.* London: Palgrave Macmillan. http://dx.doi.org/10.1057/978-1-349-95121-5_1298-2

Samuelson, P.A. (1979). Paul Douglas's Measurement of Production Functions and Marginal Productivities. *Journal of Political Economy*, 87(5), 923-939. http://dx.doi.org/10.1086/260806

Shaikh, A. (1974). Laws of Production and Laws of Algebra: The Humbug Production Function. *Review of Economics and Statistics*, 56(1), 115-120. http://dx.doi.org/10.2307/1927538

Shaikh, A. (1980). Laws of Production and Laws of Algebra: Humbug II. In: E.J. Nell (ed.), *Growth, Profits and Property: Essays in the Revival of Political Economy*. Cambridge: Cambridge University Press. http://dx.doi.org/10.1017/CBO9780511571787.007

Simon, H.A. (1979a). Rational Decision-Making in Business Organizations. *American Economic Review,* 69(4), 493-513. https://www.jstor.org/stable/1808698

Simon, H.A. (1979b). On Parsimonious Explanations of Production Relations. *Scandinavian Journal of Economics*, 81(4), 459-474. http://dx.doi.org/10.2307/3439661

Simon, H.A., and Levy, F.K. (1963). A Note on the Cobb-Douglas Function. *Review of Economic Studies*, 30(2), 93-94. http://dx.doi.org/10.2307/2295806

Solow, R.M. (1956). A Contribution to the Theory of Economic Growth. *Quarterly Journal of Economics*, 70(1), 65-94. http://dx.doi.org/10.2307/1884513

Solow, R.M. (1957). Technical Change and the Aggregate Production Function. *Review of Economics and Statistics*, 39(3), 312-320. http://dx.doi.org/10.2307/1926047

Solow, R.M. (1958). Technical Progress and Production Functions: Reply. *Review of Economics and Statistics,* 40(4), 411-413. http://dx.doi.org/10.2307/1926346

Solow, R.M. (1974). Laws of Production and Laws of Algebra. The Humbug Production Function: A Comment. *Review of Economics and Statistics,* 56(1), 121. http://dx.doi.org/10.2307/1927539

Solow, R.M. (1987). Second Thoughts on Growth Theory. In: A. Steinherr and D. Weiserbs (eds.), *Employment and Growth: Issues for the 1980s. International Studies in Economics and Econometrics*. Vol. 16. Dordrecht: Springer.

Solow, R.M. (1988). Growth Theory and After. *American Economic Review*, 78(3), 307-317. https://www.jstor.org/stable/1809135

Walters, A.A. (1963). Production and Cost Functions: An Econometric Survey. *Econometrica*, 31(1/2), 1-66. http://dx.doi.org/10.2307/1910949

Wibe, S. (1984). Engineering Production Functions. A Survey. *Economica,* 51(204), 401-411. http://dx.doi.org/10.2307/2554225

Wicksell, K. (1896). Eim neues Prinzip der gerechen Besteurung. Finanz-theoretische Untersuchungenm. Translated as: A New Principle of Just Taxation. In: R.A. Musgrave and A.T. Peacock (eds.), *Classics in the Theory of Public Finance*. London: Macmillan, 1958.

Wilson, T., and Andrews, P.W.S. (1951). *Oxford Studies in the Price Mechanism*. Oxford: Clarendon Press.

Whiteman, J.L. (1988). The Efficiency of Labour and Capital in Australian Manufacturing. *Applied Economics*, 20(2), 243-261. http://dx.doi.org/10.1080/00036848800000008