Review of Political Economy
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/crpe20

The Aggregate Production Function: 'Not Even Wrong'

J. Felipe[a] & J.S.L. McCombie[b]
[a] Asian Development Bank, Manila, Philippines
[b] Cambridge Centre for Economic and Public Policy, University of Cambridge, Cambridge, UK
Published online: 19 Mar 2014.

To cite this article: J. Felipe & J.S.L. McCombie (2014) The Aggregate Production Function: 'Not Even Wrong', Review of Political Economy, 26:1, 60-84, DOI: 10.1080/09538259.2013.874192

To link to this article: http://dx.doi.org/10.1080/09538259.2013.874192

Review of Political Economy, 2014
Vol. 26, No. 1, 60–84, http://dx.doi.org/10.1080/09538259.2013.874192

The Aggregate Production Function: 'Not Even Wrong'

J. FELIPE* & J.S.L. McCOMBIE**

**Asian Development Bank, Manila, Philippines; **Cambridge Centre for Economic and Public Policy, University of Cambridge, Cambridge, UK*

(Received 18 December 2012; accepted 7 March 2013)

ABSTRACT *The foundations of the aggregate production function were long ago thrown into doubt by problems of aggregation and the Cambridge capital theory controversies. Yet the aggregate production function, whether in the familiar form of the Cobb-Douglas, the CES, or the translog, continues to be widely used in both theoretical and applied analysis. The reason for its continued use rests on the instrumental position that 'it works'. The aggregate production function sometimes yields good statistical fits with plausible estimates of the coefficients. However, for some time, it has been realised that the existence of an underlying accounting identity can explain the regression results, even if the aggregate production function does not exist. This argument has been widely ignored. This paper, drawing on a rhetorical approach, assesses why this is the case. It shows that the few criticisms that have been made of the critique involve fundamental misunderstandings that represent a failure of the economic method.*

> The physicist Wolfgang Pauli was, with Heisenberg, Schrödinger and Dirac, one of the early leaders in the development of quantum mechanics. He was renowned for being a tough audience, exclaiming 'wrong' (*falsch*), or 'completely wrong' (*ganz falsch*) when he disagreed with a speaker. Near the end of his life, when asked his opinion of a recent article by a younger physicist, he sadly said 'it is not even wrong' (*Das ist nicht einmal falsch*) . . . *A scientific idea is 'not even wrong' if it is so incomplete that it cannot be used to make predictions that could be compared to observations to see if the idea is wrong*. (Woit, 2006, p. 6; emphasis added)

1. Introduction

The aggregate production function is at the heart of mainstream macroeconomics, whether it is neoclassical growth theory, real business cycle models

Correspondence Address: J. S. L. McCombie, Department of Land Economy, Cambridge Centre for Economic and Public Policy, University of Cambridge, 19 Silver Street, Cambridge CB3 9EP, UK, E-mail: jslm2@cam.ac.uk

or the short-run demand for labour. It is now treated in neoclassical economics as a useful and unproblematical, albeit simplified, representation of the underlying technology of an economy. It is widely used, together with the marginal productivity theory of factor pricing, in innumerable applied and theoretical articles on macroeconomics.

Yet the concept has been subject for many years to fundamental criticisms that question its theoretical foundations.[1] The conditions under which one can aggregate different heterogeneous outputs and factors of production and sum across different micro-production functions to give an aggregate production function are so stringent as to make it difficult to believe that such a function can exist in reality. This 'aggregation problem' has been extensively examined since the 1940s. Franklin Fisher (2005, p. 490) summarised the conclusions of this literature as follows:

> Even under constant returns, the conditions for aggregation are so very stringent as to make the existence of aggregate production functions in real economies a non-event. . . . One cannot escape the force of these arguments 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 technical relations require close approximation to stringent conditions, and this is not a sensible thing to suppose.

The implications of this work have been totally ignored by mainstream economists, who continue to estimate aggregate production functions and to use them in theoretical models.[2]

The Cambridge capital theory controversies, which lasted from about 1965 to 1975, certainly drew more attention (Cohen & Harcourt, 2003, 2005). Nevertheless, in spite of the somewhat acrimonious nature of the debate and Samuelson's (1966) capitulation, the capital theory controversies have come to be largely forgotten by neoclassical economists. While from time to time the debate has been re-examined (although not rekindled) by heterodox economists, it is now generally seen by the mainstream as a contribution to the history of economic thought.

So why do neoclassical economists persist with using the aggregate production function? The implicit defence undoubtedly rests on Milton Friedman's instrumentalism as expressed in his essay on 'The Methodology of Positive Economics' (1953). The mark of 'truly important and significant hypotheses' is that although their assumptions will be 'wildly inaccurate descriptive representations of reality', they nevertheless have good predictive power (p. 14). The reason that the aggregate production function is so widely used, no matter what the theoretical shortcomings are, is that 'it works' in the sense that Friedman had in mind. That is to say, it usually gives plausible and statistically significant estimates.

[1]The initial reception of the Cobb-Douglas production function was so hostile, from both the *a priorists* at the University of Chicago and the econometricians, that Douglas nearly gave up on the endeavour (McCombie, 1998a). However, in 2011, the paper in which Cobb & Douglas (1928) originally formulated their production function was listed as one of the top twenty 'most admirable and important articles' published in the *American Economic Review* (Arrow *et al*., 2011).

[2]For an accessible survey of this somewhat technical literature see Felipe & Fisher (2003).

What is 'plausible' begs the question to some extent, but as Robert Solow once remarked to Franklin Fisher, 'had Douglas found labor's share to be 25% and capital's 75% instead of the other way around, we would not now be discussing aggregate production functions' (Fisher, 1971, p. 305). However, it should be noted that cross-sectional data normally give the best statistical fits. Estimates using time-series data can provide implausible estimates of the coefficients with, especially, the coefficient of capital sometimes being insignificant. The reason, as will be discussed below, is the difficulty of modelling 'technical progress' (or total factor productivity growth, in the neoclassical parlance). Because of this, there is a large literature that still regards the problems of estimating aggregate production functions as mostly econometric questions (see, for example, Blundell & Bond, 2000; Griliches & Mairesse, 1998; Levinsohn & Petrin, 2003; Olley & Pakes, 1996).

But here we come to the crux of this paper. These studies about the econometric considerations in estimating aggregate production functions are irrelevant. With sufficient ingenuity and the knowledge of the data given by an underlying accounting identity, we can always find a specification of the 'aggregate production function' that cannot fail to give a close statistical fit, with the estimated output elasticities very close to the factor shares. Although the micro-production function is a technological relationship and should be estimated using physical data, because of the heterogeneity of both output and capital these variables have to be summed using constant-price value data.[3] This distinction is rarely, if ever, mentioned in the neoclassical literature on production theory. This change from physical to value measurement units is not innocuous, as it introduces an underlying accounting identity that *definitionally r*elates value added to the total compensation of labour and total profits. This identity, $V \equiv wL + rJ$ (where V is value added, w is the wage rate in constant prices, L is employment, r is the rate of profit and J is the constant-price value of the capital stock), can always be mathematically transformed into a specific functional form resembling a production function that gives a very close, or indeed, a perfect statistical fit to the data, even though no aggregate production function actually exists.[4] It is 'not even wrong' in Pauli's sense of the term, because the identity renders a meaningful empirical interpretation of the estimated parameters of the aggregate production function impossible. This damaging critique of the neoclassical aggregate

[3]Ferguson (1971, p. 250), the author of what for many years was considered the definitive book on neoclassical production theory (Ferguson, 1969), wrote that he 'assumed a production function relating physical output to the physical inputs of heterogeneous labour, heterogeneous machines and heterogeneous raw materials'. The problem that many sectors—for example, defence, health, education and financial services—do not have a well-defined physical output is usually glossed over. See Carter (2011a) for a detailed discussion of Ferguson's views.

[4]Somewhat unconventionally, an aggregate production function is defined in this paper as one that requires constant-price monetary data for output and capital in its estimation. It thus encompasses production functions for the whole economy to, say, the four-digit industry level of aggregation. The term micro-production function is reserved for one where the inputs are measured in physical terms (i.e. engineering production functions). The arguments in this article hold equally for when gross output is used and intermediate inputs are included in the production function.

production function is one of logic and not of subjective plausibility. The purpose of this paper is to analyse why the critique has been ignored.[5]

Our argument will make use of Kuhn's concept of paradigm and McCloskey's insights on the rhetoric of economics; these ideas are discussed in the next section. In Section 3 we consider how the concept of the neoclassical aggregate production function is derived and briefly recapitulate the identity argument. We then look at some of the early reactions to the critique and in particular at Solow's (1974) one-page dismissal of Shaikh's (1974) landmark paper on the Humbug production function. We conclude by considering some further fundamental misunderstandings of the accounting identity critique. These misunderstandings are often a *petitio principii*, or guilty of circular reasoning; that is to say, they use the assumption that the aggregate production function does exist as a defence against the argument that it doesn't.

2. Paradigms and the Rhetoric of Economics[6]

What vantage point is likely to give a workaday economist a deeper understanding of the causes of unresolved disputes in economics such as this one? Thomas Kuhn has pointed out the role of the paradigm, or disciplinary matrix, in protecting scientists from facing severe methodological problems in their daily experiments. The scientist learns the appropriate method by demonstration, by the worked examples from the textbooks that reflect the prevailing disciplinary consensus. Hence, Kuhn's distinction between 'puzzle solving', which is how normal science proceeds, and the 'crisis' when unexplained results, previously shelved as anomalies, can no longer be ignored. The paradigm has both a negative and positive role. By seeming to legitimise the puzzles, it protects scientists from continuously having to examine the fundamental postulates of their approach. As Kuhn points out, if this were not the case then little scientific work would be done. But in economics, what is it that causes an economist to accept one approach, or model, and to reject another mutually exclusive explanation? One obvious, and powerful, influence is the training the economist has received. But the more general question is what persuades economists, bearing in mind that their view will be shaped by the paradigm in which they were trained? This brings us to the role of rhetoric in shaping theory choice in economics, an issue that has been explored by McCloskey (1994, 1997, 1998, 2008).

McCloskey, at least in her early writings, suggests that normal debate within the economics profession, the 'economics conversation', generally ensures that economics progresses, however one defines progress.[7] The academic debate is sufficient, she argues, maintaining that 'overlapping conversations provide the standards. It is a market argument. There is no need for philosophical law-making or

[5]Although we inevitably address some of the technical literature, our focus will be on the reasons why the accounting identity argument has proved so unpersuasive to mainstream economists. For a detailed discussion see Felipe & McCombie (2013).

[6]This section draws upon and extends McCombie (1998b).

[7]McCloskey's book on *The Rhetoric of Economics* was originally published in 1985; we quote here from the 1998 second edition.

methodological regulation to keep the economy of the intellect running just fine' (McCloskey, 1998, p. 163). Rhetoric and disputation have ensured that 'Economics in its modern and mathematical form has grown into a brilliantly successful science' (McCloskey, 1994, p. xi). The problems of the aggregate production function were extensively debated between 1950 and 1980; according to McCloskey's rhetoric of economics framework, the fact that they are no longer discussed must mean that no important issues concerning the production function remain to be resolved. This view, often espoused by mainstream economists, has a Kuhnian flavour, in so far as it reflects Kuhn's idea that science progresses through discourse within scientific community (although he admits that there is no objective way of defining progress). The crucial difference between the natural sciences and economics is that in the former, controlled experiments can arbitrate between theories. This option is generally absent in economics, and some economists view econometrics as a poor and unconvincing substitute.[8] McCloskey's rhetoric framework, therefore, begs the question of how we can be sure that debates between economists will always lead to progress in economics.

This free market analogy has been challenged by Mirowski (1987), among others. But more recently McCloskey appears to have changed her view. For over 20 years, she has been criticising the economics profession for its failure to distinguish between Fisherian economic significance and economic importance. According to McCloskey, econometricians agree the latter is important in private, but repeatedly fail to acknowledge this in print.[9] McCloskey & Ziliak (2008, p. 47) contend that 'all the econometric findings since the 1930s need to be done over again.' If this is correct then it represents a major failure of the econometric method. There are striking parallels between this and the treatment of the aggregate production function, as we shall see.

McCloskey (1994, pp. 87–88) concedes that:

> For students of science in the here and now it is naïve to think that power, analogy, upbringing, story, prestige, style, interest, and passion cannot block science for years, decades, centuries. The naïve view is that science is rational in a rationalist sense, that is, non-rhetorical and non-sociological, understandable in our rationalist terms now, not at dusk. The history and sociology and the rhetoric of science says it isn't so.

As Kuhn (1999) has shown, one of the hallmarks of competing paradigms is that there is likely to be partial incommensurability. Not *all* concepts employed in two paradigms change in the translation from one to another. But there is no neutral observational language: 'Two languages are incommensurable in just those

[8]See, for example, Summers (1991, p. 130): 'I invite the reader to try and indentify a single instance in which a "deep structural parameter" has been estimated in a way that has affected the profession's beliefs about the nature of preferences or production technologies or to identify a meaningful hypothesis about economic behavior that has fallen into disrepute because of a formal statistical test'.

[9]The impact of Ziliak & McCloskey's critique on the economics profession has been minimal. One of the few formal responses was put forth by Hoover & Siegler (2008a) two decades after McCloskey first raised the issue; see also the reply by McCloskey & Ziliak (2008) and the rejoinder by Hoover & Siegler (2008b). To be fair, their book (Ziliak & McCloskey, 2008) was extensively reviewed.

areas where they project nature in incompatible ways' (Kuhn, 1999, p. 34). Yet, in economics there are cases where economists disagree about the implications of models even using the same symbols and theoretical concepts and, as in the Cambridge capital theory controversies, still talk past each other. We term this 'weak' incommensurability, as opposed to 'strong' incommensurability. With the latter, concepts in one paradigm literally have no meaning in another (e.g. utility has no explanatory meaning in Marxian economics and social class has no meaning in the neoclassical paradigm). Many continuing debates in economics (such as those between the adherents to the Post-Keynesian and the neoclassical paradigms) result from weak incommensurability. In other words, the competing paradigms have a shared vocabulary and, in many cases, use the same econometric tools and mathematical techniques, but nevertheless have different fundamental assumptions.

An example is the different emphasis placed on the Cambridge capital theory controversies by Fisher and Harcourt. Fisher, who has done more than most to determine the full extent of the aggregation problem, sees the Cambridge capital theory controversies as merely a subset of the former:

> Reswitching and reverse capital deepening only appear paradoxical if one supposes that aggregates should behave the way intuition suggests they should behave—the way that factors of production and outputs behave at the micro level. But the non-existence of aggregate production functions means that such intuition simply does not apply. No further consequence can be read from its failure. (Fisher, 2005, p. 491)

Harcourt (1976, p. 29), however, sees more to it than this.

> What is involved is the relevant 'vision' of the economic system and the historical processes associated with its development. In particular, stress is laid by the Cambridge (England) school on production and distribution as involving underlying social relations, especially their implications for production and distribution, accumulation and growth, that the capitalists own the means of production (a necessary condition for undertaking production) and the wage earning classes can only sell their labour services to the capitalists. Capital hires labour but labour does not hire capital.

It can be seen here that there is weak incommensurability between the views of Fisher and Harcourt. This is also reflected in the Cambridge capital theory controversies. The incommensurability is 'weak' because the debate centred on the same models and concerned matters of logical inference, not empirical issues that could be subject to different interpretations. In this regard, there was agreement over the formal results, but not over their implications. In the neoclassical paradigm, the aggregation problem is a mere technical problem that caused some difficulties for the aggregate production function, but in the long run had no effect on its ubiquitous (and uncritical) use (notwithstanding Fisher's arguments). Discussion about the meaning of capital and the social relations of production are dismissed as merely ideological by, for example, Solow (1988, p. 309) (but how could they be otherwise?) in the pejorative sense of the word.

There is no guarantee that the use of rhetoric is an infallible process. In later writings, McCloskey (e.g. 1997) implies that rhetoric does not ensure progress; it

cannot even ensure that fundamental criticisms will be discussed: they may simply be ignored. Appeals to authority, where the 'authority' is the dominant paradigm, may cause the critique to be dismissed as unimportant. But the paradigm is itself partially determined by sociological forces, by the dictates of peer review. There is no 'universal economics conversation'. Given the thousands of articles submitted each year for publication, the editors of the leading journals are the filters; they decide who is worth being heard and who is not. This choice is irreducibly social; the decision is in the hands of the relatively few editors and their appointed referees. If a particular argument is not heard, the interpretation is that it can't be important. Consequently, there is a Catch-22 problem.

We shall argue below that many of the erroneous criticisms of the accounting identity critique can be best explained in terms of weak incommensurability. The aggregate production function is, to use a Kuhnian term, a *paradigmatic pseudo-assumption* (Hoyningen-Huene, 1993, p. 210). It is seen within the neoclassical framework as an empirical concept that is so central to neoclassical macroeconomics that it is implicitly deemed by fiat within the paradigm to be unfalsifiable. Aggregate production functions are estimated not in an attempt to test the hypothesis that they exist, but rather to estimate the parameter values that are assumed to reflect the underlying technology of the plant, industry, or economy. A consequence of this is that in a number of cases the assumption of the existence of a well-defined aggregate production function, whose functional form can be determined by econometric methods and value data, is used to defend the concept against criticisms that it does not exist. We first digress somewhat to consider the orthodox approach to the production function as a background to a critique of the concept (see also Felipe & McCombie, 2013).

3. From Microeconomic to Aggregate Production Functions

The fact that the output of commodities of goods or services is determined by the flow of inputs is uncontroversial. Intuitively, we may think of an engineer designing a production process for, say, an oil refinery. The plans will indicate the most efficient techniques, that is to say, the least costly combinations of labour and machines, associated with different scales of output. The aggregate production function reduces these complex production processes to a simple functional form:

$$Q = f(K_1, K_2, K_i \ldots L_1, L_2, L_j \ldots M_1, M_2, M_k) \tag{1}$$

where Q is the volume of homogeneous output measured in physical terms, and K, L, and M are the flows of labour services, physical capital services (broadly defined to include buildings, etc) and raw materials. The formalisation for these complex production processes of the firm is given by the familiar equation for a production function: $Q = Af(K, L)$ where Q is the maximum amount of output (numbers of homogeneous 'widgets') that can be produced from any combination of K and L (where capital K is measured as numbers of 'leets', to use Joan Robinson's term). In other words, there is a unique mapping from K and L to Q. A necessary assumption for aggregation is that the inputs are used optimally. But

already a further simplification has been made: *A* denotes the level of technological knowledge, assumed to grow exponentially over time.[10]

Certain tacit rules are inculcated in the economics student from this simple analysis. They are tacit because, as Kuhn has argued, the methodological rules of the game are set out by demonstration, through worked examples and the textual exposition, *per se*. The student learns implicitly to treat capital and labour as factors of production standing on an equal footing. The marginal productivity theory of factor pricing that is central to the neoclassical theory of production is uncritically accepted. Just as the wage rate is the reward for the contribution that labour makes to the production of output, so the payment to the owners of the capital stock is a reward for the latter's contribution to production (or abstinence from consumption). Textbooks reflect the view of John Bates Clark (1891, p. 313) that 'what a social class gets is, under natural law, what it contributes to the general output of industry'.

The concept of the production function then moves seamlessly from the narrow engineering definition to a macroeconomic concept relating the inputs to the output of the whole economy, total manufacturing or individual industries. Most textbooks make no attempt to show how an aggregate production function can be derived from the underlying micro-production functions (see, for example, Aghion & Howitt, 2009 or Barro & Sala-i-Martin, 2003). In the microeconomic textbook by Estrin & Laidler (1995, p. 134) we do find a discussion of the properties of the production function, including a brief mention of the aggregation problem, but this is quickly brushed aside: 'The results of the two input/output special case are both useful and often capable of being generalised and therefore well worth the reader's attention.'

There is, to some, a persuasive metaphor at work here. Nobody would deny that there is a physical relationship between outputs and inputs (remember the technical blueprints). Of course, as we move up the hierarchy from plant to firm to the whole economy, we are employing increasingly aggregated conceptualisations. Particularly crucial for empirical purposes are value measures of output and capital that stand in for the vectors of heterogeneous units of these variables. One of the points of economics is to explain a lot from a little. As Solow (1956, p. 65) remarks at the start of his modestly titled 'A Contribution to the Theory of Growth', 'All theory depends upon assumptions that are not quite true. That is what makes it theory'. His companion article on applied aspects of growth theory contains a masterful and, to many, convincing piece of rhetoric.

> 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 .... The new wrinkle I want to describe is an elementary way of segregating variations in output per head due to technical change from those due to the availability of capital per head. Either that kind of aggregate economics appeals or it doesn't. Personally, I belong to both schools .... It is convenient to begin with the special case of

[10]This may be modelled as a knowledge production function, as in endogenous growth theory, but we will not consider it here for reasons of space. It does not affect our argument.

> *neutral* technical change. In that case the production function takes the special form $Q = A(t)f(K, L)$ and the multiplicative factor $A(t)$ measures the cumulative effect over time. (Solow, 1957, p. 312, emphasis in the original)

This paragraph has been analysed rhetorically by McCloskey (1998, pp. 48–51) who shows that the four master tropes are at work here: 'The argument depends at once on a metaphor. The "aggregate production function" Solow diffidently introduces (he is not really diffident: he is pretending to be for rhetorical effect) says the making of our daily bread is like a mathematical function' (McCloskey, 1998, p. 48). Put like this, it seems incredible that a function with only two arguments, K and L, together with a shift factor, can adequately represent the total output of an economy. McCloskey (1998, p. 49) continues that 'Economists are used to such figures of speech ... to the point of not recognizing that they are, but non-economists will agree that the figures are bold.' The use of the metonym K obviates the need to worry about all the problems of the aggregation of such diverse capital goods as steel girders and personal computers. The use of 'symbolic generalizations' (Kuhn, 1970, p. 182) such as $Q = F(K, L, t)$, $p_K^F = \rho$, $p_L^F = w$, $F_i > 0$, $F_{ii} < 0$ $(i = K, L)$ needs no further explanation; indeed the symbols hardly need defining for the intended audience. The underlying assumptions need no explicit justification in either articles or seminars. We next consider in more detail the problems that this metaphor of the aggregate production function ignores.

4. On Behavioural Equations and Identities

A fundamental problem arises from the fact that with a variety of outputs and different machines and structures, aggregation requires the use of constant-price value data for output (V) and the capital stock (J). These are derived from National Income and Product Accounts and are generally treated as unproblematic proxies for their physical counterparts. Thus, the general form of the production function is specified as $V = f(J, L, t)$. The underlying accounting identity is given by:

$$V \equiv wL + rJ \tag{2}$$

where w is the real wage rate (in monetary terms) and r is the ex-post rate of profit. Note that this is just what its name suggests; it is an identity and is compatible with any or none of the usual neoclassical assumptions, including whether or not markets are perfectly competitive, factors are paid their marginal products, and there are constant or increasing returns to scale. It is not possible to recover any physical magnitudes from equation (2). V and J are also used in the estimation of the aggregate production function. This causes the neoclassical analogy that output and capital can be regarded as *physical* magnitudes (reference is often made to the 'volume' of output) to breakdown. V and J are value measures related through the identity, equation (2). An identity is a tautology; it is definitionally true. There is simply no point in estimating statistically an identity or regarding such estimation as a legitimate test of any economic hypothesis.

Let us eschew any economics for the moment and just consider the simple identity given by the equation $Y \equiv X + Z$. Totally differentiating this identity, we obtain $\mathrm{d}Y \equiv \mathrm{d}X + \mathrm{d}Z$. This may be equivalently written as $\mathrm{d}Y/Y \equiv \theta \mathrm{d}X/X + (1-\theta)\mathrm{d}Z/Z$, where $\theta \equiv X/Y$ and $(1-\theta) \equiv Z/Y$. Integrating this equation gives $Y \equiv CX^{\theta}Z^{(1-\theta)}$ where C is the constant of integration and is equal to $\theta^{-\theta}(1-\theta)^{-(1-\theta)}$. It should be emphasised that $Y \equiv CX^{\theta}Z^{(1-\theta)}$ is not an approximation of $Y \equiv X + Z$: it is an exact mathematical transformation.[11]

Now apply the same arithmetical manipulation to equation (2). The following identity is obtained:

$$V \equiv a^{-a}(1-a)^{-(1-a)}r^{a}w^{(1-a)}J^{a}L^{(1-a)} \equiv AJ^{a}L^{(1-a)} \tag{3}$$

Taking any data for a given period, say a year, equation (3) is formally equivalent to equation (2) (see McCombie, 2011, p. 177). Hence, it is obvious why Douglas had such spectacular results in the 1930s by estimating the Cobb-Douglas production function using cross-sectional (industry) data, and likewise Hildebrand & Liu (1965) and Moroney (1972) using US cross-state data. As wages, the rate of profit and the factor shares show relatively little variation between the units of observation compared with V, K and L, the Cobb-Douglas function must give a very good statistical fit to the data (McCombie, 1998a; Simon, 1979a).

But does this mean that the argument only applies when factor shares are constant, as, for example, Temple (2006, 2010) seems to think? The answer, even intuitively, is no.[12] Suppose cross-sectional data are used and shares vary. A more flexible transformation of the accounting identity is needed. We could, for example, consider a Box-Cox transformation, which is a mathematical transformation with no economic implications. Depending on the degree of variation of the factor shares, this technique will give the most general functional form that is, in effect, a CES function; but, depending upon the variation in the data, the estimates will also give the Cobb-Douglas or a linear function. Again, there is no economics in this analysis.

We may also demonstrate this by considering the identity in terms of exponential growth rates. This dynamic accounting identity may be expressed as:

$$\dot{V}_t = a_t\dot{r}_t + (1-a_t)\dot{w}_t + a_t\dot{J}_t + (1-a_t)\dot{L}_t \tag{4}$$

where a dot over a variable denotes an exponential growth rate. The neoclassical economist, starting with the general 'production function' $V = f(J, L, t)$, derives:

$$\dot{V}_t = \lambda_t + \alpha_t\dot{J}_t + \beta_t\dot{L}_t \tag{5}$$

[11]The reader may easily confirm this by using any values for X and Z.

[12]What is the necessary degree of variation in factor shares before we can be confident that a 'true' production function is being estimated? This suggests there is something like a statistical identification problem here, which is not the case.

where λ is the rate of technical progress. The isomorphism is again apparent with $\lambda_t \equiv a_t\dot{r}_t + (1 - a_t)\dot{w}_t$. The requirement is to find a specific functional form for equations (4) and (5) that accurately tracks the movement of the factor shares (or, equivalently, output elasticities) over time. This is the accounting identity critique in a nutshell, although it inevitably abstracts from many of the adjunct complex issues that it raises. The importance of this argument is that it completely undermines the instrumental defence of the use of the aggregate production function on the grounds that it gives good predictions. We can always find a specification using any data set of V, J and L that will give a good fit to a putative aggregate production function, which consequently, is 'not even wrong'.

5. Early Reactions: A Failure to Appreciate the Full Implications

The above argument is a matter of logic, and as such represents a challenge to adherents of the aggregate production function. Yet it has been almost totally ignored in the literature because of weak incommensurability. Consequently, we are applying the term incommensurability to questions of logic, as the accounting identity critique is a logical rather than empirical issue. This might be stretching the usage of the notion of incommensurability somewhat, since a logical argument is either right or wrong. Therefore, one may ask: where does the incommensurability problem come into play? It comes into play to the extent that the critique is sometimes erroneously interpreted as depending on special assumptions such as constant factor shares (e.g. Temple, 2006, 2010) even in the face of arguments that show logically that it does not (Felipe & McCombie, 2005a, 2010, 2011b). The view of the aggregate production function as a paradigmatic pseudo-assumption has led to three misleading rhetorical devices that are often used in attempting to refute the accounting identity critique.

The first is *petitio principii*, or the use of circular reasoning or an argument that begs the question. This is often found where the implicit assumption that the aggregate production function exists is used to counter the critique's argument that it is not possible to test this hypothesis.

The second is *ignoratio elenchi*, which is the fallacy of using a chain of reasoning which in itself is valid but that has no relevance for the issue being addressed. Temple (2010), for example, in attempting to minimise the importance of the accounting identity critique, cites numerous works that have solved various econometric problems connected to the estimation of aggregate production functions, even though such problems have no bearing on the critique. The latter has nothing to do with the correct method of statistical estimation.

Finally, there is *appeal to authority*, not necessarily the authority of an individual, but to the authority of the paradigm. The paradigm is used to defend the paradigm. The critique is seen as a mere technical anomaly and hence to be shelved, while the paradigm continues to legitimise the use of the aggregate production function in both theoretical and empirical work, albeit as a rough, but useful, parable of the underlying technological structure of the economy.

In order to understand the reasons for this, it is useful to consider the early reactions to the critique before considering later misunderstandings. The starting point of the accounting identity critique is Phelps Brown's (1957, p. 557)

argument in the *Quarterly Journal of Economics*. Here, in a multifaceted criticism of Douglas's work on the aggregate production function, he noted that all that was being estimated was the accounting identity, and that 'the Cobb-Douglas [$1 - \alpha$] and the share of earnings in income, will be only two sides of the same penny'. (He noted that differentiating the identity, given by equation (2), gives $\partial V/\partial L = w$). This had virtually no impact on the profession. The paper was somewhat obscure, and even Simon & Levy (1963), who wrote a short and clear note on the problem, were not completely sure of Phelps Brown's argument. Starting with the traditional Cobb-Douglas production function, they used a Taylor-series expansion to show that it could be expressed as a linear function, which was identical to the accounting identity at the point of evaluation. The extremely brief note attracted little attention.

And even when the note by Simon & Levy has attracted attention it has been misinterpreted. The counter-argument is that Simon & Levy (1963) had merely shown that the aggregate production function could be approximated as a linear function, i.e. as the accounting identity. However, this is an example of *petitio principii* and is a misreading of the argument. The only underlying relationship postulated by Simon & Levy (1963) was the linear function $V = wL + rJ$ (in our notation), and it was this that could be approximated to a Cobb-Douglas function. In other words, the 'causation' was from the linear accounting identity to the Cobb-Douglas relationship, and not vice versa. Simon & Levy may not have fully appreciated the implications of their argument, for they referred to this linear relationship as a 'production function' when their analysis indicates clearly that it is not a function at all but an identity.

Simon was deeply sceptical of the marginal productivity theory of factor pricing, as his correspondence in the early 1970s with Solow, recently unearthed by Carter (2011b), shows (see also Felipe & McCombie, 2011a). He thought the problem so serious that he wrote in his Nobel Prize acceptance speech:

> Fitted Cobb-Douglas functions are homogeneous generally of degree close to unity with a labor exponent of about the right magnitude. These findings, however, cannot be taken as strong evidence for the [neo]classical theory, for 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) (see E. H. Phelps Brown [1957]). The same comment applies to the SMAC [CES] production function .... (Simon, 1979b, p. 497)

There is no reference to a 'true' production function here and Simon's argument is not just confined to the Cobb-Douglas function. But the impact of Simon's views was minimal. One reason why Simon was sceptical about the aggregate production function may be that he was a polymath with publications in such diverse fields as physics, philosophy of science and organisational decision-making; that is to say, he was not the typical economist brought up in, and trained exclusively within, the neoclassical paradigm (see Simon, 1978).

Somewhat earlier, Walters (1963) in his then definitive survey of cost and production functions alluded to Phelps Brown's critique that the production function could be explained in terms of a 'simple linear relationship'. However, the

force of his argument was undermined by his erroneous insistence, immediately after discussing Phelps Brown (1957), that 'aggregate industry data have been used with considerable success in *interstate* (or *international*) studies of the SMAC [CES] function' (p. 38, emphasis in the original). Moreover, while he conceded that 'the theoretical foundations of the aggregate production function give one grounds for doubting whether the concept is at all useful,' Walters still held that 'the temptation to discuss movements in indices and output in terms of a function is difficult to resist' (Walters, 1963, p. 425). No justification was advanced for this assertion.

Samuelson (1979) in his far from hagiographic tribute to his former teacher Paul Douglas rediscovered the argument as it applied 'tautologically' to cross-section data and gives an eloquent example of it, but—as far as we are aware—never followed it up. The paper, although published in the *Journal of Political Economy*, had no impact. After all, it only referred to some empirical work done many years ago with basic statistical techniques.

Fisher (1971) attempted to shed light on why the aggregate Cobb-Douglas production function worked when, in theory, it should not. He resorted to simulation analysis. If the aggregation theorems are not persuasive, then perhaps an empirical counterexample would be more effective. The advantage of simulation analysis is that the underlying structure of the economy is known to the reader with certainty, but the hypothetical researcher with access only to the simulation results does not know this. This approach shows how the researcher can draw erroneous inferences if the person only has knowledge of the generated data. In Fisher's simulation, the economy consists of four micro Cobb-Douglas production functions with output growing over a 20-year period, but the conditions for successful aggregation are deliberately violated so an aggregate production function does not exist. Fisher runs the simulation for the micro-production functions over this period. Aggregating the data and then estimating a wage equation shows paradoxically that the statistically significant results are compatible with an aggregate Cobb-Douglas production function. But as Fisher repeatedly emphasised, the direction of causation is from the stability of the aggregate factor shares to the aggregate Cobb-Douglas production function and not vice versa. He did not explain why the aggregate factor shares were constant, although the Kaldorian theory of distribution or a constant price mark-up pricing theory will give this result. This result went largely unnoticed (except for the few critics of the aggregate production function, notably Shaikh, 1980), although it was published in one of the leading journals, the *Review of Economics and Statistics*.

A jointly written follow-up paper using micro CES production functions came to a similar result (Fisher *et al*., 1977). The aggregated simulated data gave plausible estimates (although not always) of the value of the 'aggregate elasticity of substitution', which in the simulations did not exist. In other words, 'the elasticity of substitution in these production functions is an "estimate" of nothing; there is no "true" aggregate parameter to which it corresponds' (Fisher *et al*., 1977, p. 312). Furthermore, 'the aggregative data themselves do not tell you very clearly whether the estimated parameters are likely to have average meaning or not' (Fisher *et al*., 1977, p. 319). From a rhetorical point of view, the insertion of the words 'very clearly' weakens the main conclusion, which is that the aggregate

data really do not tell us anything at all. And then, in the closing paragraph of the paper, the authors further dilute the force of their analysis:

> For many problems, aggregate production functions are simply too useful to pass up, especially as they can work, as our experiment shows. Our parting advice is to handle them the way the old garbage man tells the young garbage man to handle garbage wrapped in plastic bags of unknown provenance: 'Gingerly, Hector, gingerly'. (Fisher *et al*., 1977, p. 319)

The rhetoric in this passage is at variance with the main thrust of the paper. Reassuringly for the neoclassical economists, it implies that they can continue to use aggregate production functions. But the imperative to handle the results 'gingerly' gives no practical guidance at all: it's business as usual; carry on estimating aggregate production functions.

But what are we to understand by the statement that aggregate production functions 'work'? It is that they have good *predictive* power, which is not the same thing as having good explanatory power. The whole thrust of these two papers by Fisher (1971) and Fisher *et al*. (1977) is to show that aggregate production functions can have predictive power even when no true aggregate production function exists.

6. The Rhetoric of Shaikh's Humbug Production Function and Solow's Reply[13]

A key paper in the development of the accounting identity argument was that of Anwar Shaikh (1974). Solow's (1974) reply provides a good example of how the use of rhetoric can obfuscate rather than illuminate an argument. Shaikh's title grabs the attention: 'Laws of Production and Laws of Algebra: The Humbug Production Function'. In the article he links his subsequent discussion to the Cambridge capital theory controversies, illustrating immediately the potential importance of his argument. He made two major points.

First, he set out the accounting critique in clear terms and applied it, for the first time, to time-series data—earlier articles had confined the critique to cross-sectional data. If the data display constant factor shares, the 'production function' must be a Cobb-Douglas. Secondly, he turned his attention to the procedure that Solow (1957) followed, which is also a tautology, but for different reasons. Solow estimated the Cobb-Douglas specification:

$$\ln(V_t/L_tA_t) = c + b\ln(J_t/L_t) \tag{6}$$

Considering the problems associated with measuring the capital stock, the result was a remarkably tight fit. The reason is that $A(t)$ was calculated from the identity $\dot{A}_t \equiv (\dot{V}_t - \dot{L}_t) - a_t(\dot{J}_t - \dot{L}_t)$ and $\dot{A}_t$ was used to calculate the index A_t. (Solow actually used discrete growth rates.) But as Shaikh points out, as long as factor shares are roughly constant (as was the case with Solow's data) the estimation

[13]This section draws on McCombie (1998b).

of equation (6) must give a perfect fit to the data, due to the way A was calculated. This actually has nothing to do with the accounting identity, but is simply a manifestation of the 'laws of algebra'. Furthermore, when the accounting identity is taken into account, Solow is actually estimating the identity:

$$\ln(V_t/L_t r_t^a w_t^{(1-a)}) \equiv c + b \ln(J_t/L_t) \tag{7}$$

where the estimate of the slope coefficient b is the average value of capital's share, i.e. a.

To drive home the point that with constant shares Solow's procedure of deflating V by $A(t)$ will give an excellent fit to the Cobb-Douglas with any underlying data for V, J and L, Shaikh constructs an artificial data set where the scatter plot of the observations of V/L against J/L traces out the word HUMBUG. Capital's share was the same as in Solow's actual data. Not surprisingly, following Solow's procedure, Shaikh obtains a very good fit to the Cobb-Douglas function; moreover, the estimate of the slope of the regression line equals the average value of capital's share. The format of his note in the *Review of Economics and Statistics* mimics the paper of Solow (1957) published in the same journal, with analogous figures and the data set out in an appendix for the Humbug data in a similar way to Solow's statistics.

In his one-page rejoinder, Solow (1974) begins with the unequivocal statement that 'Mr Shaikh's article is based on misconception pure and simple'. This putative error by Shaikh was that Solow had never intended to *test* the aggregate production function. The 1957 paper 'merely shows how one goes about interpreting given time series if one starts by *assuming* that they were generated by a production function and that competitive-marginal product relations apply' (emphasis in the original). Solow's supposed *coup de grâce* was to estimate a Cobb-Douglas production function without imposing the marginal productivity conditions using the Humbug data (i.e. Solow estimates $\ln(V_t/L_t) = c + b_1 t + b_2 \ln(J_t/L_t)$ and finds no significant statistical relationship. Hence, the 'humbug seems to be on the other foot.' The inescapable conclusion, Solow suggests, is that Shaikh's note is trivial and hardly worth taking seriously.[14]

The editors of the *Review of Economics & Statistics* had accepted Shaikh's paper with ill-disguised discomfort. They insisted that Shaikh revise it down to a brief note, and then declined to allow him the usual opportunity to respond to Solow's criticisms (see Turner, 1989, p. 196, citing Joan Robinson).[15]

Six years later, Shaikh (1980) did publish a compelling reply that raised some further serious questions, but went largely unnoticed: it was published in a book that was not likely to draw the attention of neoclassical economists. If Solow (1974) claimed that all along he was not testing a production function and that

[14]In the span of a single page, Solow applies the following dismissive phrases to Shaikh's argument: 'misconception pure and simple'; 'the point is even simpler'; 'hardly a deep thought'; the 'cute humbug'; 'bowl you over at first'; 'but when you think about it for a minute'; and any 'educated mind'.

[15]Anwar Shaikh has confirmed this in a personal communication of 17 November 2011.

'if the observed factor shares are exactly constant then the method would then yield an exact Cobb-Douglas', then, as Shaikh (1980) pointed out, why would Solow even bother to remark that the 'fit is remarkably tight'? How could it be otherwise? Why did he estimate five different specifications of the production function, including the accounting identity which, although it assumes the wage rate and the profit rate are constant, also gives a very close statistical fit? The very close statistical fits give the reader the misleading impression that the aggregate production function performs well, rather than that it was bound to do so. As for the poor statistical results found by Solow (1974), and noted above, using the Humbug data, this is due to the imposition of the assumption that the rate of technical change is constant. But there is nothing in neoclassical production theory that requires this. Moreover, in Solow's own calculations, the rate of technical change takes a saw-tooth path. Shaikh (1980) showed that a more complex time trend that captured the path of this so-called technical change would both rescue the identity and give a very close fit to the data.

McCombie (2000–01) used irony in his assessment of Solow's rejoinder to Shaikh, the 'trope of tropes' according to McCloskey. McCombie started from Solow's comment that '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 deriving from factor shares; and (b) that the function so fitted has partial derivatives that closely mimic observed factor prices' (Solow, 1974, p. 21). McCombie found that using Solow's data, and repeating the exercise that Solow had done with Shaikh's data, there was no statistically significant relationship and the estimated output elasticities bore no correspondence to the factor shares. In other words, Solow would have been forced to conclude, by his own criterion, that there was no empirical basis for the aggregate production function. This raises the three interesting rhetorical questions. In these circumstances, would Solow have submitted the paper for publication and, if so, would it have been accepted? If this hypothetical paper had been published instead of the original, would the intellectual history of the aggregate production function have been very different?[16]

Solow (1987) returned to these issues, contending that Shaikh's argument could be applied to physical data, and as these could always be used to estimate a production function, it followed that value data could likewise be used. In fact this does not follow; Solow's conclusion is a non-sequitur. He also argued (in our notation) that substituting the equations $r = a(V/J)$ and $w = (1-a)(V/L)$ into equation (3) gives $V = (V/J^aL^{(1-a)})/(J^aL^{(1-a)})$, setting A in the Cobb-Douglas equal to unity and where V on the right-hand side of the equation equals $f(J, L)$. According to Solow (1987, p. 20), 'What Shaikh has discovered ... is that any production function can be written as a product of a Cobb-Douglas and something else; the something else is the production function divided by the Cobb-Douglas'. But this is a classic case of circular reasoning as

[16]It is ironic that the article by McCombie (2000–01) also shows that by using a plausible measure of 'capacity utilisation', a perfect fit using Solow's data can be obtained. However, it will come as no surprise to learn that he used the identity to construct this index and so comes full circle.

far as value data are concerned because it assumes that $V = f(J, L)$ is a production function that can be estimated using value data. The fact that it cannot is at the heart of the critique.

Solow (1987) can be best viewed as trapped within the neoclassical paradigm, with the default position (or the paradigmatic pseudo-assumption) that the aggregate production function is a meaningful representation of the technology of the economy, albeit as an approximation. This is because 'it seems unlikely that the stringent conditions for exact aggregation just happen to be satisfied' (Solow, 1987, p. 26). Solow considers that 'if it works, why does it work? No very good answer has been offered, and I have no startling new suggestion' (Solow, 1987, p. 26). Here he dismisses the whole question of the accounting identity, to which Herbert Simon first drew his attention in the early 1970s (Carter, 2011b). But to accept this critique means completely abandoning the growth models initiated by Solow (1956, 1957) rather than defending them as parables that do give some insights into the underling technological parameters of the economy. It is not hard to understand why Solow prefers the latter approach.

7. Other Fundamental Misconceptions of the Argument

In this section, we discuss some other major misconceptions concerning the problems that aggregation and the accounting identity pose both for the estimation of aggregate production functions and for the interpretation of the obtained coefficients. (These arguments have been put to us in a number of different forums.) All of them suffer from the methodological problems discussed earlier, and they all demonstrate the dominance of the neoclassical paradigm. Kuhn (1970, p. 94, emphasis added) succinctly put the general position as follows:

> When paradigms enter, as they must, into debate about paradigm choice, their role is necessary circular. Each group uses its own paradigm to argue that paradigm's defence ... Yet whatever its force, the status of the circular argument is *only that of persuasion*. It cannot be made logically or even probabilistically compelling for those who refuse to step into the circle.

We noted above that studying economists' rhetoric can lend some insights into why the aggregate production function continues to be widely used. However, the analysis of rhetoric has real limitations. It might help us understand the structure and type of an economic argument, including when those arguments rely upon such illegitimate concepts as *petitio principii*, *ignoratio elenchi*, *argumentum ad hominem*, appeal to authority (if only the paradigm) and so forth. It can also help us understand why a particular argument or model sometimes belatedly becomes persuasive within the paradigm (see the examples discussed in McCloskey, 1998). But where it is singularly unhelpful is in explaining paradigm choice, primarily because of the self-referential problem. In this view, we differ from McCloskey (2008), who writes:

> Rhetoric ... can be good, offering good reasons for believing that the elasticity of substitution between capital and labour in American manufacturing, say, is about 1.0. The good reasons are not confined by syllogism and number. They include good analogy (production is just like a mathematical function) good

authority (Knut Wicksell and Paul Douglas thought this way, too), good symmetry (if mining can be treated as a production function, so should manufacturing).

In the light of the arguments presented above, we cannot evade the question of what are the 'good reasons'. We know from aggregation theory that the concept of the aggregate elasticity of substitution does not exist. Is production really just like a mathematical function? What is the compelling evidence for this? Why is appeal to the authority of Knut Wicksell and Paul Douglas any more convincing than to that of Franklin Fisher, who actually provides logical reasons for the non-existence of the aggregate production function.

Turning to the more specific misconceptions, the first one is that some economists have suggested that the question at stake is merely one of aggregation. Typical of this is the view that aggregation is unavoidable, for without it empirical work becomes impossibly cumbersome: science works with approximations.

However, it is not a question of simply refusing to aggregate. The problem, of course, is that to aggregate physical inputs and outputs requires prices, and so the use of these data to estimate a technological relationship runs into the problems outlined above.[17] In this vein, Temple (2006, 2010) suggests that the problem may be solved by sufficient disaggregation, and cites the work of Jorgenson & Griliches (1967). However, Temple seems to miss the fundamental point that Jorgenson & Griliches worked with value data and so disaggregation of the data does not remove the problem. Jorgenson & Griliches's procedure also requires the explicit assumption of a constant-returns-to-scale production function, competitive markets and the marginal productivity theory of factor pricing. Indeed, they say themselves that their disaggregated data cannot be used to test the marginal productivity theory of distribution (Jorgenson & Griliches, 1967, p. 257, footnote 2).

Some economists express no surprise that the Cobb-Douglas is an approximation of the linear accounting identity. For example, in the equilibrium of the model (that is, the point at which output is maximised subject to the linear cost constraint), the Cobb-Douglas production function and the linear cost constraint have the same slope and the Cobb-Douglas at this point is given as an approximation to the national accounting identity. In this interpretation, the critique is merely restating in another way an equilibrium condition. An unstated postulate in this argument is that the aggregate production function actually exists; in other words this argument assumes what is in dispute or commits the error of *petitio principii*. The economy is also assumed to be in a state of equilibrium and is technically efficient (output is asserted always to 'be maximised subject to a constraint'). This implicitly assumes that there is a well-behaved aggregate production function that can be expressed in physical terms, $Q = f(K, L, t)$ and the accounting identity can be expressed using Euler's equation as $pQ = pf_L L + pf_K K = wL + \rho K$,

[17]This is not to deny that aggregation is often necessary and justifiable, as for example, when we make use of consumption function in macroeconomic analysis. But this sort of behavioural relationship between constant-price expenditure and income is qualitatively different from the case of the production function, which is presumed to reflect a technological relationship between physical outputs and inputs.

where p is the price per unit of homogeneous output (not a price deflator), and ρ is the rental price of the identical capital goods. The erroneous assumption, which dates back to the first paragraph of Cobb & Douglas (1928), is that the constant price value of output is a perfect proxy for the volume of output measured in physical terms and that $pQ \equiv V \equiv wL + rJ$ whereas V, in fact, measures $\sum_i p_i Q_i$ where the subscript i denotes a particular commodity.[18]

Some sceptics of the critique often commit an *ignoratio elenchi*. Papers that have no bearing on the issue are cited to support the assertion that the accounting identity critique applies to a now dated literature and does not take into account recent developments in the estimation of production functions. But as we have shown elsewhere, these cited works fail to avoid the accounting identity problem.[19] Antràs (2004) and Hsieh & Klenow (2009) are two examples of work cited as demonstration of the validity of the aggregate production function. Antràs (2004) finds that assuming Hicks neutral technical change biases the estimation of aggregate production functions towards the Cobb-Douglas. However, this result is not relevant to the critique as it assumes in the first place that the aggregate production function can be estimated, which is what is being disputed. Hsieh & Klenow (2009) assume heterogeneous firms with firm-level production technologies, from which an aggregate total factor productivity measure is constructed. Yet, even a cursory examination of this paper shows that the micro-production functions do not have heterogeneous physical inputs, but instead use constant-price output and capital measures. The 'prices' used to differentiate what is misleadingly termed 'physical productivity' from 'revenue productivity' are plant-specific price deflators, not prices. Industry price deflators are used to calculate the 'revenue productivity'. Hence, their approach suffers from the same problems as the aggregate production function.

A frequently made argument is that the aggregation problem and the Cambridge capital theory controversies cannot be correct because value measures of the aggregate capital stock are used in state-of-the-art growth accounting studies. This, it is argued, explicitly defines the relationship between aggregate inputs and outputs at the country or regional level. Moreover, calculations of total factor productivity growth are routinely made by organisations and official agencies such as the OECD and the US Bureau of Labor Statistics. This is essentially an 'appeal to authority' type of justification, but, as we noted above, not to a single, or even a group of authors, but to the paradigm itself. An implication of this argument is that deficiencies in the way Cobb and Douglas calculated the value of the indices of the capital stock used in their original regressions in 1928 do not constitute a reason to deny the existence of the aggregate production function. However, the use of the perpetual inventory method to estimate the constant-price value of the capital stock neither avoids the aggregation problem, nor vitiates

[18]As we have noted above, the argument follows through for the case of gross output, where intermediate inputs are included.

[19]In Felipe & McCombie (2002, 2005b, 2009) we deal respectively with Hall's (1988) estimation of the mark-up, Mankiw *et al.*'s (1992) test of the augmented Solow growth model, and estimates of the labour demand function (Hamermesh, 1993).

the problem of the accounting identity, which involves the inappropriate use of value instead of physical data.

Kincaid (2009) has brought a philosopher's eye to the methodology underlying neoclassical growth theory and provides a clear discussion of the nuances of causality in explaining economic growth within the neoclassical paradigm. Perhaps a philosopher of science, untrammelled by schooling in any particular economic paradigm, is the nearest we can get to Kuhn's 'uncommitted observer'. Kincaid is well aware of the capital theory controversies and the problems of aggregation and measuring capital. This leads him to conclude that 'there are serious doubts about neoclassical growth models in so far as they are supposed to be providing a causal explanation based on marginal productivities' (Kincaid, 2009, p. 466). Kincaid continues 'if the equations tested are just supporting the claim that there is some causal contribution and causal relationship from the *quantity* of capital a country has at its disposal and the total output, then the doubts are less worrisome' (Kincaid, 2009, p. 466, emphasis added). This is correct to the extent that even a brief visit to a developing country will provide evidence that the observed *physical* capital stocks (the physical infrastructure, the types and quantities of machinery and the skills of teachers and their resources) are fewer and of poorer quality than in the advanced countries.

However, this does not mean that fitting an aggregate production function using value data can tell us anything more than this casual empiricism. It cannot distinguish between the contributions to productivity growth made by technical progress and capital accumulation. The usual statistical diagnostics (such as the R^2s, t-ratios, and tests for unit roots) are of no avail in this case. No economic hypothesis is being tested. The accounting argument is that there is no point in estimating relationships based on the so-called 'aggregate production function'. Kincaid (2009, p. 473, note 4) asks: 'how do these aggregation issues bear on [growth accounting] and its evidential value?' Our answer is it that it has no evidential value. In fact, from the identity alone, it can be simply shown that the growth of so-called technical change must account for about three-quarters of productivity growth—a result Solow (1988) found 'startling', which itself is startling.[20]

Finally, Temple (2006, 2010) has presented a considered and detailed assessment of the critique, and while he accepts that it has some validity and needs to be more widely known, he argues that too much is made of it. There are a number of serious misconceptions in Temple's argument (see Felipe & McCombie, 2010, 2011b, 2013, for detailed assessments). First, as noted above, Temple believes that disaggregation will overcome the problem. As we have seen, this is not the case as long as value data are used. Secondly, Temple assumes that the critique requires a number of *ad hoc* assumptions, such as constant factor shares and

[20]If we take the stylised facts that factor shares are roughly constant and the capital-output ratio does not greatly change over time (these two stylised facts do not depend upon the existence of an aggregate production function) then, from the identity, the growth of total factor productivity ('technical change') definitionally equals the growth of the real wage rate multiplied by its factor share. From this, it follows that technical change accounts for, and is definitionally equal to, about 75% of the growth of labour productivity.

that the growth of the weighted wage rate and rate of profit is constant over time. Consequently, he considers that it only applies to the spurious Cobb-Douglas function. Related to this, he seems, at times, to believe that it is simply an identification problem. For example, he argues that if value data were generated by a 'stable production function' and the researcher were capable of controlling for total factor productivity, then there is no reason for 'the parameters to be unidentified'. Moreover, he argues that the 'dynamic version of the value added identity cannot do any better than this model' (Temple, 2010, p. 688). In fact, the converse is true: any putative production function expressed in growth rate form cannot do better than the dynamic version of the identity. But as our discussion of the critique has shown, it is not fundamentally an identification problem. Temple, like other critics, is guilty here of a *petitio principii*. In other words, he *assumes* the existence of the aggregate production function that can be estimated with value data in defence of the criticism that it cannot.

Temple, nevertheless, accepts an important implication of the criticism when he acknowledges that

> the argument shows that an applied researcher may appear to obtain meaningful results from estimating a production relationship, even when the researcher is making assumptions that do not hold in the data. One important instance arises when factors are not paid their marginal products. In that case, although researchers often interpret their results as if the estimated parameters can be used to derive output elasticities, the identity suggests that the estimates may be more closely related to factor shares. (Temple, 2010, p. 686)

This comes perilously close to conceding the validity of the critique.

8. Conclusions

This paper has discussed why some fundamental questions concerning the aggregate production function have been completely overlooked by neoclassical economists. It has examined some key arguments that show that the notion of an aggregate production function is deeply flawed. Although this critique is logically correct, it has failed to persuade mainstream economists. We have shown that this is due to the problem of weak incommensurability, which seems to constrain some neoclassical economists from perceiving that it is, in fact, a question of logic and not of plausibility. Many neoclassical criticisms of the accounting identity critique that we have encountered start from the premise that the aggregate production function does exist. In this sense, the existence of the aggregate production function may be regarded as a neoclassical 'paradigmatic pseudo-assumption'.

The term pseudo-assumption is used by Kuhn because, in the natural sciences, these assumptions are hybrid analytic-synthetic, or quasi-analytic, statements. Such assumptions are analytic because they are deemed not falsifiable by fiat. They are taken as self-evident and demarcate the paradigm. They are synthetic in that they may initially have been part of the empirical basis of the paradigm, but 'they are by no means the product of arbitrary definitional stipulations. They are rather in part the product of painstaking empirical and theoretical research' (Hoyningen-Huene, 1993, p. 210). The difference here is that the

myriad of empirical estimations of aggregate production functions can give no support for the hypothesis that the latter exist. Hence, weak incommensurability presents difficulties for economists in addressing the accounting identity critique. Not for nothing did Keynes (1936, p. viii) comment on the difficulty of a 'long struggle of escape . . . from habitual modes of thought and expression'.

The argument is sometimes made that one should persevere with the aggregate production function because there is no alternative, but this is not compelling. The aggregate production function is not an approximation that economists can be expected to improve upon over time with the development of new data sets and estimation techniques. It is a concept that is fundamentally theoretically flawed: it is 'not even wrong' in that it cannot be empirically tested. The 'there is no alternative' defence merely serves to perpetuate an error of logic. Nevertheless, the aggregate production function has become so well-established in the neoclassical paradigm that it is assumed to exist by fiat. Perhaps the fact that the specific arguments raised by the critique are never addressed, but simply ignored is not surprising given their damaging implications for significant portions of neoclassical macroeconomic theory that rely on the aggregate production function.

Acknowledgment and Disclaimer

The authors are grateful to an anonymous referee for helpful comments. The views in this paper are the authors' own, and not those of the Asian Development Bank, its member countries, or those of its Executive Directors.

References

Aghion, P. & Howitt, P. (2009) *The Economics of Growth* (Cambridge, MA: MIT Press).

Antràs, P. (2004) Is the US aggregate production function Cobb-Douglas? New estimates of the elasticity of substitution, *B.E. Journal of Macroeconomics, Contributions to Macroeconomics*, 4, pp. 1–34. Available at http://scholar.harvard.edu/files/antras/files/cespublished.pdf

Arrow, K.J., Bernheim, B.D., Feldstein, M.S., McFadden, D.L., Poterba, J.M. & Solow, R.M. (2011) 100 years of the American economic review: The top 20 articles, *American Economic Review*, 101, pp. 1–8.

Barro, R. & Sala-i-Martin, X. (2003) *Economic Growth*, 2nd edn (Cambridge, MA: MIT Press).

Blundell, R. & Bond, S. (2000) GMM estimation with persistent panel data: an application to production functions, *Econometric Reviews*, 19, pp. 321–340.

Carter, S. (2011a) C. E. Ferguson and the neoclassical theory of capital: a matter of faith, *Review of Political Economy*, 23, pp. 339–356.

Carter, S. (2011b) 'On the Cobb-Douglas and all that . . . ': The Solow–Simon correspondence over the aggregate neoclassical production function, *Journal of Post Keynesian Economics*, 34, pp. 255–274.

Clark, J.B. (1891) Distribution as determined by the law of rent, *Quarterly of Journal of Economics*, 5, pp. 289–318.

Cobb, C.W. & Douglas, P.H. (1928) A theory of production, *American Economic Review*, 18, pp. 139–165.

Cohen, A. & Harcourt, G.C. (2003) Whatever happened to the Cambridge capital theory controversies?, *Journal of Economic Perspectives*, 17, pp. 199–214.

Cohen, A. & Harcourt, G.C. (2005) Capital theory controversy: scarcity, production, equilibrium and time, in: C. Bliss, A. Cohen & G.C. Harcourt (Eds) *Capital Theory*, Vol. 1 (Northampton: Edward Elgar).

Estrin, S. & Laidler, D.E.W. (1995) *Introduction to Microeconomics*, 4th edn (London: Harvester-Wheatsheaf).
Felipe, J. & Fisher, F.M. (2003) Aggregation in production functions: what applied economists should know, *Metroeconomica*, 54, pp. 208–262.
Felipe, J. & McCombie, J.S.L. (2002) A problem with some recent estimations and interpretations of the markup in manufacturing industry, *International Review of Applied Economics*, 16, pp. 187–215.
Felipe, J. & McCombie, J.S.L. (2005a) How sound are the foundations of the aggregate production function, *Eastern Economic Journal*, 31, pp. 467–488.
Felipe, J. & McCombie, J.S.L. (2005b) Why are some countries richer than others? A sceptical view of Mankiw-Romer-Weil's test of the neoclassical growth model, *Metroeconomica*, 56, pp. 360–392.
Felipe, J. & McCombie, J.S.L. (2009) Are estimates of labour demand functions mere statistical artefacts?, *International Review of Applied Economics*, 23(2), pp. 147–168.
Felipe, J. & McCombie, J.S.L. (2010) What is wrong with aggregate production functions. On Temple's 'Aggregate production functions and growth economics', *International Review of Applied Economics*, 24, pp. 665–684.
Felipe, J. & McCombie, J.S.L. (2011a) On Herbert Simon's criticisms of the Cobb-Douglas and the CES production functions, *Journal of Post Keynesian Economics*, 34, pp. 275–294.
Felipe, J. & McCombie, J.S.L. (2011b) The tyranny of the accounting identity works full-time. A further rejoinder to Temple, Cambridge Centre for Economic and Public Policy, University of Cambridge.
Felipe, J. & McCombie, J.S.L. (2013) *The Aggregate Production Function and the Measurement of Technical Change: 'Not Even Wrong'* (Cheltenham: Edward Elgar).
Ferguson, C.E. (1969) *The Neoclassical Theory of Production and Distribution* (Cambridge: Cambridge University Press).
Ferguson, C.E. (1971) Capital theory up to date: a comment on Mrs Robinson's article, *Canadian Journal of Economics*, 4, pp. 250–254.
Fisher, F.M. (1971) Aggregate production functions and the explanation of wages: a simulation experiment, *Review of Economics and Statistics*, 53, pp. 305–325.
Fisher, F.M. (2005) Aggregate production functions—A pervasive, but unpersuasive, fairytale, *Eastern Economic Journal*, 31, pp. 489–491.
Fisher, F.M., Solow, R.M. & Kearl, J.M. (1977) Aggregate production functions: some CES experiments, *Review of Economic Studies*, 44, pp. 305–320.
Friedman, M. (1953) The methodology of positive economics, in: M. Friedman (Ed) *Essays in Positive Economics* (Chicago: University of Chicago Press).
Griliches, Z. & Mairesse, J. (1998) Production functions: the search for identification, in: S. Strøm (Ed) *Econometrics and Economic Theory in the 20th Century* (Cambridge: Cambridge University Press).
Hall, R.E. (1988) The relation between price and marginal cost in US industry, *Journal of Political Economy*, 96, pp. 921–947.
Hamermesh, D.S. (1993) *Labor Demand* (Princeton: Princeton University Press).
Harcourt, G.C. (1976) The Cambridge controversies: Old ways and new horizons—or dead end, *Oxford Economic Papers*, 28, pp. 25–65.
Hildebrand, G. & Liu, T.C. (1965) *Manufacturing Production Functions in the United States, 1957* (Cornell: Cornell University Press).
Hoover, K.D. & Siegler, M.V. (2008a) Sound and fury: McCloskey and significance testing in economics, *Journal of Economic Methodology*, 15, pp. 1–37.
Hoover, K.D. & Siegler, M.V. (2008b) The rhetoric of 'signifying nothing': a rejoinder to Ziliak and McCloskey, *Journal of Economic Methodology*, 15, pp. 57–68.
Hoyningen-Huene, P. (1993) *Reconstructing Scientific Revolutions. Thomas S. Kuhn's Philosophy of Science* (Chicago: Chicago University Press).
Hsieh, C.-T. & Klenow, P.J. (2009) Misallocation and TFP in China and India, *Quarterly Journal of Economics*, 124, pp. 1403–1448.
Jorgenson, D.W. & Griliches, Z. (1967) The explanation of productivity change, *Review of Economic Studies*, 34, pp. 249–283.

Keynes, J.M. (1936) *The General Theory of Employment, Interest and Money* (London: Macmillan).
Kincaid, H. (2009) Explaining growth, in: H. Kincaid & D. Ross (Eds) *The Oxford Handbook of Philosophy of Economics* (Oxford: Oxford University Press).
Kuhn, T.S. (1970) *The Structure of Scientific Revolutions*, 2nd edn (Chicago: Chicago University Press).
Kuhn, T.S. (1999) Remarks on incommensurability and translation, in: R.R. Favretti, G. Sandri & R. Scazzeri (Eds) *Incommensurability and Translation. Kuhnian Perspectives on Scientific Communication and Theory Change* (Cheltenham: Edward Elgar).
Levinsohn, J. & Petrin, A. (2003) Estimating production functions using inputs to control for unobservables, *Review of Economic Studies*, 70, pp. 317–341.
Mankiw, N.G. Romer, D. & Weil, D.N. (1992) A contribution to the empirics of economic growth, *Quarterly Journal of Economics*, 107, pp. 407–437.
McCloskey, D.N. (1994) *Knowledge and Persuasion in Economic* (Cambridge: Cambridge University Press).
McCloskey, D.N. (1997) *The Vices of Economists; The Virtues of the Bourgeoisie* (Amsterdam: Amsterdam University Press).
McCloskey, D.N. (1998) *The Rhetoric of Economics*, 2nd edn (Madison: University of Wisconsin Press).
McCloskey, D.N. (2008) Rhetoric of economics, in: S.N. Durlauf & L.E. Blume (Eds) *The New Palgrave Dictionary of Economics Online* (Palgrave Macmillan) [http://www.dictionaryofeconomics.com/article?id=pde2008_R000142].
McCloskey, D.N. & Ziliak, S.T. (2008) Signifying nothing: reply to Hoover and Siegler, *Journal of Economic Methodology*, 15, pp. 39–55.
McCombie, J.S.L. (1998a) 'Are there laws of production?': an assessment of the early criticisms of the Cobb-Douglas production function, *Review of Political Economy*, 10, pp. 141–173.
McCombie, J.S.L. (1998b) Rhetoric, paradigms, and the relevance of the aggregate production function, in: P. Arestis & M.C. Sawyer (Eds) *Method, Theory, and Policy in Keynes. Essays in Honour of Paul Davidson*, Vol. III, (Aldershot: Edward Elgar).
McCombie, J.S.L. (2000–01) The Solow residual, technical change and aggregate production functions, *Journal of Post Keynesian Economics*, 23, pp. 267–297 [Errata, Vol. 23, p. 544].
McCombie, J.S.L. (2011) 'Cantabrigian economics' and the aggregate production function, *Intervention*, 8, pp. 165–182.
Mirowski, P. (1987) Shall I compare thee to a Minkowski-Ricardo-Leontief-Metzler matrix of the Mosak-Hicks type?: or, rhetoric, mathematics and the nature of neoclassical theory, *Economics and Philosophy*, 3, pp. 67–95.
Moroney, J.R. (1972) *The Structure of Production in American Industry* (Chapel Hill: University of North Carolina Press).
Olley, S. & Pakes, A. (1966) The dynamics of productivity in the telecommunications equipment industry, *Econometrica*, 64, pp. 1263–1298.
Phelps Brown, E.H. (1957) The meaning of the fitted Cobb-Douglas function, *Quarterly Journal of Economics*, 71, pp. 546–560.
Samuelson, P.A. (1966) A summing up, *Quarterly Journal of Economics*, 80, pp. 568–583.
Samuelson, P.A. (1979) Paul Douglas's measurement of production functions and marginal productivities, *Journal of Political Economy*, 87, pp. 923–939.
Shaikh, A. (1974) Laws of production and laws of algebra, the humbug production function, *Review of Economics and Statistics*, 56, pp. 115–120.
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).
Simon, H.A. (1978) Herbert A. Simon—Biographical [http://www.nobelprize.org/nobel_prizes/economics/laureates/1978/simon-autobio.html].
Simon, H.A. (1979a) On parsimonious explanations of production relations, *Scandinavian Journal of Economics*, 81, pp. 459–474.
Simon, H.A. (1979b) Rational decision-making in business organizations, *American Economic Review*, 69, pp. 459–474.

Simon, H.A. & Levy, F.K. (1963) A note on the Cobb-Douglas function, *Review of Economic Studies*, 30, pp. 93–94.
Solow, R.M. (1956) A contribution to the theory of economic growth, *Quarterly Journal of Economics*, 70, pp. 65–94.
Solow, R.M. (1957) Technical change and the aggregate production function, *Review of Economics and Statistics*, 39, pp. 312–320.
Solow, R.M. (1974) Laws of production and laws of algebra: the humbug production function: a comment, *Review of Economics and Statistics*, 56, p. 121.
Solow, R.M. (1987) Second thoughts on growth theory, in: A. Steinherr & D. Weiserbs (Eds) *Employment and Growth: Issues for the 1980s* (Dordrecht: Martinus Nijhoff).
Solow, R.M. (1988) Growth theory and after, *American Economic Review*, 78, pp. 307–317.
Summers, L. (1991) The scientific illusion in empirical macroeconomics, *Scandinavian Journal of Economics*, 93, pp. 129–148.
Temple, J.R.W. (2006) Aggregate production functions and growth economics, *International Review of Applied Economics*, 20, pp. 301–317.
Temple, J.R.W. (2010) Aggregate production functions, growth economics, and the part-time tyranny of the identity: a reply to Felipe and McCombie, *International Review of Applied Economics*, 24, pp. 685–692.
Turner, M.J. (1989) *Joan Robinson and the Americans* (Armonk: M. E. Sharpe).
Walters, A.A. (1963) Production and cost functions. An econometric survey. *Econometrica*, 31, pp. 1–66.
Woit, P. (2006) *Not Even Wrong* (London: Jonathan Cape).
Ziliak, S.T. & McCloskey, D.N. (2008) *The Cult of Statistical Significance. How the Standard Error Cost Us Jobs, Justice and Lives* (Ann Arbor: University of Michigan Press).