Cambridge Journal of Economics 2008, **32**, 441–459
doi:10.1093/cje/bem043
Advance Access publication 13 December, 2007

Correcting for biases when estimating production functions: an illusion of the laws of algebra?

Jesus Felipe, Rana Hasan and J. S. L. McCombie*

This paper shows that the endogeneity bias that allegedly appears when estimating production functions using value data, and which the literature has tried to deal with since the 1940s, is simply the result of omitted-variable bias due to a poor approximation to an accounting identity. This problem has no econometric solution. As a result, recent attempts to solve the problem by developing new estimators are questioned. The only possible way to estimate the technological parameters of the production function is to use physical quantities.

Key words: Accounting identity, Endogeneity, Omitted-variable bias, Production function
JEL classifications: C13, D24, O47

1. Introduction

Estimation of production functions is central to a large body of empirical work in economics. Work in areas such as the sources of growth, returns to scale or technological progress, among others, proceeds under the assumption that a production function exists and then either assumes the values of its parameters (in the case of growth accounting) or estimates them. The data used in these estimations are, with few exceptions, monetary values—the *value* of output, or revenues, and the *values* of the various inputs, appropriately deflated.[1] A pre-eminent concern of this literature has been to address the endogeneity biases that appear as a result of the supposed contemporaneous correlation between the error term and the factor

Manuscript received 21 May 2007; final version received 5 September 2007.
Address for correspondence: Jesus Felipe, Asian Developement Bank, Manila, Philippines, Email: Jfelipe@adb.org

*Jesus Felipe: Asian Development Bank, Manila, Philippines, Cambridge Centre for Economic and Public Policy, University of Cambridge, UK, Centre for Applied Macroeconomic Analysis, College of Business and Economics, Australian National University Australia; Rana Hasan: Asian Development Bank (Manila, Philippines); and J. S. L. McCombie: Cambridge Centre for Economic and Public Policy, University of Cambridge, University of Cambridge and Downing College, Cambridge, UK. The authors are grateful to Steven Durlauf and to two anonymous referees for their comments and suggestions. Any remaining errors are solely the authors'. This paper represents the views of the authors and does not represent those of the Asian Development Bank, its Executive Directors or the countries that they represent.

[1] The only true physical quantity used in the estimation of production functions is labour, introduced in terms of either number of workers or number of hours worked. This suffers from an aggregation problem (see Felipe and Fisher, 2003).

inputs due to simultaneity.[1] As is well-known, the ordinary least squares (OLS) estimates of the parameters of the production function are biased and inconsistent in the presence of such correlation. The main approach to dealing with this problem has been to use an instrumental variable (IV) estimator. The IV estimator achieves consistency by instrumenting the factor inputs with regressors that are correlated with them but uncorrelated with the error term. The IV estimator approach has been extended recently in a number of ways with new methods proposed for conditioning out serially correlated shocks to the production technology. See, for example, Olley and Pakes (1996), Blundell and Bond (2000) and Levinsohn and Petrin (2003).[2]

The purpose of this paper is to demonstrate that the alleged endogeneity problem that these papers propose to tackle is a red herring. The endogeneity problem is not the cause of the implausible results that often appear in empirical estimations of production functions. Instead, the very nature of the data that are used in estimating production functions is the source of the problem. Data typically used are deflated monetary values of output and inputs (rather than *physical quantities* of output and inputs) that are, in fact, related definitionally through an accounting identity. This, as we shall show, prevents meaningful econometric estimates of the production technology being obtained and reinforces the results on aggregation of Blackorby and Schworm (1984, p. 647), who indicated that in empirical work using aggregate data 'it is almost always inappropriate to treat regression coefficients as estimates of parameters characterising an aggregate technology'. Put differently, the concern of much of the literature with endogeneity issues in the estimation of production functions is misplaced, and the problem of estimating production functions is substantially more serious than one requiring the correction for possible simultaneity (or any other putative) biases. It is one that has no econometric solution.

This is an implication of an argument that appeared in an early paper by Phelps Brown (1957), who outlined a criticism of Douglas's (1948) cross-industry results. This critique was later formalised in a note by Simon and Levy (1963). Simon (1979A) extended it into a full paper and elaborated upon its implications. He considered it to be of sufficient importance to warrant mention in his Nobel Lecture (Simon, 1979B). Paradoxically, the same year, Samuelson (1979) also rediscovered the same argument. Shaikh (1974) gave the *coup de grâce* with his memorable 'Humbug' production function, and Shaikh (1980) offers a detailed and incisive discussion of the critique. Shaikh (2005) further shows, by means of a simulation analysis, how it is not possible to distinguish between a fixed-coefficients technology in a Goodwin growth model and a conventional Cobb–Douglas production function because of the accounting identity. The literature is surveyed by Felipe and McCombie (2005).

As this important argument seems to have been neglected in the literature, notwithstanding its profound and damaging implications for the estimation and interpretation of production functions, it is useful to summarise it as well as discuss its implications for applied econometrics and the concerns that it raises for practitioners.[3]

[1] The problem was studied in detail by Marschak and Andrews (1944), who were aware of the endogeneity problem in the estimation of production functions and clearly argued that the endogeneity of inputs implies that there is a system of simultaneous equations.

[2] Olley and Pakes (1996) use investment as a control function to mimic the fixed effect. Levinsohn and Petrin (2003) use a mix of instrumental variables and control functions. Finally, Blundell and Bond (2000) propose a model where the error term is divided into three parts. They use GMM estimation where the moments used for identification are the lagged values of the inputs.

[3] We consider that this neglect has occurred because the implications have not been totally understood and/or clearly spelled out, even by those who are aware of the argument. Some neoclassical economists who use the aggregate production function are also naturally reluctant to concede that much of their academic work has been a pointless exercise. See fn. (2) on p. 10 for a discussion of how the argument has been treated in some econometrics textbooks.

In essence, this critique states that the income accounting identity that relates value added (gross output) to the sum of the wage bill plus total profits (plus intermediate materials and energy) can be algebraically rewritten as a form that resembles a production function. The implications are far reaching: if the functional form chosen (Cobb–Douglas, CES, etc.) for fitting a production function provides a good approximation to the identity, the statistical fit obtained will generally be very high, potentially with a correlation coefficient of unity; the estimated putative factor elasticities will be equal to the factor shares; and because what one is essentially estimating is an accounting identity (or a very good approximation to it), the whole exercise is extremely problematic, if not useless. Theoretical problems such as possible biases due to endogeneity of the inputs, or the effects of the presence of unit roots in the variables (in the case of time series) are, at most, of secondary importance.

It is important to note that using micro-level data alone does not get round the problem. Researchers in many studies use firm-level or even plant-level data pooled at the 3- or 4-digit-level industry data. However, as Levinsohn and Petrin (2003, p. 323, fn. 17; see also Griliches and Mairesse, 1998, p. 195) acknowledge, the measure of output used in these studies is not a physical quantity (the data that should be used to estimate production functions), but deflated gross output or value added.[1] Moreover, there is considerable differentiation of products at even the 4-digit industry level, and even a firm or plant can be producing disparate products (which most likely use different production processes). Thus, the only way to report an 'aggregate' measure of output is in monetary or value terms (however deflated). At higher levels of aggregation (e.g., sectors of the economy, total economy), the problem is even more severe, and the well-known aggregation problems become an insurmountable difficulty (Felipe and Fisher 2003, 2006A, 2006B; Felipe and McCombie, 2003).[2]

The rest of the paper is structured as follows. Section 2 summarises and discusses the main implications of using accounting data to estimate production functions. Section 3 provides a rationalisation of the findings of researchers who estimate production functions using value data and employ a variety of econometric techniques to overcome 'implausible' results. Section 4 uses firm-level data to illustrate the arguments more concretely. We show how estimates of production functions can be interpreted as estimates of a 'mis-specified accounting identity.' Section 5 concludes.

2. The production function and the tyranny of the accounting identity

For expositional clarity, the argument can be summarised most easily for the case of value added and time-series data. The arguments for the cases of gross output and/or cross-section data can be similarly derived (see Section 3). The income accounting identity may be written as

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

[1] The problems regarding capital are equally damaging. Owing to space constraints, we do not address them here. Suffice to say that the standard measures of the stock of capital used in the estimation of production functions (calculated through the perpetual inventory method at aggregate levels or as the book value of plant and equipment if firm-level data are used) most likely do not have much to do with the true notion of physical capital that, theoretically, should go into the production function (Cohen and Harcourt 2003).

[2] The aggregation problems affect not only capital, but also output and labour.

where V_t is output (value added) in constant prices. Π_t and W_t are the total wage bill and total profits, respectively; w_t is the real wage rate; L_t is the labour input (number of hours worked or number of workers); and r_t is the profit rate, defined as $r_t \equiv \pi_t / \mathcal{J}_t$ (measured as dollars of profit per dollar of capital, i.e., a pure number), where $\mathcal{J}_t$ is the *value* of the stock of capital in constant prices.

In proportionate growth rates (and in real terms) equation (1) becomes

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

where ^ denotes a growth rate, $a_t \equiv r_t\mathcal{J}_t / V_t$ is the share of capital in output, $(l - a_t) \equiv w_t L_t / V_t$ is the share of labour, and $g_t \equiv (1 - a_t)\hat{w}_t + a_t\hat{r}_t$.

Now let us assume that factor shares are constant, i.e., $a_t = a$; $(1 - a_t) = (1 - a)$. Then equation (2) becomes

$$\hat{V}_t \equiv (1-a)\hat{w}_t + a\hat{r}_t + (1-a)\hat{L}_t + a\hat{\mathcal{J}}_t \equiv g_t^\star + (1-a)\hat{L}_t + a\hat{\mathcal{J}}_t \tag{3}$$

where $g_t^\star \equiv (1 - a)\,\hat{w}_t + a\hat{r}_t$. Integrating equation (3) yields

$$\ln V_t \equiv \ln B_0 + (1-a)\ln w_t + a\ln r_t + (1-a)\ln L_t + a\ln \mathcal{J}_t \tag{4}$$

and taking anti-logarithms, we obtain

$$V_t \equiv B_0 w_t^{1-a} r_t^a L_t^{1-a} \mathcal{J}_t^a \tag{5}$$

Moreover, suppose also that wage and profit rates grow at constant rates, i.e., $\hat{w}_t = \hat{w}$ and $\hat{r}_t = \hat{r}$. This, together with the assumption that factor shares are constant, implies that $g^\star \equiv a\hat{r} + (1 - a)\hat{w}$. Equation (5) can now be written as

$$V_t \equiv B_0 e^{g^\star t} L_t^{1-a} \mathcal{J}_t^a \tag{6}$$

In growth rates, this is

$$\hat{V}_t \equiv g^\star + (1-a)\hat{L}_t + a\hat{\mathcal{J}}_t \tag{7}$$

The important point of this derivation is that equation (6) (or (7) in growth rates) is *not* a Cobb–Douglas production function. It is simply the income accounting identity, equation (1), rewritten under the two assumptions of constant factor shares and a constant weighted growth rate of the wage and profit rates.

The implications of this derivation are very important and, ultimately, prevent any unambiguous interpretation of the estimated parameters of an alleged production function (estimated using value data) as the true technological coefficients. They are discussed in the next section.

3. Rationalising the estimates of production functions with value data

Suppose one gathers data for output V_t, employment L_t and capital $\mathcal{J}_t$, and estimates by OLS the regression $V_t = A_0 e^{\lambda t} L_t^{\alpha} \mathcal{J}_t^{\beta} exp(\varepsilon_t)$ (i.e., in an unrestricted form), where ε_t is the random disturbance assumed to have the standard properties and where V_t, L_t and $\mathcal{J}_t$ are the same series as in equation (1).[1] If the two assumptions above about the factor shares

[1] The reader may argue that for this to be the case, the data must satisfy the accounting identity (1) above. The data on output, labour and capital, however, *must* satisfy the identity, in the sense that the series are the same. It is obvious that the identity has to hold.

and the paths of the wage and profit rates happen to be empirically correct, it is obvious then, by comparison with equation (6), the accounting identity, that the statistical fit will be perfect, and the estimates will be $\lambda \equiv a\hat{r} + (1 - a)\hat{w} \equiv g^{\star}$, $\alpha \equiv (1 - a)$ and $\beta \equiv a$. Under a neoclassical interpretation, the equality of the elasticities and the factor shares would be interpreted as a failure to refute the neoclassical theory of factor pricing and, consequently, the assumption that markets are competitive. It can also be seen that the estimate of the trend λ is a weighted average of the (constant) growth rates of the wage and profit rates.[1] Moreover, the result indicates the putative presence of 'constant returns to scale'. However, this data set could well correspond to, for example, a command economy where factors are not paid their marginal products. All we have used in deriving equation (6) is the identity that output equals the payment to the factors of production, together with the two empirical assumptions.[2] The fact that the estimated 'output elasticities' closely approximate the factor shares does not imply that markets are competitive, and that there are constant returns to scale. This correspondence merely follows from the accounting identity.

What happens if the results obtained are implausible, e.g., the estimated coefficients are significantly different from the observed factor shares? Implausible results have led most researchers to interpret them as being due to endogeneity bias. However, the implausible results simply mean that one or both assumptions used to derive equation (6) from equation (2) do not correspond empirically to the data:

(i) Factor shares vary substantially. Assume the researcher is not aware of the aggregation problems and mistakenly believes that there is an underlying production function, which in growth rates can be written as

$$\hat{V}_t = \lambda_t + \alpha_t \hat{L}_t + \beta_t \hat{J}_t + \mu_t \tag{8}$$

where α_t and β_t are the alleged output elasticities of labour and capital (which, in general, may vary in time), λ_t is the growth rate of total factor productivity, and μ_t is the random term associated with the growth rate specification. However, the accounting identity is given by equation (2), i.e., ${}_t\hat{V}_t \equiv g_t + (1 - a_t)\hat{L}_t + a_t \hat{J}_t$. A comparison of both equations indicates that the results obtained using any estimation procedure that gives a good approximation to the identity (such as a time-varying coefficients estimation method or simply recursive or rolling estimations) could be mistakenly interpreted as the result of having estimated a production function. It

[1] We realise that the equation $g_t \equiv (1 - a_t)\hat{w}_t + a_t\hat{r}_t$ is what neoclassical analyses refer to as the dual measure of total factor productivity growth. Our argument does not deny this, in the sense that the equation can be referred to as such. However, it must not be forgotten that our derivation does not assume the existence of any production function, and/or that factor markets are competitive, necessary assumptions in the standard derivation. This equation is simply part of the accounting identity. This is true always by definition. Also recall that the derivation uses the profit rate, and not the user cost of capital, often used in the neoclassical measure of the dual (this does not affect the argument). This is the only difference. The user cost of capital, unlike the concept of profit rate, is theory-dependent (it follows from neoclassical theory) and has to be calculated making a series of assumptions. This means that there is no way of knowing and testing whether the number computed is correct (see Felipe and McCombie, 2007).

[2] The assumption about the constancy of the factor shares could be mistaken to imply that we have implicitly assumed a Cobb–Douglas production function. Constant factor shares are indeed consistent with a Cobb–Douglas production function. However, Fisher (1971) showed using simulation experiments that this conclusion need not necessarily follow. In fact, Fisher showed that the Cobb–Douglas form tends to work well in empirical analyses *because* factor shares are constant; and *not* the other way around, that is, that factor shares are constant because the underlying production function is Cobb–Douglas.

turns out that, when factor shares vary substantially, one can derive approximations to the accounting identity different from the one in the previous section (that led to a form resembling the Cobb–Douglas in equation (6)). These other approximations lead to mathematical forms that resemble other production functions and which will yield better results in terms of statistical fit and elasticities that equal the factor shares.[1]

(ii) Wage and profit rates do not grow at constant rates. It turns out that, empirically, the assumption that they do is likely to be incorrect, and is the one that most often produces 'implausible' estimates of the coefficients in the Cobb–Douglas specification with an exponential time trend. That is, although factor shares do vary in reality, the variation displayed by actual time-series data does not cause the implausible results. As the term $g_t \equiv (1 - a_t)\hat{w}_t + a_t\hat{r}_t$ (see equation (2)) cannot be approximated correctly by a constant, it will have to be approximated by a more flexible functional form, for example, a trigonometric function. Alternatively, it may be better proxied by another variable that fluctuates significantly and that is highly correlated with g_t.

In some more detail, what happens is as follows. Let us assume for simplicity that indeed factor shares are constant so that the accounting identity in growth rates corresponds to equation (3), that is $\hat{V}_t \equiv g_t^\star + (1 - a)\hat{L}_t + a\hat{\mathcal{J}}_t$, where $g_t^\star$, it will be recalled, equals $(1 - a)\hat{w}_t + a\hat{r}_t$. If the said researcher who mistakenly believes that there exists a Cobb–Douglas function estimates it in growth rates and in an unrestricted form as

$$\hat{V}_t = \lambda + \alpha\hat{L}_t + \beta\hat{\mathcal{J}}_t + \phi_t \qquad (9)$$

where λ is a constant rate of growth, the econometric problem is akin to one of omitted-variable bias. In other words, a relevant variable has been omitted from the regression (ϕ_t is the error term associated with this specification). To see this, compare equations (3) and (9). Why does so often the latter fail to yield plausible estimates of the coefficients, a problem researchers interpret usually as the result of endogeneity bias (Griliches and Mairesse, 1998, Levinsohn and Petrin, 2003)? A comparison of both equations indicates that what the constant term λ 'tries to do' in the Cobb–Douglas regression is to approximate the term $g_t^\star \equiv (1 - a)\hat{w}_t + a\hat{r}_t$ in the accounting identity (3). However, this seldom works as empirically $g_t^\star$ displays pronounced cyclical fluctuations. Hence, the estimates of α and β appear to be biased. This, however, is not the result of endogeneity of the regressors (in the standard sense) but of a poor approximation of the term $g_t^\star$ using the constant λ. What this means is that the accounting identity (3) could be very well approximated (in the limit the approximation would be perfect) by $\hat{V}_t \equiv \hat{A}(t) + (1 - a)\hat{L}_t + a\hat{\mathcal{J}}_t$, where $\hat{A}(t)$ is, for example, a complex (trigonometric) function of time that accurately tracks the path of $g_t^\star$ (not necessarily a smooth function of time). Integration would lead to $V_t = C_0 e^{A(t)} L_t^\alpha \mathcal{J}_t^\beta$. Of course, this new equation approximating the accounting identity should not be confused either with the Cobb–Douglas production function (as is also the case with equations (6) and (7)).

[1] For example, assume that $a_t\hat{w}_t + (1 - a_t)\hat{r}_t \equiv \lambda$ and that factor shares evolve as $(1 - a_t) \equiv \delta L_t^{-\rho}/V_t$ and $a_t \equiv (1 - \delta)\mathcal{J}_t^{-\rho}/V_t$. Substitute these three paths into the identity in growth rates, equation (2), and integrate it. The result is $V_t \equiv A_0 e^{\lambda t}[\delta L_t^{-\rho} + (1 - \delta)\mathcal{J}_t^{-\rho}]$, which resembles a CES production function (see Felipe and McCombie, 2001).

The 'degree of bias' can be easily derived algebraically because what is being omitted in this regression is the term $g_t^{\star} \equiv (1-a)\,\hat{w}_t + a\hat{r}_t$. The OLS estimator of the capital coefficient in the Cobb–Douglas regression in growth rates (equation (9)) is given by:

$$\beta_{OLS} = \frac{Cov(\hat{V}_t,\hat{\mathcal{J}}_t)\,Var(\hat{L}_t) - Cov(\hat{V}_t,\hat{L}_t)Cov(\hat{L}_t,\hat{\mathcal{J}}_t)}{Var(\hat{L}_t)\,Var(\hat{\mathcal{J}}_t) - [Cov(\hat{L}_t,\hat{\mathcal{J}}_t)]^2} \tag{10}$$

To calculate the bias, substitute (3) (what we refer to as the 'true model') into equation (10) for $\hat{V}_t$. Working out the algebra and taking expectations yields

$$E(\beta_{OLS}) = a + E\left[\frac{Cov(\hat{\mathcal{J}}_t,g_t^{\star})\,Var(\hat{L}_t) - Cov(\hat{L}_t,g^{\star}_t)Cov(\hat{L}_t,\hat{\mathcal{J}}_t)}{Var(\hat{L}_t)\,Var(\hat{\mathcal{J}}_t) - [Cov(\hat{L}_t,\hat{\mathcal{J}}_t)]^2}\right] \tag{11}$$

where, since equation (3) is an accounting identity, its *theoretical disturbance* term (which does not exist, as the equation is an accounting identity) is uncorrelated with the regressors. Hence, one cannot speak of an endogeneity problem.[1] An equation can be derived in a similar manner for the labour coefficient.

Equation (11) shows that the estimated coefficient equals the share of capital in value added a, plus a term that captures the 'bias' (the second part of equation (11)), which is due to the incorrect specification of $g_t^{\star} \equiv (1-a)\hat{w}_t + a\hat{r}_t$ as a constant λ, plus the error term ϕ_t. If the estimate of β equals a, this will imply that the bias in equation (11) is zero. However, this will simply mean either that $g_t^{\star}$ is a constant (which would imply that wage and profit rates grow at constant rates so that $g_t^{\star} \equiv (1-a)\hat{w} + a\hat{r} \equiv g^{\star}$), and hence $Cov(\hat{\mathcal{J}}_t, g_t^{\star}) = Cov(\hat{L}_t, g_t^{\star}) = 0$; or that, by coincidence, the numerator of equation (11) equals zero. Given that the denominator of the bias is always positive, the sign is determined by the numerator.[2] Why is the capital coefficient often 'biased downward' (see Levinsohn and Petrin, 2003, p. 333)? The reason is that, on the one hand, the variable that drives the fluctuations in $g_t^{\star}$ is $\hat{r}_t$ (wage rate and factor shares do not fluctuate that much), which is a highly procyclical variable, hence highly correlated with $\hat{L}_t$. On the other hand, the growth of capital $\hat{\mathcal{J}}_t$, shows virtually no cyclical fluctuations.[3]

Often researchers use a capacity-utilisation adjusted capital stock $\hat{\mathcal{J}}_t^{\star}$, instead of the unadjusted one $\hat{\mathcal{J}}_t$. This solution tends to produce plausible estimates (Lucas, 1970; Barro, 1999, p. 123 also refers to this option). Returning to the omitted-variable $g_t^{\star}$ in the Cobb–Douglas regression, the reason why the use of $\hat{\mathcal{J}}_t^{\star}$ works is that, fluctuations in the wage rate $\hat{w}_t$ are much smaller than those in the profit rate $\hat{r}_t$, a highly procyclical variable; and most often $\hat{r}$ is highly correlated with $\hat{V}_t$ and $\hat{L}_t$. Therefore, $\hat{\mathcal{J}}_t^{\star}$ will be

[1] At the expense of labouring the obvious, the orthogonality condition is $Cov(u_t\hat{\mathcal{J}}_t) = Cov(u_t\hat{L}_t) = 0$, where u_t is the 'disturbance term' of the accounting identity (3), which obviously does not exist; that is why the correlation is zero. This point should not be mistaken with the fact that, in the standard model, input prices are assumed to drive input choices, and hence wage and profit rates are correlated with labour and capital. The issue here is simply that the 'true model', namely, the accounting identity, equation (3), does not contain a disturbance term. Hence the alleged endogeneity problem is a red herring.

[2] Equation (11) is the same as Levinsohn and Petrin's (2003) unnumbered equation on p. 319 in their paper. The difference is that, here, we clearly show that the first part of the expected value of the coefficient of capital has to be, precisely, the capital share (a).

[3] Playing devil's advocate, a counter argument to ours could be that the estimates of the parameters in the production function are biased because of the omission of technical progress in the regression. This is precisely how $g_t^{\star}$ is interpreted in the neoclassical analysis; namely, the dual expression for total factor productivity growth. However, the reader must not forget that all our derivation stems from an accounting identity (there is no behavioural model here), and that if $g_t^{\star}$ is appropriately included or proxied in the regression, one would always obtain the same suspicious results: a perfect fit and elasticities equal to the factor shares.

procyclical by construction and thus the specification of the production function with $\hat{\mathcal{J}}_t^*$ will more closely approximate the underlying identity, as this variable 'proxies' the movements of the omitted variable $\hat{r}_t$.[1]

What about the conventional argument that estimation of the production function should be part of a system of equations that includes the marginal productivity conditions? This argument overlooks the point that the marginal productivity conditions are also related definitionally through accounting identities. In the case of the Cobb–Douglas production function, the marginal productivity conditions are $w_t = \alpha V_t/L_t$ and $r_t = \beta V_t/\mathcal{J}_t$, which would be estimated together with the production function by adding their corresponding disturbances. However, it will be noted that the equations for the factors shares implicit in the accounting identity, equation (1), are $a_t \equiv r_t\mathcal{J}_t/V_t$, the share of capital in output, and $(1 - a_t) \equiv w_t L_t/V_t$, the share of labour (also identities). These two equations can be rewritten as $w_t \equiv (1 - a_t)V_t/L_t$ and $r_t \equiv a_t V_t/\mathcal{J}_t$. If in this economy factor shares are sufficiently constant, the marginal productivity conditions would give plausible results when estimated econometrically, in the sense that results would yield $\alpha = (1 - a)$ and $\beta = a$. This result, once again, is driven by the accounting identity.

The upshot of this discussion is that the regressors' endogeneity (or any other standard econometric problem) is not the problem that confronts the estimation of production functions, because all that is being estimated is an accounting identity, or an approximation to it. Hence, the problem that an IV estimator will try to overcome is the one discussed here, namely, a poor approximation to the accounting identity, but in a complicated and artificial way. Given that for most econometric applications factor shares are indeed relatively constant, the problem boils down to how well the particular approach used in estimation helps approximate empirically $g_t^* \equiv (1 - a)\hat{w}_t + a\hat{r}_t$. When will an IV estimator yield plausible estimates? We know precisely what the error term ϕ_t is when one estimates equation (9) (i.e., the unrestricted form corresponding to the identity equation (7)), but the 'true model' is identity (3). It is, i.e., the difference between equations (7) and (3). We also know, given the discussion above, that ϕ_t is procyclical. Therefore, an instrument that is correlated with $\hat{L}_t$ and $\hat{\mathcal{J}}_t$ but that is uncorrelated with ϕ_t will give unbiased estimates. As materials are likely to vary procyclically, they will probably be correlated with the error term. However, in general, it is very difficult to find such an instrument.

Levinsohn and Petrin (2003), for example, claim that the differences between the OLS estimates and those obtained with their approach are consistent with simultaneity (Levinsohn and Petrin, 2003, p. 318). However, paraphrasing (Simon 1979A, 1979B), we believe our arguments provide a more parsimonious explanation for the results obtained and for the true cause of the alleged biases that researchers have claimed to encounter when they estimate production functions. This explanation, however, leads to the uncomfortable conclusion that estimates of production functions do not reflect in any way the underlying technology.

The main issue underlying this accounting identity critique is that *value data*, rather than *physical data* are used in the estimation of production functions. This introduces an insurmountable problem. To see what is so important about the accounting identity, let us

[1] Barro (1999, pp.122–3) has argued that the econometric estimation of the production function suffers from some serious disadvantages (as opposed to growth accounting) as a method of estimating total factor productivity growth. In particular, he lists the following three: (i) the growth rates of capital and labour are not exogenous variables with respect to the growth of output; (ii) the growth of capital is usually measured with error (this often leads to low estimates of the contribution of capital accumulation); and (iii) the regression framework must be extended to allow for variations in factor shares and the TFP growth rate. The arguments here show why this is not the case and what the true nature of the problem is.

assume that firms pursue a pricing policy where the price is determined by a fixed, or constant, mark-up on unit labour costs. We assume for expositional simplicity that no materials are used in production, although, in practice, firms mark-up on normal unit costs (see Lee (1999) for a detailed discussion). Thus, $p_i = (1 + \pi_i)w_i L_i / Q_i$, where p_i is the price in dollars, π_i is the mark-up, and Q_i is the physical measure of output (omitting the time subscripts for notational convenience). Value added for production process i is given by $V_i = p_i Q_i = (1 + \pi_i)w_i L_i$. Labour's share is $(1 - a_i) = 1/(1 + \pi_i)$ and capital's share $a_i = \pi_i/(1 + \pi_i)$, and they will be constant to the extent that the mark-up does not vary. In practice, it is likely to vary owing to changes in the individual mark-ups, which may be temporary, as a result of the wage-bargaining process over the cycle.

Consequently, value added for any industry or firm j, is the sum of the value added of the individual production processes of the firms, that is, $V_j \equiv \sum_i V_i \equiv \sum_i p_i Q_i \equiv \sum_i w_i L_i + \sum_i r_i J_i \equiv \bar{w}_j L_j + \bar{r}_j j_j$, where the bar over w_j and r_j denotes the average value. As all firms are assumed to pursue a constant mark-up pricing policy, the aggregate factor shares for firm, industry or sector are likely to be constant. In fact, as Solow (1958) pointed out, when there is variation in the factor shares, this is likely to be smaller the more aggregated are the data. It can be seen that there are relatively few aggregation problems in the summation of the linear accounting identities above. Note that the aggregate accounting identity holds regardless of the form of the physical micro-production functions at the firm level. It should also be noted that this holds for all industries, even those at the 4-digit SIC level, where output is measured in value terms.

Let us suppose that w and r are the same across the units of observation in a cross-sectional data set, or that they do not change over time when time-series data are used. In other words, we abstract from the problem of finding a good proxy for technical change that was a central concern of the analysis above. If factor shares are constant, then, as we have seen above, an approximation to the accounting identity, using either cross-section or time-series data, will be given by $V = B_0 w^{1-a} r^a L^{1-a} J^a = A L^{1-a} J^a$, where A is a constant. The researcher would obtain a close fit to the data in spite of the fact that the aggregation problems are so severe that no aggregate production function actually exists (Felipe and Fisher, 2003, 2006A, 2006B), even if physically homogeneous data were available for all firms. But the causation is now from the identity to the multiplicative power function, not the other way round. The values of the putative elasticities are determined by the value of the mark-up and do not reflect any technological relationship.

Felipe and McCombie (2006) undertook a simple simulation experiment that illustrated the problems posed by the accounting identity in these circumstances. They assumed that the underlying micro-production functions were given by identical Cobb–Douglas production functions in physical terms, but where the elasticity of labour was 0.25 (instead of the usual 0.75) and the elasticity of capital was 0.75 (instead of 0.25). For expositional ease, they assumed that firms employ a constant mark-up of 1.333 on unit labour costs, which means that the share of labour in the *value* of output is hence 0.75. The resulting constant-price value data (as these are the only data assumed available to the researcher) were used to estimate a cross-industry production function. This has the advantage of showing that the problem is not one of the way in which technical change is modelled. They found that the estimate of the 'output elasticities' of labour and capital were 0.75 and 0.25, respectively, and exactly equal to the relevant factor shares. What is happening is that the causation, as noted above, runs from the accounting identity and the values of the factor shares to the estimated 'output elasticities'. The direction of causality is not from the estimates of the underlying micro-production functions and the marginal productivity theory of factor pricing to the factor shares. (See Felipe and McCombie, 2006, Table 1, for a further elaboration.)[1]

It should also be noted that as the output and capital stock are homogeneous physical measures, a cross-industry production function could be estimated using these data. In these circumstances, of course, the estimates of the output elasticity of labour and capital would be 0.25 and 0.75, respectively. An inference could, in this case, be made that factors were not being paid their marginal products. In practice, we can never recover the physical quantities due to the heterogeneity of output and capital noted above, and so the output for each firm and industry can only be measured in value terms.

Summing up, the reason why the coefficients of estimated value-based production functions often closely approximate the factor shares stems from the fact that the weighted growth of the real wage and the rate of profit closely approximate a constant. To the extent that it differs, it does so procylically. Conventional estimation of production functions is often undertaken with the growth of the factor inputs corrected for (cyclical) changes in utilisation rates. This has the effect of giving a very close approximation to the accounting identity. In the case of cross-section data, $w_j^{1-a}r_j^a$ generally shows little variation or is orthogonal to labour and capital, and here again the regression is merely estimating the accounting identity.[2]

4. An illustration using firm-level data

In this section, we present empirical evidence to support the arguments in Section 3. We take the accounting identity as the primitive and corroborate that its estimation leads to the results discussed above and known *ex ante*. Then we compare these results with those of the 'misspecified identity', i.e., the production function. We use firm-level data from publicly listed Indian firms belonging to a single industry (textiles).[3] First, we show the

[1] Franklin Fisher (1971) once mentioned that Solow remarked to him that if Douglas had found capital's share to be equal to 0.75 instead of 0.25, 'we would not now be talking about aggregate production functions'. What we have shown is that, with value data, the supposed output elasticities must always equal the factor shares.

[2] The argument that all estimations of production functions with value data do is to approximate the income accounting identity has certainly not passed unnoticed by the profession. However, for reasons difficult to explain, it has not been taken to its ultimate conclusions or has been incorrectly dismissed. Econometrics textbooks such as those of Cramer (1969), Intriligator (1978) or Wallis (1980) explain this critique, although the latter two authors, inexplicably, do not take it to its logical conclusion. Intriligator (1978, p. 270), while discussing Cramer's argument, only notes that it leads to a bias in the estimates towards constant returns to scale, and that the factor shares will be approximately equal to the output elasticities. Wallis (1980, pp. 61–3) goes further and concludes that 'perhaps these Cobb–Douglas results and the apparent support for constant returns or marginal productivity theory are not as persuasive as was first supposed'. The clear-cut implication of the argument is that the problem removes entirely the possibility of interpreting the result of estimating a production function as a test of a technological relationship.

[3] We thank the Institute for Studies in Industrial Development, New Delhi, for kindly making these data available to us. The data included the following firm-specific information: gross-output, book value of plant and equipment, total wage bill and expenditures on raw materials, intermediates, fuel and energy. Industry-specific wage rates were used to divide firms' total wage bill to arrive at the number of workers. Real values for gross output, capital stocks and total intermediate inputs were derived by deflating gross output, the book value of plant and machinery and total intermediates (raw materials plus intermediates plus fuel plus energy) by price deflators for textile products, investment goods and intermediate inputs, respectively, obtained from Chandhok (1990). Because the industry-specific deflators pertain to the calendar year, while firms' data pertain to their fiscal years, each of the deflators was adjusted for the fiscal year of the firm.

Table 1. *Time-series data: value-added regression and the accounting identity; Indian textile firms, 1976–1989; OLS estimates; dependent variable is value added*

	Log levels		Growth rates			
	(i)	(ii)	(iii)	(iv)	(v)	(vi)
Constant	0.588 (12.59)***	14.716 (4.60)***	0.000 (0.53)	0.035 (0.96)	−0.003 (3.57)***	0.009 (1.38)
Trend		0.080 (3.78)***				
w	0.289 (43.19)***		0.281 (25.35)***			
r	0.709 (168.00)***		0.706 (145.88)***			
L	0.292 (57.57)***	−0.017 (0.07)	0.278 (35.78)***	0.146 (0.40)	0.308 (36.94)***	0.246 (4.03)***
J	0.710 (127.30)***	−0.816 (2.17)*	0.705 (69.02)***	−0.202 (0.48)	0.744 (64.85)***	0.697 (8.38)***
TFPGR						1.056 (19.02)***
TFPG					1.019 (141.54)***	
No. obser.	14	14	13	13	13	13
D.W. stat.	2.81	2.29	2.65	2.38	2.10	2.09
$\bar{R}^2$	0.999	0.826	0.999	−0.159	0.999	0.968

Notes: Absolute *t*-values in parentheses; * significant at 10%; ** significant at 5%; *** significant at 1%. The average factor shares are 0.70 for capital and (1–0.70) = 0.30 for labour; with ranges 0.67–0.74 for capital and 0.26–0.33 for labour.

results for the time-series case (the case considered so far in the paper) with data for one firm, for both value added and gross output; and then for a cross-section of firms.[1]

4.1 Time-series value-added data

Table 1 summarises the relevant results pertaining to value-added production data for the time-series case. It shows the regressions in log-levels, numbered (i) and (ii), and then in growth rates, numbered (iii), (iv), (v) and (vi). Regression (i), corresponding to equation (4) above, and regression (iii), corresponding to equation (2) above, are the reference cases. All regressions are estimated with the coefficients unrestricted. Regression (i) is the identity rewritten under the only assumption that factor shares are constant. Estimation results of this reference case confirm that what has been estimated must be the identity: the estimates of the coefficients of the wage rate and labour are equal to each other, and equal to the average labour share (provided in the notes to Table 1); and, likewise, the estimates of the coefficients of the profit rate and the capital stock are also equal to each other, and equal to the average capital share (provided in the notes to Table 1). The very high *t*-values and statistical fit (not exactly equal to one because factor shares vary slightly) reinforce the belief that all we are estimating is an identity.

[1] The firm used in the estimation was chosen on the basis of the relative constancy of the factors shares. This choice is deliberate, of course, since the point of our exercise is to illustrate the argument made in Section 3.

The interesting case appears in the second column, regression, (ii) which corresponds to equation (6) above. This is the standard Cobb–Douglas with a linear time trend. The comparison between both regressions is truly revealing: now the estimates of the coefficients of labour and capital are negative. Given the results in regression (i), we know why this has happened: all the time trend does in regression (ii) is to try to approximate the weighted average of the logarithms of the wage and profit rates, i.e., $(1-a)\ln w_t + a\ln r_t$. It was shown above that for this approximation to be correct, wage and profit rates would have to grow at constant rates. All this regression shows is that this is not true. Hence, this approximation to the identity turns out to be extremely poor. This is not an econometric problem in the standard sense. All that is needed is to find a good approximation to the term $(1-a)\ln w_t + a\ln r_t$.

Turning now to the regressions in growth rates, regression (iii) corroborates that all we have is an identity, and the results are virtually identical to those in regression (i). Regression (iv) corresponds to equation (7) above. Once again, results are very poor, with the estimates of labour and capital statistically insignificant, and the latter again negative. In regression (v), we have added the variable $g_t \equiv (1-a_t)\hat{w}_t + a_t\hat{r}_t$ (total factor productivity growth or TFPG) which, by construction, must enter the regression with a coefficient of unity. The estimates of labour and capital are, again, equal to the average factor shares. This regression confirms, as argued in Section 3, that all that is needed is to find a variable that tracks correctly the path of TFPG (e.g., appropriate trigonometric function). This will take us back to the identity. Regression (iv) shows that a constant does not do a good job. Indeed, Figure 1 graphs TFPG and shows that a constant cannot track the path of this variable. Finally, regression (vi) uses as a regressor the term $a_t\hat{r}_t$, denoted TFPGR. This is the most important component of TFPG in determining its fluctuations. Figure 1 shows that indeed this is the case: most of the variation in TFPG is accounted for by TFPGR, and this is confirmed by regression (vi).

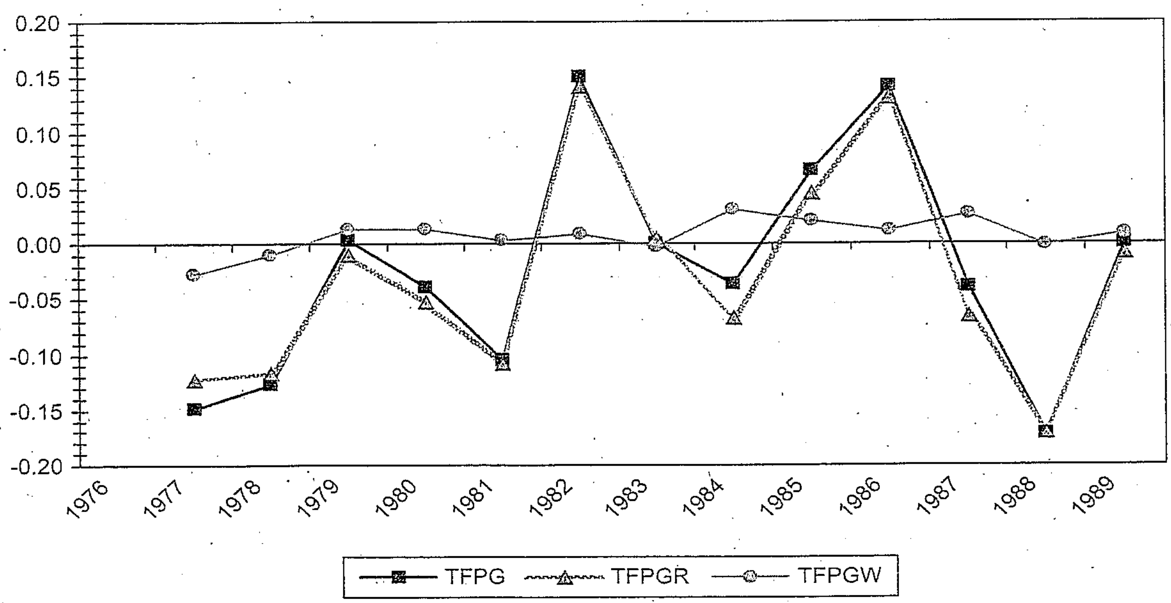

Fig. 1. *Value-added. Total factor productivity growth and its components.*
Note: $TFPG \equiv g_t \equiv (1-a_t)\hat{w}_t + a_t\hat{r}_t$; $TFPGR \equiv a_t\hat{r}_t$; $TFPGW \equiv (1-a_t)\hat{w}_t$.

4.2 Time-series gross-output data

We use now gross output data for the same firm to prove the generality of the argument. In this case, the accounting identity is

$$Y_t \equiv W_t + \pi_t + M_t \equiv w_t L_t + r_t J_t + z_t M_t \tag{12}$$

where Y_t denotes real gross output, M_t is the real value of materials (to simplify we aggregate the value of intermediate materials and energy) and z_t denotes the price of materials in real terms. Denoting $b_t = r_t J_t / Y_t$ and $c_t = z_t M_t / Y_t$ the shares of capital and materials in gross output, respectively, and $(1 - b_t - c_t) = w_t L_t / Y_t$ the share of wages in gross output, and proceeding as above, equation (12) can be rewritten as

$$Y_t \equiv D_0 w_t^{(1-b-c)} r_t^b z_t^c L_t^{(1-b-c)} J_t^b M_t^c \tag{13}$$

where D_0 is a constant.

One can derive similar equations to the ones derived above in growth rates. Now the weighted average of the growth rates of the factor prices is $\tilde{g}_t \equiv (1 - b_t - c_t)\hat{w}_t + b_t \hat{r}_t + c_t \hat{z}_t$.

Table 2 shows the estimation results of equation (13), the approximation to the gross output identity under the sole hypothesis that factor shares are constant, in log levels (regression (i)) and growth rates (regression (iii)). The result that the estimates are close to the three shares (and of approximately equal magnitude for each pair w_t and L_t; r_t and J_t; and z_t and M_t) together with the extremely good fit can only be interpreted as empirical validation of the hypothesis that we are estimating an identity. Indeed, the results indicate that factor shares must be sufficiently constant in the data set so that equation (13), whether estimated in log levels (regression (i) in Table 2) or in growth rates (regression (iii) in Table 2)—both regressions yield virtually the same estimates—provides an excellent approximation to the accounting identity (12). The statistical fit is virtually unity and, most importantly, there is no apparent econometric problem that needs to be taken care of.

We now discuss the results that researchers obtain estimating the standard production function for gross output as $Y_t = A_0' e^{\lambda t} L_t^{\alpha} J_t^{\beta} M_t^{\gamma}$ (regression (ii) in log levels and regression (iv) in growth rates). The differences with the previous results are startling. The estimates are now not plausible and include negative values. The estimated coefficients are substantially different when the specification is in levels and in growth rates. Most researchers would argue that these estimates are the result of endogeneity bias and spuriousness due to the presence of unit roots. Hence, they would search for econometric solutions such as finding a suitable instrumental variable.[1]

What has happened? As argued above, our parsimonious explanation is that the weighted average of the factor prices has not been correctly proxied, thus causing an omitted variable bias. It is not a case of true endogeneity bias. To see why this is the case (for the regression in growth rates), Figure 2 shows the graph of $\tilde{g}_t \equiv (1 - b_t - c_t)\hat{w}_t + b_t \hat{r}_t + c_t \hat{z}_t$ (denoted TFPG), which is being proxied by the constant term in the regression in growth rates. It is obvious that this approximation is so poor that, for all practical purposes, $\tilde{g}_t$ is omitted from the regression. Hence, the other coefficients are biased. Regression (v) in growth rates introduces the variable $\tilde{g}_t \equiv (1 - b_t - c_t)\hat{w}_t + b_t \hat{r}_t + c_t \hat{z}_t$ as a regressor. By construction, the coefficient of this variable has to be unity. This indicates that all is needed is to search for a variable highly correlated with $\tilde{g}_t$.

[1] Regression (ii) in levels displays a very good fit, despite the poor estimates of the coefficients. The fact that the variables display trends contributes to it; but the main reason is still that what is being estimated is an approximation to the identity.

Table 2. *Time-series data: gross output regression and the accounting identity; Indian textile firms, 1976–1989; OLS estimates; dependent variable is gross output*

	Log levels		Growth rates			
	(i)	(ii)	(iii)	(iv)	(v)	(vi)
Constant	1.071 (19.30)***	8.674 (4.86)***	−0.001 (0.90)	0.019 (1.07)	−0.002 (1.97)*	0.009 (1.01)
Trend		0.052 (4.42)***				
w	0.162 (14.72)***		0.166 (12.93)***			
r	0.328 (59.98)***		0.330 (52.20)***			
z	0.518 (27.26)***		0.508 (25.59)***			
L	0.142 (21.33)***	−0.038 (0.24)	0.135 (17.46)***	0.021 (0.11)	0.158 (16.50)***	0.152 (1.59)
J	0.315 (32.78)***	−0.513 (2.32)**	0.329 (28.42)***	−0.124 (0.61)	0.349 (26.20)***	0.388 (2.69)**
M	0.535 (65.28)***	0.547 (2.97)**	0.532 (64.49)***	0.673 (4.67)***	0.535 (68.31)***	0.429 (4.89)***
TFPGR						1.076 (5.14)***
TFPG					1.040 (57.98)***	
No. obser.	14	14	13	13	13	13
D.W. stat.	2.41	2.56	1.87	1.78	2.32	2.33
$\bar{R}^2$	0.999	0.947	0.999	0.632	0.999	0.904

Notes: Absolute t-values in parentheses; * significant at 10%; ** significant at 5%; *** significant at 1%. The average factor shares are 0.33 for capital, 0.53 for materials and (1–0.33–0.53) = 0.14 for labour; with ranges 0.30–0.38 for capital, 0.47–0.56 for materials and 0.13–0.16 for labour.

Figure 2 also shows the three components of $\tilde{g}_t$, namely $(1 - b_t - c_t)\hat{w}_t$ (TFPGW), $b_t\hat{r}_t$ (TFPGR) and $c_t\hat{z}_t$ (TFPGPM). It is worth noting that the variable driving the fluctuations in TFPG is $b_t\hat{r}_t$, while the other two variables contribute very little. Moreover, given the constancy of the shares, movements in $b_t\hat{r}_t$ are driven basically by $\hat{r}_t$.[1] By introducing in the regression $b_t\hat{r}_t$ (regression (vi)), results already improve substantially.

4.3 Cross-section value-added data

Finally, and to dissipate any doubts about the generality of the argument, we now discuss the empirical evidence using cross-sectional data. For a cross-section (and using value-added data to simplify the exposition), the labour share can be written as $(1 - a_j) = w_jL_j/V_j$; and similarly the capital share as $a_j = r_jK_j/V_j$ (where j denotes the units of the cross-section). For a low dispersion in factor shares, the approximation

[1] The correlation between TFPG and $b_t\hat{r}_t$ is 0.91; between TFPG and $\hat{r}_t$ is 0.92; and between $b_t\hat{r}_t$ and $\hat{r}_t$ is 0.99. On the other hand, the correlations between TFPG and $(1 - b_t - c_t)\hat{w}_t$, and between TFPG and $c_t\hat{z}_t$ are much lower (0.58 and 0.10, respectively).

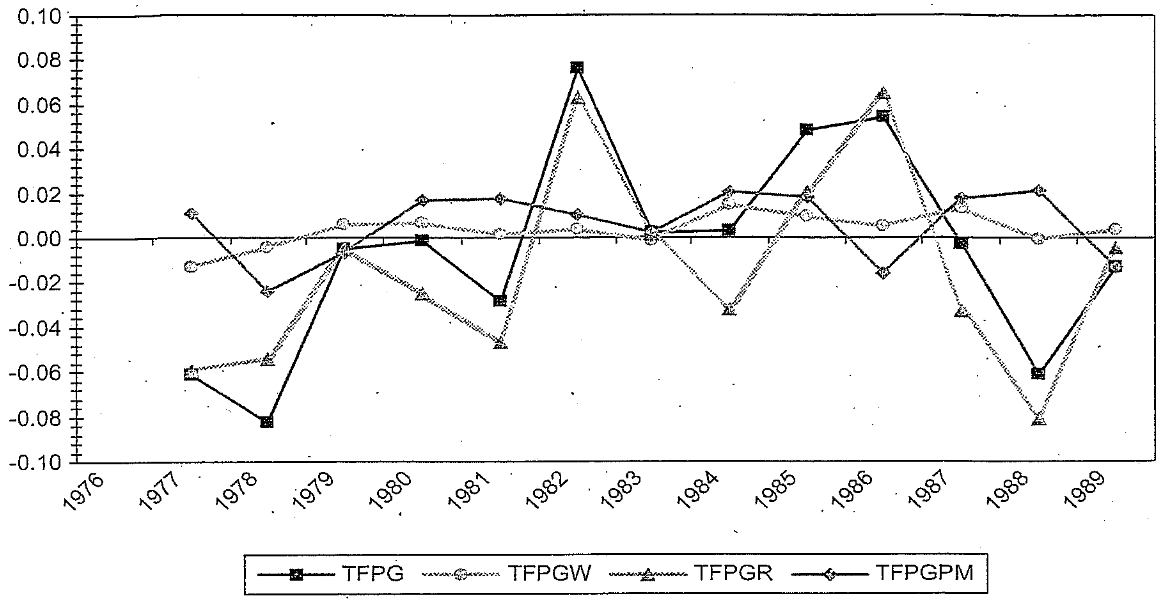

Fig. 2. *Gross output. Total factor productivity growth and its components.*
Note: $TFPG \equiv \tilde{g}_t \equiv (1 - b_t - c_t)\hat{w}_t + b_t\hat{r}_t + c_t\hat{z}_t$; $TFPGR \equiv b_t\hat{r}_t$; $TFPGW \equiv (1 - b_t - c_t)\hat{w}_t$; $TFPGPM \equiv c_t\hat{z}_t$.

$(1 - \bar{a}) = \bar{w}\bar{L}/\bar{V}$, where a bar denotes the average value of the variable, holds. Then the following also holds for the labour share[1]

$$\frac{(1 - a_j)}{(1 - \bar{a})} \cong \left(\frac{w_j L_j}{V_j}\right) / \left(\frac{\bar{w}\bar{L}}{\bar{V}}\right) \tag{14}$$

and a similar equation follows for the capital share:

$$\frac{a_j}{\bar{a}} \cong \left(\frac{r_j \mathcal{J}_j}{V_j}\right) / \left(\frac{\bar{r}\bar{\mathcal{J}}}{\bar{V}}\right) \tag{15}$$

For small deviations of a variable X_j from its mean $\bar{X}$, it follows that $\ln(X_j/\bar{X}) \cong (X_j/\bar{X}) - 1$. Thus, taking logs in equations (14) and (15) and using this approximation, we can write

$$\ln\left(\frac{w_j}{\bar{w}}\right) + \ln\left(\frac{L_j}{\bar{L}}\right) - \ln\left(\frac{V_j}{\bar{V}}\right) \cong \frac{(1 - a_j)}{(1 - \bar{a})} - 1 \tag{16}$$

and

$$\ln\left(\frac{r_j}{\bar{r}}\right) + \ln\left(\frac{\mathcal{J}_j}{\bar{\mathcal{J}}}\right) - \ln\left(\frac{V_j}{\bar{V}}\right) \cong \frac{a_j}{\bar{a}} - 1 \tag{17}$$

Multiplying equations (16) and (17) by $(1 - \bar{a})$ and $\bar{a}$, respectively, adding them up and rearranging the result yields

$$\begin{aligned} \ln V_j &\cong B + (1 - \bar{a})\ln w_j + \bar{a}\ln r_j + (1 - \bar{a})\ln L_j + \bar{a}\ln \mathcal{J}_j \\ &= \ln A(j) + (1 - \bar{a})\ln L_j + \bar{a}\ln \mathcal{J}_j \end{aligned} \tag{18}$$

[1] This derivation follows the arguments in Cramer (1969).

where $B = (\ln\bar{V} - (1-\bar{a})\ln\bar{w} - \bar{a}\ln\bar{r} - (1-\bar{a})\ln\bar{L} - \bar{a}\ln\bar{J}) = (1-\bar{a})\ln(1-\bar{a}) + \bar{a}\ln\bar{a}$ is a constant and $\ln A(j) = \ln B + (1-\bar{a})\ln w_j + \bar{a}\ln r_j$.

As before, the implication of this derivation is that equation (18) may be mistaken for a Cobb–Douglas production function. It is self-evident, however, that equation (18) is an approximation to the income distribution identity.

It must be noted that, in general, it is easier to obtain plausible results with cross-sectional data than with time series. The reason is that, often, wage and profit rates in a cross-section (e.g., regions in a country, firms in a sector) vary relatively little. This implies that the term $\ln A(j)$ in equation (18) will be accurately approximated by the constant term, so that $\ln A(j) \cong \ln A$ will be, effectively, the constant in the regression. This means that the cross-sectional regression $V_j = A_0 L_j^{\alpha} J_j^{\beta} e^{\epsilon_j}$ should work very well provided only that factor shares in the cross-section do not vary excessively.

Table 3 provides the results based on value-added production data using a cross-section of 48 Indian textile firms for 1980.[1] As before, we show first the full approximation to the identity, equation (18) in regression (i), and then the production function, in regression (ii). The results for the full regression corroborate that equation (18) provides an excellent approximation to the accounting identity: virtually a perfect fit and estimates highly significant and equal to the average factor shares. The important difference now with respect to the time-series case is that the Cobb–Douglas regression (ii) works very well, with estimates of the coefficients of L_j and J_j that anyone would take as plausible (although substantially different from those in regression (i)). The reason *must be* that the term $\ln A(j) = \ln B + (1-\bar{a})\ln w_j + \bar{a}\ln r_j$ is minimally and sufficiently (though not perfectly) well approximated by the constant term. Indeed, Figure 3 shows that $(1-\bar{a})\ln w_j + \bar{a}\ln r_j$ does not vary excessively with respect to its average.[2]

Table 3. *Cross-section data: value-added regression and the accounting identity; Indian textile firms, 1980; OLS estimates; dependent variable is value added*

	Log levels	
	(i)	(ii)
Constant	0.578 (2.94)***	0.282 (1.37)
w	0.257 (1.52)	
r	0.736 (71.41)***	
L	0.281 (44.12)***	0.529 (9.15)***
J	0.720 (105.31)***	0.440 (7.27)***
No. obser.	48	48
$\bar{R}^2$	0.999	0.963

Notes: Absolute value of *t*-values in parentheses; * significant at 10%; ** significant at 5%; *** significant at 1%.
The average factor shares are 0.69 for capital and (1–0.69)=0.31 for labour; with ranges 0.57–0.85 for capital and 0.14–0.42 for labour.

[1] This is the industry with the largest number of firms in the data available to us.

[2] The regression of $(1-\bar{a})\ln w_i + \bar{a}\ln r_i$ on a constant yields a value of 7.31, the mean, which is statistically significant at the 1% level. The maximum value of the series is 8.99 and the minimum is 6.05.

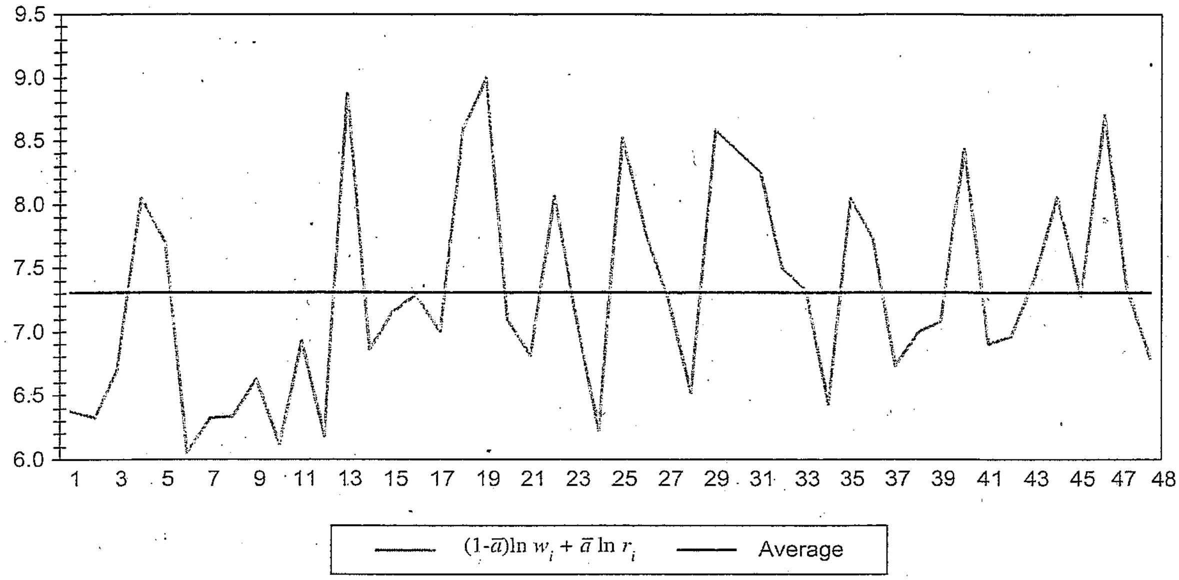

Fig. 3. *Cross section. Weighted logarithm of factor prices. Cross-section data for 48 Indian textile firms.*

5. Conclusions

This paper has shown that the endogeneity bias discussed in the literature that appears when estimating value-based production functions is simply the result of omitted-variable bias due to an incorrect approximation to the accounting identity that relates output to the sum of the total payments to labour (the wage bill) and capital (total profits) (plus materials in the case of gross output). As a result, we have questioned the attempts to solve the problem by developing new estimators. The problem has no econometric solution.

We emphasise that we are not saying that it is impossible to estimate a production function. This could be done if all variables entering the production function were measured in terms of physical quantities. In this case, although there is also an accounting identity for value added (and gross output), the physical data can be recovered from the identity (something impossible in the case above). Under these circumstances, it is possible to estimate a production function.

Physical data are scarce, and estimations of production functions with physical quantities for homogeneous products are the exception, outside of agricultural economics. For example, the recent study of Foster *et al.* (2005) uses plant-level data for 11 highly homogeneous manufacturing products (e.g., corrugated and solid fibre boxes, white pan bread, ready-mixed concrete, etc.). This paper, as well as some of the literature discussed in it, represents an improvement, in that the authors realise that the use of value measures as opposed to physical quantities leads to productivity mismeasurement. Indeed, using simulation analysis, Felipe and McCombie (2006) have shown that the true rate of total factor productivity growth estimated using physical quantities for both output and input differs substantially from that estimated using value data. The Foster *et al.* (2005) paper does not avoid the use of value data for the stock of capital, materials and energy. Thus, effectively, only output and labour are measured in terms of true physical quantities. While these authors do realise the importance of measuring output in physical terms, inputs also have to be measured in physical terms. In fact, in footnote 10 in their paper, they argue: 'in practice, comprehensive input quantity data are rarely available, so expenditures are used instead'. While their approach removes from the estimate of total factor productivity growth fluctuations in a firm's relative price due to, for example, idiosyncratic demand shocks, the contribution of the inputs to output growth is still calculated using value data. Katayama *et al.* (2003), in a related but different context, acknowledge the necessity to work with physical quantities, and discuss (from a different point of view) some of the problems derived from using data expressed in monetary values.

It must be added that, even with physical data, one is not trouble free, as, in practice, the true production function could only be estimated if one knew not only its functional form but also the true path of technical change. Despite these problems, efforts must be made to collect data for homogeneous products, the only ones for which physical quantities can be used for estimating production functions in a meaningful way, rather than spending effort devising new estimation methods. Otherwise, estimation of production functions with variables expressed in monetary terms (however deflated) will continue being a problematic exercise. In contrast, with physical quantities (for both output and inputs) the problems that these new estimators try to address do apply and their attempt at correcting for endogeneity bias is valid.

Bibliography

Barro, R. 1999. Notes on growth accounting, *Journal of Economic Growth*, vol. 64, no. 2, 119–37

Blackorby, C. and Schworm, W. 1984. The structure of economies with aggregate measures of capital: a complete characterization, *Review of Economic Studies*, vol.51, no. 4, 633–50

Blundell, R. and Bond, S. 2000. GMM estimation with persistent panel data: an application to production functions, *Econometric Reviews*, vol. 19, no. 3, 321–40

Chandhok, H. L. 1990. *India Database: The Economy*, vols I–II, New Delhi, Living Media

Cohen, A. and Harcourt, G. C. 2003. Whatever happened to the capital controversies? *Journal of Economic Perspectives*, vol. 17, no. 1, 199–214

Cramer, J. S. 1969. *Empirical Econometrics*, Amsterdam, North Holland

Douglas, P. H. 1948. Are there laws of production? *American Economic Review*, vol. 38, no. 1, 1–41

Felipe, J. and Fisher, F. M. 2003. Aggregation in production functions: what applied economists should know, *Metroeconomica*, vol. 54, nos 2–3, 208–62

Felipe, J. and Fisher, F. M. 2006A. Aggregate production functions, neoclassical growth models and the aggregation problem, *Estudios de Economia Aplicada*, vol. 24, no. 1, 127–63

Felipe, J. and Fisher, F. M. 2008. Aggregation: production, *New Palgrave Dictionary of Economics*, 2nd edn, Basingstoke, Palgrave MacMillan(forthcoming)

Felipe, J. and McCombie, J. S. L. 2001. The CES production function, the accounting identity, and Occam's razor, *Applied Economics*, vol. 33, no. 10, 1221–32

Felipe, J. and McCombie, J. S. L. 2003. Comment on 'Whatever happened to the Cambridge capital theory controversies?', *Journal of Economic Perspectives*, vol. 17, no. 4, 229–31

Felipe, J and McCombie, J. S. L. 2005. How sound are the foundations of the aggregate production function?, *Eastern Economic Journal*, vol. 31, no.3, 467–88

Felipe, J. and McCombie, J. S. L. 2006. The tyranny of the identity: growth accounting revisited, *International Review of Applied Economics*, vol. 20, no. 3, 283–99

Felipe, J. and McCombie, J. S. L. 2007. On the rental price of capital and the profit rate, the perils and pitfalls of total factor productivity growth, *Review of Political Economy*, vol. 19, no. 3, 317–45

Fisher, F. M. 1971. Aggregate production functions and the explanation of wages: a simulation experiment, *Review of Economics and Statistics*, vol. 53, no. 4, 305–25

Foster, L., Haltiwanger, J. and Syverson, C. 2005. 'Reallocation, firm turnover and efficiency: selection on productivity or profitability?', NBER Working Paper, no. 11555

Griliches, Z. and Mairesse, J. 1998. Production functions: the search for identification, pp. 169–203 in Strøm, S. (ed.), *Econometrics and Economic Theory in the 20th Century*, Cambridge, Cambridge University Press

Intriligator, M. D. 1978. *Econometric Models, Techniques and Applications*, Englewood Cliffs, N.J., Prentice Hall

Katayama, H., Lu, S. and Tybout, J. R. 2003. 'Why plant-level productivity studies are often misleading, and an alternative approach to inference?', NBER Working Paper, no. 9617

Lee, F. S. 1999. *Post Keynesian Price Theory*, Cambridge, Cambridge University Press

Levinsohn, J. and Petrin, A. 2003. Estimating production functions using inputs to control for unobservables, *Review of Economic Studies*, vol. 70, no. 2, 317–41

Lucas, R. E. 1970. Capacity, overtime, and empirical production functions, *American Economic Review, Papers and Proceedings*, vol. 60, no. 2, 23–7

Marschak, J. and Andrews, W. H. 1944. Random simultaneous equations and the theory of production, *Econometrica*, vol. 12, nos 32–4, 143–205

Olley, S. and Pakes, A. 1996. The dynamics of productivity in the telecommunications equipment industry, *Econometrica*, vol. 64, no. 6, 1263–98

Phelps Brown, E. H. 1957. The meaning of the fitted Cobb–Douglas function, *Quarterly Journal of Economics*, vol. 71, no. 4, 546–60

Samuelson, P. A. 1979. Paul Douglas's measurement of production functions and marginal productivities, *Journal of Political Economy*, vol. 87, no. 5, 923–39

Shaikh, A. 1974. Laws of production and laws of algebra: the Humbug production function, *Review of Economics and Statistics*, vol. 56, no. 1, 115–20

Shaikh, A. 1980. Laws of production and laws of algebra: Humbug II, pp. 80–95 in Nell, E. J. (ed.), *Growth, Profits and Property. Essays in the Revival of Political Economy*, Cambridge, Cambridge University Press

Shaikh, A. 2005. Nonlinear dynamics and pseudo-production functions, *Eastern Economics Journal*, vol. 31, no. 3, 447–66

Simon, H.A. 1979A. On parsimonious explanations of production relations, *Scandinavian Journal of Economics*, vol. 81, no. 4, 459–74

Simon, H.A. 1979B. Rational decision making in business organizations, *American Economic Review*, vol. 69, no. 4, 493–513

Simon, H.A. and Levy, F. 1963. A note on the Cobb–Douglas function, *Review of Economic Studies*, vol. 30, no. 2, 93–4

Solow, R.M. 1958. A skeptical note on the constancy of relative shares, *American Economic Review*, vol.48, no. 4, 618–31

Wallis, K. F. 1980. *Topics in Applied Econometrics*, 2nd en revised, Minneapolis, University of Minnesota Press