



Why is the Cobb–Douglas production function so popular?

Bert M. Balk¹

Received: 1 February 2024 / Accepted: 29 February 2024 / Published online: 5 April 2024
© The Author(s) 2024

Abstract

It is well known that, in continuous time, the Cobb–Douglas function can be derived from the underlying, data governing, accounting identity under some reasonable assumptions (factor shares are constant, and the weighted growth of the labour input price and the capital input price is constant). In this article these results are generalized in three ways: (1) the accounting identity contains a (pure) profit term; (2) continuous time is replaced by discrete time periods; (3) additional assumptions appear to be superfluous. The article also discusses extensions: from two to multiple inputs, from value added to gross output, and from a single production unit to an ensemble of those units.

Keywords Productivity · Production function · Cobb–Douglas · Index number theory

JEL Classification C43 · D24 · E01 · E23

1 Introduction

One of the famous workhorses in economic theory, used in uncountable many applications at all levels of aggregation (economy, industry, enterprise, plant, or whatever), despite its limitations, is the Cobb–Douglas (CD) (production) function.¹ Its

¹ Two examples from recent handbooks: According to Chambers and Ray (2021, 22) “The Cobb–Douglas production function remains a classic example of empirical evidence inspiring a theoretical formulation of a production function that has served as the gold standard in neoclassical production

A part of this article was initially presented at the 12th European Workshop on Efficiency and Productivity Analysis (EWEPA), Verona, 22 June 2011. A more complete version was presented at the 17th EWEPA, Porto, 29 June 2022. The author thanks Jesus Felipe and Erwin Diewert for providing comments on these earlier versions.

✉ Bert M. Balk
bert.m.balk@outlook.com

¹ Rotterdam School of Management, Erasmus University Rotterdam, Rotterdam, The Netherlands

‘discovery’ was reported by Cobb and Douglas (1928)² and the subsequent success story told in Douglas’s presidential address at the 1947 meeting of the American Economic Association, as reported in Douglas (1948). The years after 1947 were covered in a posthumous article by Douglas (1976) — both articles are still fascinating pieces to read. The colourful history from its start in the mid-1920s to the late 1960s is to be found in the recent book by Biddle (2021).

The universal ‘law’ one believed to have uncovered – Biddle (2021) makes us aware of the fact that the concept of ‘production function’ only gradually crept into the language used – relates production P , measured as an index number of real (*i.e.*, deflated) value added, to an index number of labour input L and an index number of capital stock K by means of the function $P = aL^\lambda K^{1-\lambda}$, where a is some constant. Based on empirical evidence the labour exponent λ was set equal to 0.75, whereas a appeared to be approximately equal to 1. In time-series as well as cross-sectional studies of the manufacturing industry at various levels of aggregation time and again it appeared that λ was remarkably well approximated by the (nominal) share of labour cost in value added (which in turn was conceived as the sum of labour and capital input cost). In the 1947 address Douglas considered the slightly more general formula $P = aL^\lambda K^\mu$, but then more often than not it turned out that the sum of the exponents λ and μ was close to 1, which was taken as signalling constant returns to scale. Quoting Douglas (1976, 914):

“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.”

In a relatively unknown article, Tinbergen (1942a) pointed out a number of restrictive implications of the CD function as well as a number of statistical estimation problems, especially in a cross-section context. One of these implications is that labour productivity, P/L , is related to capital intensity, K/P , by $P/L = a^{1/\lambda}(K/P)^{(1-\lambda)/\lambda}$; for $a = 1$ and $\lambda = 0.75$ this means that $P/L = (K/P)^{1/3}$, which implies that in a time-series context the increase of capital intensity would be the sole driving force for the increase of labour productivity. Where is technological progress?

It was more or less customary among economists of those days to see technological progress as part of accumulation of the capital stock and/or as increase of the quality of labour. Tinbergen however suggested to introduce time as an additional variable in the production function; for instance, by replacing the scalar a by the simple function a^t , as in Tinbergen (1942b). This can be seen as the beginning of a development in which the simple function a^t is replaced by a more intricate function

Footnote 1 (continued)

economic theory for decades and has retained much of its popular appeal despite the advent of more flexible functional forms even as it nears its centenary.” De Loecker and Syverson (2021, 161, 193) say “When working with production functions, we usually use the Cobb-Douglas form. It conveys most of the necessary intuition and makes notation easier.” and “The predominant functional form for production functions in applied work is Cobb-Douglas.”

² The mathematical form of the CD function had already materialized in some Swedish and German publications of Wicksell between 1900 and 1923, as documented by Sandelin (1976).

$A(t)$, defining ‘total factor productivity’, or ‘efficiency’, or ‘state of technology’, as in Solow (1957).

From the start, the project of Douglas and his followers has met with approval and criticism, both from econometric and conceptual perspectives, as documented extensively in Biddle (2021) and Felipe and McCombie (2013, Chapter 4).³ Many empirical applications took place in agricultural economics because of the availability of more and better (firm-level) data; here estimation could be based on less dubious assumptions.

In the meantime theoretical work showed that the conditions for the existence of an aggregate production function are unlikely to being satisfied in any realistic situation — see the lemma on “Aggregation (production)” in *The New Palgrave* by Felipe and Fisher (2008). But then: why always the good fit? Is there a sort of law of production that supports the input–output data, or is the observed, remarkably stable, relationship between those data some statistical artefact? But what precisely causes this artefact then to emerge?

Important contributions to the solution of this puzzle were provided first by Phelps Brown (1957),⁴ followed by Simon and Levy (1963), Fisher (1971), Shaikh (1974), and again Simon (1979). It gradually began to dawn that the responsibility rests with the underlying accounting relation stating that (nominal) value added equals (nominal) labour input cost and (nominal) capital input cost. Simon (1979, 473) concluded by saying that

“Since the observed phenomena can as readily be explained on the weaker assumption that what is being observed is simply the accounting relation equating value of output to the sum of factor costs, the criterion of parsimony would lead us to prefer the latter explanation to the classical one.”

In an article commemorating Douglas’s work, Samuelson (1979, 933) also pointed to the role of “the accounting identity involved in the residual definition of profit”.

Further evidence, approaching a body of data with a CES production function, of which the CD function is a special case, was provided by Felipe and McCombie (2001). Their conclusion was that

“... Occam’s razor suggests that the CES production function,..., merely reflect the underlying accounting identity that value added equals the wage bill plus profits.”

³ Another restrictive implication of the CD production function is that the elasticity of substitution $\sigma \equiv d \ln(K/L) / d \ln((\partial P / \partial L) / (\partial P / \partial K)) = 1$. To remedy this, the CES function was invented in 1961. Knoblauch et al. (2020) considered 77 studies, published during the years 1961–2017, for the US economy and/or its industries. In all these studies the CES production function was used as basis for estimation. A meta-regression suggested that “a CD production function is unlikely to be a good representation of the US economy.” The majority of the studies suggested that $\sigma < 1$. A more recent meta-analysis with broader scope, reported by Gechert et al. (2022), covered 121 studies. It turned out that removal of publication bias let the mean estimate of σ drop from 0.9 to 0.5.

⁴ On Phelps Brown’s critique, see Biddle (2021, 148–155.)

Interesting insights can be found in a collection of symposium papers published in the Summer 2005 issue of the *Eastern Economic Journal*; in particular Felipe and Adams (2005), Shaikh (2005), and Felipe and McCombie (2005). Summaries of the state of the art were provided by Felipe and McCombie (2007) and Felipe and Fisher (2008), and much additional material by Felipe and McCombie (2013). Especially the final chapter of this book, entitled “Why have criticisms of the aggregate production function generally been ignored? On further misunderstandings and misinterpretations of the implications of the accounting identity”, is well worth (re-) reading.

Basically in all these publications it is shown that, in continuous time, the CD function can be derived from the accounting identity under some reasonable assumptions (factor shares are constant, and the weighted growth of the labour input price and the capital input price is constant). In this paper these results are generalized in three ways: (1) the accounting identity contains a (pure) profit term; (2) continuous time is replaced by discrete time periods; (3) additional assumptions appear to be superfluous.

The layout of this article is as follows. Section 2 first provides a sketch of the classic argument, and then shows in a more rigorous way how a CD function emerges from observable data. Section 3 provides a number of comments and extensions. Section 4 discusses the generalization from real value added to gross output. Section 5 goes beyond a single production unit and considers an ensemble of those. Section 6 contains some concluding thoughts.

2 The emergence of a function

Consider a production unit, operating on the market so that output prices are given (or can be imputed). Let t denote an accounting period of a certain length (say, a year). The *ex post* accounting relation according to the KL-VA model reads

$$C_{KL}^t + \Pi^t = VA^t, \quad (1)$$

where C_{KL}^t denotes (nominal) primary inputs cost, VA^t denotes (nominal) value added (that is, revenue minus intermediate inputs cost; assumed to be positive), and profit Π^t is defined as the difference between value added and primary inputs cost. Thus the accounting relation is merely a definitional identity. Profit may be positive, zero, or negative. For a more detailed exposition the reader is referred to Balk (2021, Chapter 2). In particular one should note that revenue, intermediate inputs cost, capital input cost, and labour input cost are sums of prices times quantities. Usually the number of commodities (goods and services) making out each of these aggregates is huge. Data are therefore usually presented in the form of aggregate nominal values and corresponding price index numbers.

The reader is warned that there is no standardized terminology here; *e.g.*, Felipe and Adams (2005) call value added minus labour cost, in our notation $VA^t - C_L^t$, total profits, which is then seen as remuneration of the capital stock.

Primary inputs concern owned capital and labour (including self-employed persons). Since nominal values are additive, we have, in obvious notation, that

$$C_{KL}^t = C_K^t + C_L^t. \tag{2}$$

We consider two periods, which may or may not be adjacent. These periods are labeled 0 (base period) and 1 (comparison period).⁵ Let us start by summarizing the classic argument.

2.1 The classic argument

The *first* assumption is that profit equals zero, $\Pi^t = 0$. Then, by combining expressions (1) and (2), the fundamental accounting relation reads

$$C_K^t + C_L^t = VA^t. \tag{3}$$

The three elements in this equation are *nominal* values. In order to turn them into *real* values we need price indices (*aka* deflators). Let $P_K(t, b)$, $P_L(t, b)$, and $P_{VA}(t, b)$ be suitable deflators for capital cost, labour cost, and value added, respectively, for period t relative to a certain reference period b (somewhere in the past); that is, $P_K(b, b) = P_L(b, b) = P_{VA}(b, b) = 1$. Real values $X_K(t, b)$, $X_L(t, b)$, and $RVA(t, b)$ are then defined by the three identities

$$C_K^t = P_K(t, b)X_K(t, b), \tag{4}$$

$$C_L^t = P_L(t, b)X_L(t, b), \tag{5}$$

$$VA^t = P_{VA}(t, b)RVA(t, b). \tag{6}$$

The notation highlights the fact that unlike nominal values, which are observable, real values are outcomes of functions. Relative deflators for capital and labour cost, respectively, are defined by

$$\tilde{P}_K(t, b) \equiv P_K(t, b)/P_{VA}(t, b) \tag{7}$$

$$\tilde{P}_L(t, b) \equiv P_L(t, b)/P_{VA}(t, b). \tag{8}$$

By inserting expressions (4), (5), and (6) into expression (3), dividing both sides of the resulting identity by $P_{VA}(t, b)$, and using expressions (7) and (8), we see that the fundamental accounting relation in terms of real values reads

$$\tilde{P}_K(t, b)X_K(t, b) + \tilde{P}_L(t, b)X_L(t, b) = RVA(t, b). \tag{9}$$

The *second* assumption is that all the functions occurring in expression (9) are continuously differentiable functions of continuous time t . By differentiating both sides of this expression we then obtain the following identity,

⁵ Though this paper is framed in terms of temporal (time-series) comparisons, it is straightforward to reframe the content in terms of spatial (cross-sectional) comparisons.

$$\begin{aligned} & \tilde{P}_K(t, b)dX_K(t, b) + \tilde{P}_L(t, b)dX_L(t, b) + X_K(t, b)d\tilde{P}_K(t, b) + X_L(t, b)d\tilde{P}_L(t, b) \\ & = dRVA(t, b), \end{aligned} \quad (10)$$

which can be rewritten as

$$\begin{aligned} & \frac{\tilde{P}_K(t, b)X_K(t, b)}{RVA(t, b)}d \ln X_K(t, b) + \frac{\tilde{P}_L(t, b)X_L(t, b)}{RVA(t, b)}d \ln X_L(t, b) \\ & + \frac{\tilde{P}_K(t, b)X_K(t, b)}{RVA(t, b)}d \ln \tilde{P}_K(t, b) + \frac{\tilde{P}_L(t, b)X_L(t, b)}{RVA(t, b)}d \ln \tilde{P}_L(t, b) \\ & = d \ln RVA(t, b). \end{aligned} \quad (11)$$

The *third* assumption is that the two factor shares are constant; that is, using expressions (4)–(8),

$$\frac{\tilde{P}_K(t, b)X_K(t, b)}{RVA(t, b)} = \frac{P_K(t, b)X_K(t, b)}{VA^t} = \frac{C_K^t}{VA^t} = \alpha_K \quad (12)$$

$$\frac{\tilde{P}_L(t, b)X_L(t, b)}{RVA(t, b)} = \frac{P_L(t, b)X_L(t, b)}{VA^t} = \frac{C_L^t}{VA^t} = \alpha_L. \quad (13)$$

The fundamental accounting relation in expression (3) then implies that $\alpha_K + \alpha_L = 1$, which in turn implies that expression (11) reduces to

$$\begin{aligned} & \alpha_K d \ln X_K(t, b) + (1 - \alpha_K) d \ln X_L(t, b) \\ & + \alpha_K d \ln \tilde{P}_K(t, b) + (1 - \alpha_K) d \ln \tilde{P}_L(t, b) \\ & = d \ln RVA(t, b). \end{aligned} \quad (14)$$

The *fourth*, and final, assumption is that the weighted mean of the relative deflators is constant; that is,

$$\alpha_K d \ln \tilde{P}_K(t, b) + (1 - \alpha_K) d \ln \tilde{P}_L(t, b) = \beta. \quad (15)$$

The consequence is that expression (14) reduces to

$$d \ln RVA(t, b) = d \ln (X_K(t, b)^{\alpha_K} X_L(t, b)^{1-\alpha_K}) + \beta. \quad (16)$$

On the right-hand side of this identity, between the brackets, we meet the traditional CD function. It is interesting to notice that the fourth assumption basically means that the rate of value-added based price change, $d \ln P_{VA}(t, b)$, differs from the rate of capital plus labour price change, $d \ln (P_K(t, b)^{\alpha_K} P_L(t, b)^{1-\alpha_K})$, by a fixed percentage. Another interpretation is that dual total factor productivity change, as introduced by Jorgenson and Griliches (1967, 252), is constant.

This concludes the reproduction of the classic argument: if the *nominal* data are governed by an accounting identity like expression (3) then a CD function will almost necessarily give a good ‘fit’ to the *real* data, because the relation between these input–output data exhibits a CD-like form. In the next subsection it

will be shown that the first three assumptions of the classic approach are unnecessary, and that the fourth assumption can be replaced by a more modest one.

2.2 A more rigorous approach

We drop all the assumptions made in the previous subsection. Specifically, we return to discrete time periods. It is assumed that the (nominal) value-added ratio of period 1 to period 0 can be decomposed into a price index and a quantity index,

$$\frac{VA^1}{VA^0} = P_{VA}(1, 0)Q_{VA}(1, 0), \tag{17}$$

where (1, 0) is used as shorthand notation for the vectors of prices and quantities of the two periods. Likewise, it is assumed that the joint primary inputs cost ratio is decomposed as

$$\frac{C^1_{KL}}{C^0_{KL}} = P_{KL}(1, 0)Q_{KL}(1, 0). \tag{18}$$

As in Balk (2021, 34) the value-added based (total factor) productivity index for period 1 relative to period 0 is then defined as

$$ITFP_{VA}(1, 0) \equiv \frac{Q_{VA}(1, 0)}{Q_{KL}(1, 0)}. \tag{19}$$

This index can be interpreted as the ‘quantity’ change of value added relative to the joint quantity change of primary inputs; or, as the index of real value added relative to the index of real primary inputs; or, as the real component of value-added-based profitability change.

Using the logarithmic mean.⁶ and the relation in expression (2), the primary inputs cost ratio can be decomposed as follows:

$$\begin{aligned} \ln\left(\frac{C^1_{KL}}{C^0_{KL}}\right) &= \frac{C^1_{KL} - C^0_{KL}}{LM(C^1_{KL}, C^0_{KL})} \\ &= \frac{C^1_K - C^0_K}{LM(C^1_{KL}, C^0_{KL})} + \frac{C^1_L - C^0_L}{LM(C^1_{KL}, C^0_{KL})} \\ &= \frac{LM(C^1_K, C^0_K) \ln(C^1_K/C^0_K)}{LM(C^1_{KL}, C^0_{KL})} + \frac{LM(C^1_L, C^0_L) \ln(C^1_L/C^0_L)}{LM(C^1_{KL}, C^0_{KL})}. \end{aligned} \tag{20}$$

⁶ For any two strictly positive real numbers a and b , the logarithmic mean is defined by $LM(a, b) \equiv (a - b) / \ln(a/b)$ if $a \neq b$ and $LM(a, a) \equiv a$. The logarithmic mean can be used to convert a difference $a - b$ into (the logarithm of) a ratio a/b , and *vice versa*. It has the following properties: (1) $\min(a, b) \leq LM(a, b) \leq \max(a, b)$; (2) $LM(a, b)$ is continuous; (3) $LM(\lambda a, \lambda b) = \lambda LM(a, b)$ ($\lambda > 0$); (4) $LM(a, b) = LM(b, a)$; (5) $(ab)^{1/2} \leq LM(a, b) \leq (a + b)/2$; (6) $LM(a, 1)$ is concave. See Balk (2008, 134–6) for details.

It is useful to define $\phi_K(1, 0) \equiv LM(C_K^1, C_K^0)/LM(C_{KL}^1, C_{KL}^0)$ and $\phi_L(1, 0) \equiv LM(C_L^1, C_L^0)/LM(C_{KL}^1, C_{KL}^0)$. Thus $\phi_K(1, 0)$ is the mean (over two periods) share of capital in total primary inputs cost, and $\phi_L(1, 0)$ is the mean (over two periods) share of labour in total primary inputs cost. The previous expression can then be simplified to

$$\ln \left(\frac{C_{KL}^1}{C_{KL}^0} \right) = \phi_K(1, 0) \ln \left(\frac{C_K^1}{C_K^0} \right) + \phi_L(1, 0) \ln \left(\frac{C_L^1}{C_L^0} \right) \quad (21)$$

The next step is to assume that the capital input cost ratio and the labour cost ratio can be decomposed in price and quantity indices as

$$\frac{C_K^1}{C_K^0} = P_K(1, 0)Q_K(1, 0). \quad (22)$$

$$\frac{C_L^1}{C_L^0} = P_L(1, 0)Q_L(1, 0). \quad (23)$$

Substituting the expressions (18), (22), and (23) into expression (21), rearranging the right-hand side, and exponentiating both sides, we get

$$\begin{aligned} P_{KL}(1, 0)Q_{KL}(1, 0) \\ = (P_K(1, 0)^{\phi_K(1, 0)}P_L(1, 0)^{\phi_L(1, 0)})(Q_K(1, 0)^{\phi_K(1, 0)}Q_L(1, 0)^{\phi_L(1, 0)}). \end{aligned} \quad (24)$$

Now, if

$$P_{KL}(1, 0) = P_K(1, 0)^{\phi_K(1, 0)}P_L(1, 0)^{\phi_L(1, 0)}, \quad (25)$$

then

$$Q_{KL}(1, 0) = Q_K(1, 0)^{\phi_K(1, 0)}Q_L(1, 0)^{\phi_L(1, 0)}. \quad (26)$$

Technically seen, the last two equations mean that the joint capital-labour input price and quantity index numbers are equal to Montgomery-Vartia (MV) index numbers of the underlying, separate capital and labour price and quantity index numbers, respectively.⁷ Thus, if the condition in expression (25) is satisfied then the productivity index defined by expression (19) becomes equal to

$$ITFPROD_{VA}(1, 0) = \frac{Q_{VA}(1, 0)}{Q_K(1, 0)^{\phi_K(1, 0)}Q_L(1, 0)^{\phi_L(1, 0)}}. \quad (27)$$

Rearranging this expression in growth-accounting form, we obtain

$$Q_{VA}(1, 0) = ITFPROD_{VA}(1, 0)Q_K(1, 0)^{\phi_K(1, 0)}Q_L(1, 0)^{\phi_L(1, 0)}, \quad (28)$$

⁷ For history, definition, and properties of MV indices, see Balk (2008).

which starkly looks like an unrestricted Cobb–Douglas production function. Notice that it is straightforward to reformulate this expression in terms of levels (real value added, real inputs). By taking logarithms we obtain the more familiar form

$$\ln Q_{VA}(1, 0) = \ln ITFPROD_{VA}(1, 0) + \phi_K(1, 0) \ln Q_K(1, 0) + \phi_L(1, 0) \ln Q_L(1, 0). \quad (29)$$

Realizing that the logarithm of an index number in the neighbourhood of 1 can be interpreted as a growth rate (or, what is the same, a forward-looking percentage change), expression (29) can be read as saying that the growth rate of output is equal to the growth rate of productivity plus a weighted sum of the growth rates of capital and labour.⁸ Notice, however, that these weights are *not constant*, but depend on the data of the two periods.

If these data – Recall that data are usually given in the form of index numbers. – were generated by a ‘genuine’ CD production function then we would have

$$\ln Q_{VA}(1, 0) = \alpha + \alpha_K \ln Q_K(1, 0) + \alpha_L \ln Q_L(1, 0) + \epsilon, \quad (30)$$

where $\alpha, \alpha_K, \alpha_L$ are constants, with the last two in the interval (0, 1), and ϵ is a kind of noise. The coefficients α_K and α_L can be interpreted as marginal productivities. Expression (30) perfectly fits the following description by Douglas (1948, 6), the production unit there being American manufacturing:

“It was twenty years ago last spring that, having computed indexes for American manufacturing of the numbers of workers employed by years from 1899 to 1922, as well as indexes of the amounts of fixed capital in manufacturing deflated to dollars of approximately constant purchasing power, and then plotting these on a log scale together with the Day index of physical production for manufacturing, I observed that the product curve lay consistently between the two curves for the factors of production and tended to be approximately a quarter of the relative distance between the curve of the index for labor, which showed the least increase over the period, and that of the index for capital which showed the most.”

Or, quoting Samuelson (1979, 927-8),

“... the data Douglas presented Cobb practically commanded the C-D formula.”

It is important to notice that whereas expression (30) is the outcome of a model, expression (29) basically reflects an accounting relation. This relation imposes on the noise term of the model that

$$\begin{aligned} \epsilon = & (\ln ITFPROD_{VA}(1, 0) - \alpha) + (\phi_K(1, 0) - \alpha_K) \ln Q_K(1, 0) \\ & + (\phi_L(1, 0) - \alpha_L) \ln Q_L(1, 0). \end{aligned} \quad (31)$$

⁸ This reading corresponds to the way in which growth-accounting tables are usually constructed. Expression (29) resembles the variable-output-elasticities Cobb–Douglas (VOE-CD) function as defined by Reynès (2019).

It is also interesting to consider what precisely happens when one assumes that the data are generated by a Translog production function; that is,

$$\ln Q_{VA}(1, 0) = \alpha + \alpha_K \ln Q_K(1, 0) + \alpha_L \ln Q_L(1, 0) + \alpha_{KK}(\ln Q_K(1, 0))^{1/2} + \alpha_{LL}(\ln Q_L(1, 0))^{1/2} + \alpha_{KL} \ln Q_K(1, 0) \ln Q_L(1, 0) + \delta, \quad (32)$$

where the α 's are constants, and δ is also a kind of noise. It is straightforward to check that the accounting relation in expression (29) imposes the following relation between the Translog noise δ and the CD noise ϵ ,

$$\epsilon = \alpha_{KK}(\ln Q_K(1, 0))^{1/2} + \alpha_{LL}(\ln Q_L(1, 0))^{1/2} + \alpha_{KL} \ln Q_K(1, 0) \ln Q_L(1, 0) + \delta. \quad (33)$$

It thus appears that the second-order terms of the Translog production function consume a great deal of the CD noise.

3 Comments and extensions

A couple of things are worth noting about the key result in expressions (28) or (29). *First*, due to the concavity of the logarithmic mean, the exponents $\phi_K(1, 0)$ and $\phi_L(1, 0)$ don't add up to 1. In fact, their sum is smaller than 1,

$$\phi_K(1, 0) + \phi_L(1, 0) \leq 1, \quad (34)$$

but the discrepancy usually appears to be small (see Balk 2008, 88 for an example). Note that there is no reason whatsoever to interpret the sum of the exponents as measuring returns to scale.

Second, $\phi_K(1, 0)$ as well as $\phi_L(1, 0)$ are by definition ratios of logarithmic means. Using the Sato (1974) approximation, it appears that

$$\phi_K(1, 0) \approx \frac{\frac{1}{3} \left(\frac{C_K^1 + C_K^0}{2} + 2(C_K^1 C_K^0)^{1/2} \right)}{\frac{1}{3} \left(\frac{C_{KL}^1 + C_{KL}^0}{2} + 2(C_{KL}^1 C_{KL}^0)^{1/2} \right)}. \quad (35)$$

Thus a close approximation of $\phi_K(1, 0)$ is provided by

$$\phi_K(1, 0) \approx \frac{(C_K^1 C_K^0)^{1/2}}{(C_{KL}^1 C_{KL}^0)^{1/2}} = \left(\frac{C_K^1}{C_{KL}^1} \frac{C_K^0}{C_{KL}^0} \right)^{1/2}; \quad (36)$$

that is, the geometric mean of the capital input cost shares of the two periods. Similarly, a close approximation of $\phi_L(1, 0)$ is provided by

$$\phi_L(1, 0) \approx \frac{(C_L^1 C_L^0)^{1/2}}{(C_{KL}^1 C_{KL}^0)^{1/2}} = \left(\frac{C_L^1}{C_{KL}^1} \frac{C_L^0}{C_{KL}^0} \right)^{1/2}; \quad (37)$$

that is, the geometric mean of the labour input cost shares of the two periods. Now, recall that due to relation (2), capital input and labour cost shares add up to 1,

$$\frac{C'_K}{C'_{KL}} + \frac{C'_L}{C'_{KL}} = 1, \tag{38}$$

so that we may expect that

$$\phi_L(1, 0) \approx 1 - \phi_K(1, 0). \tag{39}$$

This would reduce expression (28) to a lookalike of the original, restricted Cobb-Douglas function, except that $\phi_L(1, 0)$ and $\phi_K(1, 0)$ are not constant. Also, there is no reason to interpret the fact that $\phi_L(1, 0)$ and $\phi_K(1, 0)$ are cost shares in terms of competitiveness, as Douglas would have done; this is just a matter of definition.⁹

Third, it is important to note that in expression (29) the productivity factor, $ITFPROD_{VA}(1, 0)$, cannot be conceived as an independent source of output growth. Basically, expression (29) is nothing but a rewritten version of the *definition* of productivity change. Put otherwise, $ITFPROD_{VA}(1, 0)$ is the result of the interplay of output change, measured by $Q_{VA}(1, 0)$, and input change, measured by $Q_K(1, 0)$ and $Q_L(1, 0)$. Contrast this with expression (30) and imagine an estimation setting in which α represents an exogenous ‘productivity shock’. The fundamental challenge in production function estimation, already recognized in the 1940s, comes from the fact that inputs and outputs are simultaneously chosen by the production unit, in connection with a ‘productivity shock’ about which its management may or may not be informed. Thus the variables on the right-hand side of expression (30) are correlated and potentially endogenous. The recent literature contains several approaches to handle this challenge; however, none of these appears to be completely satisfactory according to Kim et al. (2019).

Fourth, it must be observed that expression (28) holds independent of the magnitude of profit Π^t ($t = 0, 1$), and also independent of the magnitudes of or eventual relations between the price indices $P_{VA}(1, 0)$, $P_K(1, 0)$ and $P_L(1, 0)$. However, if profit $\Pi^t = 0$ ($t = 0, 1$) then expression (28) may be replaced by the dual expression

$$P_{VA}(1, 0) = ITFPROD_{VA}(1, 0)^{-1} P_K(1, 0)^{\phi_K(1, 0)} P_L(1, 0)^{\phi_L(1, 0)}. \tag{40}$$

This can immediately be checked by using expressions (22), (23), (21), (17), and (3) respectively.¹⁰ Thus, if profit is zero then input and output price indices are linked by total factor productivity.

Fifth, the functional form of the quantity indices $Q_{VA}(1, 0)$, $Q_K(1, 0)$ and $Q_L(1, 0)$ is left unspecified. The only restriction is that the Product Test be satisfied according to expressions (17), (22) and (23). Dependent on the time distance between periods 0 and 1 the indices might be direct or chained.

⁹ Put otherwise, though Gechert et al. (2022) concluded that “The Cobb-Douglas *production function* contradicts the data.”[my emphasis], the data might well be represented by a Cobb-Douglas *function*.

¹⁰ Expression (40) resembles expression (18) in Reynès (2019).

Sixth, in the derivation of expression (28) the only, technical not behavioural, assumption made is that the joint capital-labour input price index number behaves as an MV index number of the underlying factors; recall expression (25). This assumption, however, is rather harmless. More generality is obtained by defining a factor $\delta(1, 0)$ such that

$$P_{KL}(1, 0) = \delta(1, 0)P_K(1, 0)^{\phi_K(1,0)}P_L(1, 0)^{\phi_L(1,0)}.$$

As one easily checks, the consequence is that to the noise ε must be added $\ln \delta(1, 0)$. Now it is well known that an MV index is a second-order differential approximation of symmetric indices such as Fisher or Törnqvist, and a first-order differential approximation of asymmetric indices such as Laspeyres and Paasche (Diewert 1978). Thus if the time span between periods 0 and 1 is short enough then it does not matter much which index is employed for deflating primary inputs cost. Put differently, if the time span between periods 0 and 1 decreases then one generally expects $\delta(1, 0)$ to tend to 1.

Seventh, the extension of expression (28) to more than two (aggregate) inputs is straightforward. Let capital K consist of I types, and let labour L consist of J types. Then, in obvious notation,

$$C_K^t = \sum_{i=1}^I C_{Ki}^t \quad (41)$$

and

$$C_L^t = \sum_{j=1}^J C_{Lj}^t. \quad (42)$$

Suppose that each lower level cost ratio can be decomposed into a price index and a quantity index; that is,

$$\frac{C_{Ki}^1}{C_{Ki}^0} = P_{Ki}(1, 0)Q_{Ki}(1, 0) \quad (i = 1, \dots, I) \quad (43)$$

and

$$\frac{C_{Lj}^1}{C_{Lj}^0} = P_{Lj}(1, 0)Q_{Lj}(1, 0) \quad (j = 1, \dots, J). \quad (44)$$

Let the aggregate capital input quantity index be defined as a MV index of the lower level indices; that is,

$$Q_K(1, 0) \equiv \prod_{i=1}^I Q_{Ki}(1, 0)^{\phi_{Ki}(1,0)} \quad (45)$$

where $\phi_{Ki}(1, 0) \equiv LM(C_{Ki}^1, C_{Ki}^0)/LM(C_K^1, C_K^0)$ ($i = 1, \dots, I$). Similarly, let

$$Q_L(1, 0) \equiv \prod_{j=1}^J Q_{L_j}(1, 0)^{\phi_{L_j}(1,0)} \tag{46}$$

where $\phi_{L_j}(1, 0) \equiv LM(C_{L_j}^1, C_{L_j}^0)/LM(C_L^1, C_L^0)$ ($j = 1, \dots, J$). Substituting these two expressions into expression (28) delivers

$$Q_{VA}(1, 0) = ITFPROD_{VA}(1, 0) \times \left(\prod_{i=1}^I Q_{K_i}(1, 0)^{\phi_{K_i}(1,0)} \right)^{\phi_K(1,0)} \left(\prod_{j=1}^J Q_{L_j}(1, 0)^{\phi_{L_j}(1,0)} \right)^{\phi_L(1,0)}. \tag{47}$$

Notice that $\phi_{K_i}(1, 0)\phi_K(1, 0) = LM(C_{K_i}^1, C_{K_i}^0)/LM(C_{KL}^1, C_{KL}^0)$ ($i = 1, \dots, I$) and $\phi_{L_j}(1, 0)\phi_L(1, 0) = LM(C_{L_j}^1, C_{L_j}^0)/LM(C_{KL}^1, C_{KL}^0)$ ($j = 1, \dots, J$), so that we would have obtained expression (47) also when we immediately had acted on the situation that there are $I + J$ inputs instead of 2. This is due to the fact that the MV indices are consistent-in-aggregation (see Balk 2008, 111).

4 Gross output

Our key expression (28) relates real value added to multiple inputs. Can the expression be extended to account for gross output? For this we have to go back to the definition of (nominal) value added as revenue minus the cost of intermediate inputs (encompassing energy E , materials M , and services S); that is,

$$VA^t \equiv R^t - C_{EMS}^t. \tag{48}$$

Following Balk (2021, 57-61) it is now supposed that the revenue ratio can be decomposed by price and quantity indices $P_R(\cdot)$ and $Q_R(\cdot)$, so that

$$\frac{R^1}{R^0} = P_R(1, 0)Q_R(1, 0), \tag{49}$$

and that the intermediate inputs cost ratio likewise can be decomposed as

$$\frac{C_{EMS}^1}{C_{EMS}^0} = P_{EMS}(1, 0)Q_{EMS}(1, 0). \tag{50}$$

Then, by following the same steps as leading to expression (21), except that a minus sign now replaces the plus sign, it appears that a value-added based quantity index can rather naturally be defined as a generalized MV index; that is,

$$Q_{VA}(1, 0) \equiv \frac{Q_R(1, 0)^{\psi_R(1,0)}}{Q_{EMS}(1, 0)^{\psi_{EMS}(1,0)}}, \tag{51}$$

where $\psi_R(1, 0) \equiv LM(R^1, R^0)/LM(VA^1, VA^0)$, that is, mean revenue over mean value added, and $\psi_{EMS}(1, 0) \equiv LM(C_{EMS}^1, C_{EMS}^0)/LM(VA^1, VA^0)$, that is, mean intermediate inputs cost over mean value added. Notice that $\psi_R(1, 0)$ is an instance of the so-called *Domar factor*.

Substituting expression (51) into expression (28) delivers

$$Q_R(1, 0)^{\psi_R(1, 0)} = ITFPROD_{VA}(1, 0) Q_K(1, 0)^{\phi_K(1, 0)} Q_L(1, 0)^{\phi_L(1, 0)} \times Q_{EMS}(1, 0)^{\psi_{EMS}(1, 0)}, \quad (52)$$

or

$$Q_R(1, 0) = ITFPROD_{VA}(1, 0)^{1/\psi_R(1, 0)} Q_K(1, 0)^{\phi_K(1, 0)/\psi_R(1, 0)} \times Q_L(1, 0)^{\phi_L(1, 0)/\psi_R(1, 0)} Q_{EMS}(1, 0)^{\psi_{EMS}(1, 0)/\psi_R(1, 0)}. \quad (53)$$

This is the extension asked for.

However, we could have followed a direct route by considering that combining expressions (3), (2) and (48) lead to the KLEMS-Y accounting identity

$$C_{KLEMS}^t + \Pi^t = C_K^t + C_L^t + C_{EMS}^t + \Pi^t = R^t. \quad (54)$$

The gross-output based (total factor) productivity index is then defined by

$$ITFPROD_Y(1, 0) \equiv \frac{Q_R(1, 0)}{Q_{KLEMS}(1, 0)}, \quad (55)$$

where $Q_{KLEMS}(1, 0)$ is the quantity index component of the total input cost ratio C_{KLEMS}^1/C_{KLEMS}^0 (see Balk 2021, 19). As before, by following the same steps as leading to expression (21) it appears that a total input quantity index can rather naturally be defined as a MV index; that is,

$$Q_{KLEMS}(1, 0) \equiv Q_K(1, 0)^{\chi_K(1, 0)} Q_L(1, 0)^{\chi_L(1, 0)} Q_{EMS}(1, 0)^{\chi_{EMS}(1, 0)}, \quad (56)$$

where the exponents are, respectively, defined by

$$\begin{aligned} \chi_K(1, 0) &\equiv LM(C_K^1, C_K^0)/LM(C_{KLEMS}^1, C_{KLEMS}^0) \\ \chi_L(1, 0) &\equiv LM(C_L^1, C_L^0)/LM(C_{KLEMS}^1, C_{KLEMS}^0) \\ \chi_{EMS}(1, 0) &\equiv LM(C_{EMS}^1, C_{EMS}^0)/LM(C_{KLEMS}^1, C_{KLEMS}^0). \end{aligned}$$

By substituting expression (56) into expression (55), and rearranging, we obtain

$$Q_R(1, 0) = ITFPROD_Y(1, 0) Q_K(1, 0)^{\chi_K(1, 0)} Q_L(1, 0)^{\chi_L(1, 0)} Q_{EMS}(1, 0)^{\chi_{EMS}(1, 0)}. \quad (57)$$

Notice the differences between expressions (53) and (57). If, however, profit $\Pi^t = 0$ ($t = 0, 1$) then the corresponding exponents become the same; that is,

$$\phi_K(1, 0)/\psi_R(1, 0) = \chi_K(1, 0) \quad (58)$$

$$\phi_L(1, 0)/\psi_R(1, 0) = \chi_L(1, 0) \tag{59}$$

$$\psi_{EMS}(1, 0)/\psi_R(1, 0) = \chi_{EMS}(1, 0), \tag{60}$$

and it turns out that

$$ITFPROD_Y(1, 0) = ITFPROD_{VA}(1, 0)^{1/\psi_R(1,0)}. \tag{61}$$

Recall that $\psi_R(1, 0)$ is an instance of the Domar factor (= revenue divided by value added). Expression (61) reflects Theorem 1 of Balk (2009) (see Balk 2021, 61-2).

5 An ensemble of production units

We now consider an ensemble \mathcal{K} of production units. Think of an economy consisting of industries, or an industry consisting of enterprises. For each unit and period the KL-VA accounting identity in nominal values reads

$$C_{KL}^{kt} + \Pi^{kt} = VA^{kt} \quad (k \in \mathcal{K}), \tag{62}$$

which differs from expression (3) in that a superscript k has been added. Provided that there are no tax wedges between input and output prices and that capital and labour inputs are specific for each production unit, the individual accounting identities may simply be added up to the accounting identity for the aggregate,

$$C_{KL}^{\mathcal{K}t} + \Pi^{\mathcal{K}t} = VA^{\mathcal{K}t}, \tag{63}$$

where $C_{KL}^{\mathcal{K}t} = \sum_{k \in \mathcal{K}} C_{KL}^{kt}$, $\Pi^{\mathcal{K}t} = \sum_{k \in \mathcal{K}} \Pi^{kt}$, and $VA^{\mathcal{K}t} = \sum_{k \in \mathcal{K}} VA^{kt}$.

The aggregate value-added ratio, for period 1 relative to period 0, can be decomposed as follows (see Balk 2008, 85-86):

$$\frac{VA^{\mathcal{K}1}}{VA^{\mathcal{K}0}} = \prod_{k \in \mathcal{K}} \left(\frac{VA^{k1}}{VA^{k0}} \right)^{\Psi^k(1,0)}, \tag{64}$$

where

$$\Psi^k(1, 0) \equiv \frac{LM\left(\frac{VA^{k1}}{VA^{\mathcal{K}1}}, \frac{VA^{k0}}{VA^{\mathcal{K}0}}\right)}{\sum_{k \in \mathcal{K}} LM\left(\frac{VA^{k1}}{VA^{\mathcal{K}1}}, \frac{VA^{k0}}{VA^{\mathcal{K}0}}\right)} \quad (k \in \mathcal{K}).$$

Thus, aggregate nominal value-added change, measured as ratio, is equal to a weighted geometric mean of individual nominal value-added changes. Each coefficient $\Psi^k(1, 0)$ is the (normalized) mean share of production unit k in aggregate nominal value added. Notice that these coefficients add up to 1; that is,

$$\sum_{k \in \mathcal{K}} \Psi^k(1, 0) = 1. \tag{65}$$

For each individual unit it is assumed that the value-added ratio, the capital input cost ratio, and the labour cost ratio can be decomposed into price indices and quantity indices as in expressions (17), (22), and (23) respectively, with superscripts k added appropriately. In addition, it is assumed that the aggregate value-added ratio can be decomposed as

$$\frac{VA^{K1}}{VA^{K0}} = P_{VA}^K(1, 0)Q_{VA}^K(1, 0). \quad (66)$$

Substituting expressions (66) and (17) into expression (64), and moving $P_{VA}^K(1, 0)$ from the left-hand side to the right-hand side, delivers

$$Q_{VA}^K(1, 0) = \prod_{k \in \mathcal{K}} \left(\frac{P_{VA}^k(1, 0)}{P_{VA}^K(1, 0)} Q_{VA}^k(1, 0) \right)^{\Psi^k(1, 0)} \quad (67)$$

as aggregation rule. Substituting now expression (28), with superscripts k added in appropriate places, into expression (67) delivers our final result:

$$\begin{aligned} Q_{VA}^K(1, 0) &= \prod_{k \in \mathcal{K}} \left(\frac{P_{VA}^k(1, 0)}{P_{VA}^K(1, 0)} \right)^{\Psi^k(1, 0)} \prod_{k \in \mathcal{K}} (ITFPROD_{VA}^k(1, 0))^{\Psi^k(1, 0)} \\ &\times \prod_{k \in \mathcal{K}} \left(Q_K^k(1, 0)^{\phi_K^k(1, 0)} \right)^{\Psi^k(1, 0)} \prod_{k \in \mathcal{K}} \left(Q_L^k(1, 0)^{\phi_L^k(1, 0)} \right)^{\Psi^k(1, 0)}. \end{aligned} \quad (68)$$

A number of comments are worth making. First of all, it must be noted that, apart from the first factor on the right-hand side, expression (68) has the same structure as expression (28). This first factor measures differential price change and becomes unity if and only if

$$P_{VA}^K(1, 0) = \prod_{k \in \mathcal{K}} (P_{VA}^k(1, 0))^{\Psi^k(1, 0)}; \quad (69)$$

that is, $P_{VA}^K(1, 0)$ is a Sato-Vartia (SV) index of the underlying indices $P_{VA}^k(1, 0)$ ($k \in \mathcal{K}$).¹¹

The second factor on the right-hand side of expression (68) is a weighted mean of individual value-added based productivity indices. The third factor is an aggregate measure of capital input quantity change, whereas the fourth factor is a similarly weighted measure of labour input quantity change. These weights deserve some attention.

Notice that the capital input weights $\phi_K^k(1, 0)\Psi^k(1, 0)$ as well as the labour input weights $\phi_L^k(1, 0)\Psi^k(1, 0)$ ($k \in \mathcal{K}$) do not add up to 1. Each capital input weight approximates a production unit's share in aggregate value added times its share of capital in joint capital input and labour cost. Similarly, each labour input weight approximates a production unit's share in aggregate value added times its share of

¹¹ For history, definition, and properties of SV indices, see Balk (2008).

labour in joint capital input and labour cost. Together, their sum is smaller than unity; that is,

$$\sum_{k \in \mathcal{K}} (\phi_K^k(1, 0)\Psi^k(1, 0) + \phi_L^k(1, 0)\Psi^k(1, 0)) \leq 1. \quad (70)$$

Again, there is no reason to interpret this inequality in terms of returns to scale as it is entirely due to the concavity of the logarithmic mean.

6 Some concluding thoughts

Why is the Cobb-Douglas (production) function so popular? Biddle (2021, 312) ascribes its success story to a number of factors: 1. the link to the neo-classical paradigm, 2. the flexibility/adaptability of the function, 3. the rhetorical talents of its protagonist, 4. the research and teaching activities of a (small) number of charismatic personalities. Add to this "the prima facie plausibility of the results of Douglas's research with the regression", and those of many of his followers in a variety of applications.

Why this plausibility? The simple answer appears to be: because the CD structure is an alternative representation of the accounting identity coming with the data. At any level of aggregation, real input and output data are not observables but constructs. Given are *nominal* data, tied together by an accounting identity that in most general terms states that revenue equals cost plus profit. In this article it is rigorously demonstrated that the *real* version of this identity exhibits the CD pattern, whatever the underlying data-generating mechanism may be. This explains the good 'fit' one usually encounters in empirical work.¹²

With hindsight, Phelps Brown (1957) must be given the credit for having noticed this for the first time, though his explanation was a bit heuristic and needed some assumptions. In deriving the key results in the previous sections of this paper no structural or behavioural assumptions were invoked. It was all a matter of accounting and playing with identities, realizing that addition and multiplication are isomorphic operations.

The history as documented by Biddle (2021) makes clear that through time the status of the CD function has changed from being a 'law', ideally holding for the input and output of aggregates such as the manufacturing industry, to a handy

¹² This explanation differs from the argument put forward by Ishikawa (2021). This author started by observing the empirical fact that the distributions of capital, labour, and output over firms follow power-laws, and concluded that "If firms' assets, labor, and production are expressed as a set of points in three-dimensional space, the Cobb-Douglas production function can be interpreted as a quasi-inverse-symmetric plane in three-dimensional space and a residual from the plane." (p. 129) Evidently for a conclusion like this a very large data-set of individual firm data covering several countries is necessary, and the conclusion as such holds only for very large aggregates. On the contrary, the explanation offered in this paper is independent of the aggregation level.

measurement tool that can be used in very different research situations.¹³ A parallel development was the transformation of the neo-classical paradigm from a set of testable propositions to a set of articles of faith.

Modern production theory¹⁴ starts with the concept of a technology *set* rather than a production *function*. A technology set contains all the input–output combinations that are feasible in a certain period of time, and to which all the production units under consideration have access. Under mild conditions such a set can be represented by a radial output distance function, of which the production function is a specific materialization. Basically, the production function specifies the boundary (*aka* frontier) of the technology set.

To be more specific, let us return to the setting of Section 5. We know that the nominal accounting identity in expression (62) implies that

$$(X_K^k(t, b), X_L^k(t, b), RVA^k(t, b)) \in S^t \quad (k \in \mathcal{K}),$$

where real values were defined in expressions (4), (5) and (6), and S^t denotes the technology set of period t . The production function¹⁵ corresponding with S^t is given by

$$F^t(X_K, X_L) \equiv \max\{RVA \mid (X_K, X_L, RVA) \in S^t\}, \quad (71)$$

where X_K , X_L , and RVA denote generic variables. It then immediately follows that

$$RVA^k(t, b) \leq F^t(X_K^k(t, b), X_L^k(t, b)) \quad (k \in \mathcal{K}). \quad (72)$$

Since $RVA^k(t, b) = VA^{kb} Q_{VA}^k(t, b)$, by substituting expression (28) we obtain

$$\begin{aligned} VA^{kb} ITFPROD_{VA}^k(t, b) Q_K^k(t, b)^{\phi_K^k(t, b)} Q_L^k(t, b)^{\phi_L^k(t, b)} \\ \leq F^t(X_K^k(t, b), X_L^k(t, b)) \quad (k \in \mathcal{K}), \end{aligned} \quad (73)$$

or

$$ITFPROD_{VA}^k(t, b) \leq \frac{F^t(X_K^k(t, b), X_L^k(t, b))}{VA^{kb} Q_K^k(t, b)^{\phi_K^k(t, b)} Q_L^k(t, b)^{\phi_L^k(t, b)}} \quad (k \in \mathcal{K}). \quad (74)$$

Using the definitions of the real values in the numerator the last expression can be rewritten as

$$ITFPROD_{VA}^k(t, b) \leq \frac{F^t(C_K^{kb} Q_K^k(t, b), C_L^{kb} Q_L^k(t, b))}{VA^{kb} Q_K^k(t, b)^{\phi_K^k(t, b)} Q_L^k(t, b)^{\phi_L^k(t, b)}} \quad (k \in \mathcal{K}). \quad (75)$$

¹³ For example, a bilateral price index that satisfies the basic axioms plus the Circularity (Transitivity) test necessarily has the generalized CD form; that is, a weighted geometric mean of price relatives with weights adding up to one. See Balk (2008, 97–99) for details.

¹⁴ See for instance Färe (1988).

¹⁵ Hildenbrand (1981) would call this the short-run efficient industry production function.

It is clear that there are several degrees of freedom here; most notably, the choice of production-unit-specific quantity indices (or, equivalently, deflators) for capital and labour. However, as these indices occur simultaneously in numerator and denominator, the upper bound on productivity growth should not be too dependent on particular choices. The specification of the function $F^t(\cdot)$, however, is more important. Thus, to avoid unwarranted impositions this calls for a specification of the technology set S^t as a close envelopment of the data, in line with the conclusion of Hildenbrand (1981).¹⁶

Funding There was no funding.

Declarations

Conflict of interest There are no potential conflicts.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Balk BM (2008) Price and Quantity Index Numbers: Models for Measuring Aggregate Change and Difference. Cambridge University Press, Cambridge
- Balk BM (2009) On the Relation Between Gross-output and Value-added Based Productivity Measures: The Importance of the Domar Factor. *Macroeconomic Dynamics* 13, Supplement No. 2 (Special Issue on Measurement with Theory), 241–267
- Balk BM (2021) Productivity: Concepts, Measurement, Aggregation, and Decomposition. *Contributions to Economics*. Springer Nature, Switzerland AG
- Biddle JE (2021) Progress through Regression: The Life Story of the Empirical Cobb-Douglas Production Function. Cambridge University Press, Cambridge
- Chambers RG, Ray SC (2021) Neoclassical Production Economics: An Introduction. In: Ray SC, Chambers RG, Kumbhakar SC (eds) *Handbook of Production Economics*. Springer, Singapore
- Cobb CW, Douglas PH (1928) A theory of production. *Amer Econ Rev* 18(Supplement):139–165
- De Loecker J, Syverson C (2021) An Industrial Organization Perspective on Productivity. In: *Handbook of Industrial Organization*, Volume 4, edited by K. Ho, A. Hortaçsu and A. Lizzeri (Elsevier)
- Diewert WE (1978) Superlative index numbers and consistency in aggregation. *Econometrica* 46:883–900
- Douglas PH (1948) Are there laws of production? *Amer Econ Rev* 38:1–41
- Douglas PH (1976) The Cobb-Douglas production function once again: its history, its testing, and some new empirical values. *J Polit Econ* 84:903–916
- Färe R (1988) *Fundamentals of Production Theory*, Lecture Notes in Economics and Mathematical Systems, vol 311. Springer-Verlag, Berlin

¹⁶ The conclusion of his analysis was “it suggests a parameter-free approach based on technological information on the microlevel.”

- Felipe J, Adams FG (2005) A theory of production, the estimation of the Cobb-Douglas function: a retrospective view. *Eastern Econ J* 31:427–445
- Felipe J, Fisher FM (2008) Aggregation (production). In: *The new Palgrave dictionary of economics*. Palgrave Macmillan, London. https://doi.org/10.1057/978-1-349-95121-5_2552-1
- Felipe J, McCombie JSL (2001) The CES production function, the accounting identity, and Occam's Razor. *Appl Econ* 33:1221–1232
- Felipe J, McCombie JSL (2005) How sound are the foundations of the aggregate production function? *Eastern Econ J* 31:467–488
- Felipe J, McCombie JSL (2007) Is a theory of total factor productivity really needed? *Metroeconomica* 58:195–229
- Felipe J, McCombie JSL (2013) *The Aggregate Production Function and the Measurement of Technical Change: 'Not Even Wrong'*. Edward Elgar, Cheltenham UK
- Fisher FM (1971) Aggregate production functions and the explanation of wages: a simulation experiment. *Rev Econ Stat* 53:305–325
- Gechert S, Havranek T, Irsova Z, Kolcunova D (2022) Measuring capital-labor substitution: the importance of method choices and publication bias. *Rev Econ Dynam* 45:55–82
- Hildenbrand W (1981) Short-run production functions based on microdata. *Econometrica* 49:1095–1125
- Ishikawa A (2021) Why Does Production Function Take the Cobb-Douglas Form? In: *Statistical Properties in Firms' Large-scale Data, Evolutionary Economics and Social Complexity Science*, vol. 26. Springer, Singapore
- Jorgenson DW, Griliches Z (1967) The explanation of productivity change. *Rev Econ Stud* 34:249–283
- Kim K, Luo Y, Su Y (2019) A robust approach to estimating production functions: replication of the ACF procedure. *J Appl Econ* 34:612–619
- Knoblauch M, Roessler M, Zwerschke P (2020) The elasticity of substitution between capital and labour in the US economy: a meta-regression analysis. *Oxford Bull Econ Stat* 82:62–82
- Phelps Brown EH (1957) The meaning of the fitted Cobb-Douglas function. *Quar J Econ* 71:546–560
- Reynès F (2019) The Cobb-Douglas function as a flexible function; a new perspective on homogeneous functions through the lens of output elasticities. *Math Social Sci* 97:11–17
- Samuelson PA (1979) Paul Douglas's measurement of production functions and marginal productivities. *J Polit Econ* 87:923–939
- Sandelin B (1976) On the origin of the Cobb-Douglas production function. *Econ History XIX*:117–123
- Sato K (1974) Ideal index numbers that almost satisfy the factor reversal test. *Rev Econ Stat* 56:549–554
- Shaikh A (1974) Laws of production and laws of algebra: the Humbug production function. *Rev Econ Stat* 56:115–120
- Shaikh A (2005) Dynamics and pseudo-production functions. *Eastern Econ J* 31:447–466
- Simon HA (1979) On parsimonious explanations of production relations. *Scandinavian J Econ* 81:459–474
- Simon HA, Levy FK (1963) A note on the Cobb-Douglas function. *Rev Econ Stud* 30:93–94
- Solow RM (1957) Technical Change and the Aggregate Production Function, *Review of Economics and Statistics* 39, 312–320. Reprinted in *National Accounting and Economic Growth*, edited by John M. Hartwick, *The International Library of Critical Writings in Economics* No. 313 (Edward Elgar, Cheltenham UK, Northampton MA, 2016)
- Tinbergen J (1942) Professor Douglas' production function. *Rev Int Stat Inst* 10:37–48
- Tinbergen J (1942) Zur Theorie der Langfristigen Wirtschaftsentwicklung. *Weltwirtschaftliches Archiv* 55:511–549

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.