CENTRE FOR APPLIED MACROECONOMIC ANALYSIS

The Australian National University

CAMA Working Paper Series **November, 2004**

CORRECTING FOR BIASES WHEN ESTIMATING PRODUCTION FUNCTIONS: AN ILLUSION OF THE LAWS OF ALGEBRA?

Jesus Felipe
Asian Development Bank
Manila
Philippines

Rana Hasan
Asian Development Bank
Manila
Philippines

J.S.L McCombie
Downing College
Cambride
United Kingdom

CAMA Working Paper 13/2004
http://cama.anu.edu.au

CORRECTING FOR BIASES WHEN ESTIMATING PRODUCTION FUNCTIONS: AN ILLUSION OF THE LAWS OF ALGEBRA?

Jesus Felipe
Economics and Research Department
Asian Development Bank
Manila, Philippines
jfelipe@adb.org

Rana Hasan
Economics and Research Department
Asian Development Bank
Manila, Philippines
rhasan@adb.org

JSL McCombie
Univeristy of Cambridge
Cambridge, UK
jslm2@hermes.cam.ac.uk

Abstract: This paper argues that the true cause of the endogeneity bias that allegedly appears when estimating production functions, and which the literature has tried to deal with since the 1940s, is simply the result of omitted-variable bias due to an incorrect approximation to an accounting identity. As a result we question recent attempts to solve the problem by developing new estimators.

This paper represents the views of the authors and do not represent those of the Asian Development Bank, its Executive Directors, or the countries that they represent. We are grateful to Dave Dole for his comments.

CORRECTING FOR BIASES WHEN ESTIMATING PRODUCTION FUNCTIONS: AN ILLUSION OF THE LAWS OF ALGEBRA?

1. Introduction

A series of recent papers (e.g., Olley and Pakes 1996, Blundell and Bond 2000, Levinsohn and Petrin 2003), has proposed new methods for conditioning out serially correlated shocks to the production technology.[1] As is well-known, the OLS estimates of the parameters of the production function are biased and inconsistent if there is contemporaneous correlation between the error term and the factor inputs due to simultaneity problems (however, see Zellner et al. 1966 for a discussion of when this is not likely to be a problem). An instrumental variable (IV) estimator achieves consistency by instrumenting the factor inputs with regressors that are correlated with them but uncorrelated with the error term. Olley and Pakes (1996) propose to instrument the regression with investment. On the other hand, Levinsohn and Petrin (2003) use intermediate materials as instrument. 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.

The purpose of this paper is to demonstrate that the alleged endogeneity problem to which these papers refer to is not the cause of the implausible results that often appear in empirical estimations of production functions. In fact, the problem is substantially more serious than one requiring the correction for possible simultaneity biases, and unfortunately it is one that has no econometric solution. This is an implication of an argument which appeared in an early paper by Phelps Brown (1957), who outlined a criticism of Douglas's (1948) cross-industry results. This critique was later formalized by Simon and Levy (1963). Later, 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 acceptance speech (Simon 1979b). Paradoxically, the same year, Samuelson (1979) also rediscovered the same argument.[2] And Shaikh (1980) offers a very provocative discussion of it. 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 summarize it and briefly discuss its implications.[3]

---

[1] See Levinsohn and Petrin (2000) for an extended set of results.

[2] Marschak and Andrews (1944), whom some of these papers mention for being aware of the endogeneity problem in the estimation of production functions, also had mentioned in passing the same issue; their analysis, however, was flawed. We discuss it later in the paper.

[3] We believe 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. See below our 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 into a form that resembles a production function. The implications are far reaching: if the functional form chosen for fitting a production function provides a good approximation to the identity, the statistical fit obtained will generally be very high, potentially 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 problematical, if not useless. Theoretical problems like 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. If, however, the results obtained are not good (e.g., implausible estimates of the elasticities), it is simply because the wrong functional form as an approximation to the identity has been estimated. The solution to this problem is embedded in the same argument, as shown below.

Despite the well-known theoretical problem of endogeneity bias, in applied work most researchers often commence by estimating the production function using OLS, hoping to obtain estimates of the labor and capital output elasticticities that look plausible and interpretable under a theory. In the simplest case of the Cobb-Douglas function (still the most widely used form) this means that the estimates of the parameters should be relatively close to the factor shares, thus adding up to unity (or perhaps exhibiting some increasing returns). If the results appear plausible, many researchers generally would not be particularly concerned and would not use more sophisticated estimation methods. Thus, they would ignore the simultaneity problem and implicitly rely on Wold's "proximity theorem" (Wold and Faxer 1957). But if, on the other hand, acceptable results were not found, as it is often the case in particular with time-series data, it would be necessary to search for an explanation for the anomalous results (e.g., estimated coefficients are biased due to endogeneity) and to use more sophisticated estimation methods (e.g., IV).

It is important to note that researchers in many studies use 3 or 4-digit level industry data. However, as Levinsohn and Petrin (2003, p. 323, footnote 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.[4] This is also the case even in studies that use firm or plant-level data, since a firm can be producing different products (which most likely use different production

[4] The problems regarding capital are worse. Due to space constraints we do not address them. Suffice to say that the standard measures of the stock of capital used in estimation of production functions (calculated through perpetual inventory at aggregate levels or as the book value of inventories if firm-level data is 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; Felipe and Fisher 2003).

processes) and thus the only way to report an "aggregate" measure of output is in monetary or value terms. Physical data are scarce, and estimations of production functions with physical quantities are the exception, except perhaps in agricultural economics.[5] Certainly, at higher levels of aggregation (e.g., manufacturing sector, total economy), the problem is even more clear and the well-known aggregation problems become an insurmountable difficulty (Felipe and Fisher 2003, Felipe and McCombie 2003). The implication of this observation about the nature of the data used in these studies is that the series of output, labor and capital are related, definitionally, through an income accounting identity.

The rest of the paper is structured as follows. Section 2 summarizes and discusses the argument's main implications. Section 3 provides empirical evidence. Section 4 concludes.

2. The Production Function and the Tyranny of the Accounting Identity

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

$$Y_t^n \equiv P_t Y_t \equiv \mathrm{W}_t^n + \Pi_t^n \equiv w_t^n L_t + r_t^n K_t \qquad (1)$$

where Y^n and Y are nominal and real value added, respectively; P is the output deflator; W^n is the total (nominal) wage bill; and Π^n denotes total (nominal) profits. The symbol $\equiv$ indicates that expression (1) is an accounting identity, not a behavioral model and holds irrespective of the state of competition and the degree of returns to scale.[6] Each of these two components (W, Π) can be rewritten as the product of the *average* factor price times the quantity of the factor. Thus w^n is average nominal wage rate, r^n is the average ex-post nominal profit rate (not the user cost of capital), L is the number of workers, and K is the stock of capital.[7] The decomposition of the wage bill and total profits into the products of the factor prices times the quantities is definitional. The identity can also be defined in real

5 See Wibe (1984) for a survey of engineering production functions.

6 Expression (1) should not be confused with the similar expression that can be derived from Euler's theorem, which assumes that the first-order conditions hold. That identity will hold *only* if the production function from which it is derived is linearly homogeneous, and if the conditions for producer equilibrium hold. No such requirements are necessary for equation (1) to hold always. Moreover, although it will always be true that the wage bill can be written as the product of the average wage rate times employment, whether the wage rate equals the marginal product of labor or not, is quite another matter and irrelevant for purposes of writing equation (1). The same argument applies to total profits.

7 The profit rate is calculated as $r_t^n = (\Pi_t^n / K_t) = (Y_t^n - w_t^n L_t) / K_t$ and correspondingly in real terms.

terms with $w = (w^n / P)$ the average real wage rate, and $r = (r^n / P)$ the average real ex-post profit rate.

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

$$\hat{Y}_t \equiv (1 - a_t)\hat{w}_t + a_t\hat{r}_t + (1 - a_t)\hat{L}_t + a_t\hat{K}_t$$
$$\equiv g_t + (1 - a_t)\hat{L}_t + a_t\hat{K}_t \qquad (2)$$

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

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

$$\hat{Y}_t \equiv (1 - a)\hat{w}_t + a\hat{r}_t + (1 - a)\hat{L}_t + a\hat{K}_t$$
$$\equiv g_t^* + (1 - a)\hat{L}_t + a\hat{K}_t \qquad (3)$$

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

$$ln\,Y_t = A_0 + (1 - a)\,ln\,w_t + a\,ln\,r_t + (1 - a)\,ln\,L_t + a\,ln\,K_t \qquad (4)$$

and taking anti-logarithms we obtain

$$Y_t \equiv B_0 w_t^{1-a} r_t^a L_t^{1-a} K_t^a \qquad (5)$$

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 $g \equiv a\hat{r} + (1 - a)\hat{w}$. Then equation (5) can now be written as

$$Y_t \equiv B_0\, e^{gt} L_t^{1-a} K_t^a \qquad (6)$$

In growth rates this is simply

$$\hat{Y}_t \equiv g + (1 - a)\hat{L}_t + a\hat{K}_t \qquad (7)$$

The important point of this derivation is that expression (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 constant growth rates of the wage and profit rates.

The implications of this derivation are serious and, ultimately, prevent any unambiguous interpretation of the estimated parameters of an alleged production function. Suppose one estimates by OLS $Y_t = C_0\, e^{\lambda t}\, L_t^{\alpha}\, K_t^{\beta}\, exp(\varepsilon_t)$ (unrestricted), where ε_t is the random disturbance and where *Y*, *L* and *K* are the same series as in (1), namely, real value added, employment and the stock of capital. If the two assumptions above happen to be correct, it is obvious, by comparison with (6), that the statistical fit will be perfect, and the estimates will be $\lambda \equiv a\hat{r} + (1 - a)\hat{w} \equiv g$, $\alpha \equiv a$, and $\beta \equiv 1 - a$. Under a neoclassical interpretation, the equality of the elasticities to 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 is a weighted average of the (constant) growth rates of the wage and profit rates.[8] 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.[9] 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 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 due to endogeneity bias. However, implausible results simply mean that one or both assumptions used to derive equation (6) from equation (2) does not correspond empirically to the data. There are several ways to "solve" the problem of implausible results.

(i) If factor shares vary substantially, what is needed is to determine the mathematical form of the empirical path of the shares. Once found, one simply has to proceed as above, that is, substitute it into equation (2) and integrate. This, of course, can give rise to functional forms that will resemble the CES or translog,

---

8 We realize that the expression $g_t \equiv (1 - a_t)\hat{w}_t + a_t\hat{r}_t$ is what standard analyses refer to as total factor productivity (TFP) growth. Our argument does not deny this, in the sense that the expression can be referred to as such. However, the reader must not forget 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 expression 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. 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 estimated making a series of assumptions. This means that there is no way to know and test whether the number computed is correct or not.

9 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.

for example, and which, when estimated, will lead to better results in terms of statistical fit and elasticities equal to the factor shares. In general, a production function *Y=A F(K,L)* may be expressed in growth rates and estimated econometrically as

$$\hat{Y}_t = \lambda_t + \alpha_t \hat{L}_t + \beta_t \hat{K}_t + \upsilon_t \quad (8)$$

where α_t and β_t are the output elasticities of labor and capital (which, in general, vary in time), λ_t is the growth rate of total factor productivity, and v_t is the error term. On the other hand, the accounting identity is given by equation (2). A comparison of both expressions indicates that any functional form (or estimation procedure such as a time-varying estimation method) that gives a good approximation to the identity could also be mistakenly interpreted as a production function.

(ii) If wage and profit rates do not grow at constant rates, one also has to proceed as in (i). It turns out that empirically this is the assumption that is more likely to be incorrect, and is the one that most often produces a poor statistical fit to the Cobb-Douglas with an exponential time trend. As $g_t \equiv (1-a_t)\hat{w}_t + a_t\hat{r}_t$ cannot be empirically approximated by a constant term, 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 that in this case factor shares are constant so that the accounting identity in growth rates corresponds effectively to expression (3). For all practical purposes, when one estimates the Cobb-Douglas function unrestricted in growth rates

$$\hat{Y}_t = \lambda + \alpha \hat{L}_t + \beta \hat{K}_t + v_t \quad (9)$$

the econometric problem is akin to one of omitted-variable bias, i.e., a relevant variable has been omitted from the regression. To see this, compare expressions (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? (Levinsohn and Petrin 2003, Griliches and Mairesse 1998). A comparison of both expressions indicates that what the constant term λ "tries to do" in the Cobb-Douglas regression is to approximate the term $g_t^* \equiv (1-a)\hat{w}_t + a\hat{r}_t$ in the accounting identity (3). However, this seldom works as g_t^* displays significant fluctuations. Hence the estimates of α and β appear to be biased, but not due to endogeneity of the regressors.

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

$$\hat{\beta} = \frac{Cov(\hat{Y}_t\hat{K}_t)Var(\hat{L}_t) - Cov(\hat{Y}_t\hat{L}_t)Cov(\hat{L}_t\hat{K}_t)}{Var(\hat{L}_t)Var(\hat{K}_t) - [Cov(\hat{L}_t\hat{K}_t)]^2} \tag{10}$$

To calculate the bias, substitute (3), the true model, into expression (10) for $\hat{Y}$. Working out the algebra and taking expectations yields:

$$E(\hat{\beta}) = a + \frac{Cov(\hat{K}_t g_t^*)Var(\hat{L}_t) - Cov(\hat{L}_t g_t^*)Cov(\hat{L}_t\hat{K}_t)}{Var(\hat{L}_t)Var(\hat{K}_t) - [Cov(\hat{L}_t\hat{K}_t)]^2} \tag{11}$$

where, since equation (3) is an identity, the theoretical error term is uncorrelated with the regressors (i.e., there is no endogeneity problem); and a similar expression can be derived for the labor coefficient. Expression (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 the sum) due to the incorrect specification of g_t^* as a constant λ plus the shock v_t. If $\hat{\beta} = a$, this will imply that the bias in (11) is zero; but this will simply mean that g_t^* is a constant, or that, by coincidence, the numerator equal zero. Given that the denominator of the bias is always positive, the sign is determined by the numerator. Likewise, the magnitude of the bias will simply be a function of how far g_t^* is from a constant.[10] Indeed, if it is a constant, the bias will be zero. Why is the capital coefficient often "biased downward" (Levinsohn and Petrin 2003, p.333)? This is the case because the variable that drives the fluctuations in g_t^* is $\hat{r}_t$ (wage rate and factor shares do not fluctuate that much), which is a highly procyclical variable, hence highly correlated with $\hat{L}$. On the other hand, the growth of capital ($\hat{K}$) shows virtually no cyclical fluctuations.[11]

(iii) Often researchers use a capacity-utilization adjusted capital stock ($\hat{K}^*$) instead of the unadjusted one ($\hat{K}$). 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^* in the Cobb-Douglas regression, the reason why use of

---

[10] Expression (11) is the same Levinsohn and Petrin (2003) show (unnumbered) 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).

[11] 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 what g_t is; namely, the dual of total factor productivity growth. However, the reader must not forget that all our derivation stems from an accounting identity (there is no behavioral model here), and that if g_t is appropriately included or proxied in the regression, one will always obtain the same suspicious results: perfect fit and elasticities equal to the factor shares.

$\hat{K}^*$ works is that fluctuations in the wage rate ($\hat{w}$) are much smaller than those in the profit rate ($\hat{r}$), a highly procyclical variable; and most often $\hat{r}$ is highly correlated with $\hat{Y}$ and $\hat{L}$. Therefore, $\hat{K}^*$ will be procyclical by construction and thus the specification of the production function with $\hat{K}^*$ will more closely approximate the underlying identity as this variable "proxies" the movements of the omitted variable $\hat{r}$.

In all these three cases the result is that the accounting identity will be more closely approximated by the chosen functional form, with the result that the estimated "output elasticities" will closely approximate the factor shares.[12]

Finally, regressors' endogeneity (or any other standard econometric problem) is not the problem because all that is being estimated is an accounting identity, or an approximation to it. What any IV estimator will do (but without any guarantee) is to overcome the problem discussed here in a complicated and artificial way. Given that for most econometric applications factor shares are indeed relatively constant, the problem boils down to correctly approximating empirically $g_t^* \equiv (1-a)\hat{w}_t + a\hat{r}_t$. When will an IV estimator improve the results? We know precisely what the error (*e*) is when one estimates equation (7) (which in unrestricted form is (9)), but the true model is the identity (3), namely, $e=(g_t^* - g)$; and we also know, given the discussion above, that it is procyclical. Therefore, an instrument that is correlated with $\hat{L}$ and $\hat{K}$ but that is uncorrelated with *e* will give unbiased estimates. However, it is very difficult to find such an instrument. Since materials are likely to vary procyclically, they will probably be correlated with the error term.

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 (1979), we believe our arguments provide a more parsimonious explanation for the results obtained and for the true cause of the alleged biases 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.[13]

---

12 Barro (1999, pp.122-123) has argued that the econometric estimation of the production function suffers from some serious disadvantages (as opposed to growth accounting) as a method to estimate total factor productivity growth. In particular, he lists the following three: (i) the growth rates of capital and labor 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 the true nature of the problem.

13 A reader may argue that for this to be the case, the data must satisfy the accounting identity (1) above. The data on output, labor and capital, however, *must* satisfy the identity. We are grateful to Dave Dole for stressing this point.

Econometrics textbooks like 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.61-63) 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."[14] 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.

Authors have quite correctly pointed out that Marschak and Andrews (1944) were aware of the endogeneity problem, as mentioned above. However, while these authors also hinted the problem discussed in this paper, they argued that Douglas had been fitting a hybrid of a cost and production function, and that he had confused it with the true production function. Marschak and Andrews' argument was that there was an *identification problem*, similar to that of supply and demand. The implication of their argument was that not all was lost, as it should be theoretically possible to find exogenous variables to identify the production function. Bronfenbrenner (1944) also seems to suggest that all estimations of production functions were doing was estimating the cost identity. Marschak and Andrews erroneously dismissed this argument as they considered that a "zero-profit" condition as required, a situation they considered was empirically unlikely. Bronfenbrenner (1971, pp.399-400) returned to the argument but now considered that the inclusion of a time trend was sufficient to distinguish the production function from the identity. These arguments are erroneous, and in our opinion, the problem is insoluble: inclusion of the time trend in the production function would identify the other equation, namely, the identity, and not the production function, though the idea of identifying an identity is certainly not a meaningful one. Strictly speaking, we do not have an identification problem, as the estimates are always, unambiguously, those of the income accounting identity, and are known in advance.

3. Empirical Evidence

In this section we present empirical evidence to support the arguments in the previous section. We use firm-level data from publicly listed Indian firms.[15] First

---

[14] Wallis (1980, p.62) derived the argument by assuming that wage and profit rates are constant. Hence we infer that he believes it holds *only* in this case. As we have shown the argument is more general.

[15] We thank the Institute for Studies in Industrial Development, New Delhi, for kindly making this 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

we show the results for the time-series case (our argument so far in the paper) with data for one firm; and then for a cross-section of firms.

Table1 summarizes the relevant results. We show the regressions in levels, numbered (1) and (2), and then in growth rates, numbered (3), (4), (5) and (6). Regression (1), corresponding to equation (4) above, and regression (3), corresponding to equation (2) above, are the reference points. All regressions are estimated unrestricted. The regression in levels is the identity rewritten under the only assumption that factor shares are constant. Estimation results show that indeed this is the case. The estimates of the coefficients of the wage rate and labor are equal to each other, and equal to the average labor share. 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. The very high t-values and the statistical fit indicate that we are in the presence of an identity.

TABLE 1 ABOUT HERE

The interesting case appears in the second column, regression (2), 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 labor and capital are negative. Given the results in regression (1), we know why this has happened: all the time trend does in regression (2) 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 that 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 is needed is to find the correct approximation to $(1 - a) \ln w_t + a \ln r_t$.

Turning now to the regression in growth rates, equation (3) corroborates that all we have is an identity and results are virtually identical to those in regression (1). Regression (4) corresponds to equation (7) above. Once again, results are very poor, with the estimates of labor and capital statistically insignificant, and the latter again negative. In regression (5) we have constructed the variable $g_t \equiv (1 - a_t) \hat{w}_t + a_t \hat{r}_t$ (*TFPG*), which, by construction must enter the regression with a coefficient of unity. The estimates of labor and capital are, again, equal to the average factor shares. This regression confirms that all is needed is to find a

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), respectively, by an industry-specific price deflator for total intermediates. Because the industry-specific deflators pertain to calendar year while firms' data pertain to their fiscal years, each of the deflators was adjusted for the fiscal year of the firm.

variable that tracks correctly the path of *TFPG*. Regression (4) shows that a constant does not do a good job. Indeed, Figure 1 graphs *TFPG*. A constant cannot track this path. Finally, regression (6) uses as a regressor the term $a_t\hat{r}_t$, denoted *TFPGR*. This is the most important component of *TFPG*, the one that makes it fluctuate the way it does. Figure 1 shows that indeed this is the case: most of the variation in *TFPG* is accounted for by *TFPGR*.

FIGURE 1 ABOUT HERE

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

$$Q_t^n \equiv P_t Q_t \equiv \mathrm{W}_t^n + \Pi_t^n + \mathrm{M}_t^n \equiv w_t^n L_t + r_t^n K_t + z_t^n M_t \qquad (12)$$

where Q_t^n and Q_t denote nominal and real gross output, respectively (for simplicity we use the same deflator *P* as for value added), M_t^n is the nominal value of materials (to simplify we lump together intermediate materials and energy), M_t is the real value of materials and z_t^n denotes the price of materials. Denoting $b_t = (r_t K_t)/Q_t$ and $c_t = (z_t M_t)/Q_t$ the shares of capital and materials in gross output, respectively, and $1 - b_t - c_t = (w_t L_t)/Q_t$ the share of wages in gross output, and proceeding as above, equation (12) can be rewritten as

$$Q_t \equiv B_0 w_t^{1-b-c} r_t^b z_t^c L_t^{1-b-c} K_t^b M_t^c \qquad (13)$$

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

Table 2 shows the estimation of equation (13), the approximation to the gross output identity under the sole assumption that factor shares are constant, in levels (first regression) and growth rates (third regression). The result that the estimates are close to the three shares (and of approximately equal magnitude for *w* and *L*, *r* and *K*, and *z* and *M*) together with the extremely high fit can only be interpreted as empirical validation of the said hypothesis. The results indicate that indeed factor shares must be sufficiently constant in the data set so that (13), whether estimated in levels (regression (1)) or in growth rates (regression (3)) -both regressions yield virtually the same estimates-, provides an excellent approximation to the accounting identity (12). The statistical fit is virtually unity, and very important, there is no econometric problem that needs to be taken care of.

TABLE 2 ABOUT HERE

We now discuss the results that researchers obtain estimating the standard production function for gross output as $Q_t = B_0 e^{\lambda t} L_t^{\alpha} K_t^{\beta} M_t^{\gamma}$ (regression (2) in log levels and regression (4) in growth rates). The differences with the previous results are startling. Estimates are now not plausible, including negative values, and coefficients are substantially different in levels and 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 some solution is needed.

What has happened? As argued above, our parsimonious explanation is that the weighted average of the factor prices has not been correctly proxied 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 $v_t \equiv (1 - b_t - c_t)\hat{w}_t + b_t\hat{r}_t + c_t\hat{z}_t$ (*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, v_t is omitted. Hence the other coefficients are biased. Regression (5) in growth rates introduces the variable $v_t \equiv (1 - b_t - c_t)\hat{w}_t + b_t\hat{r}_t + c_t\hat{z}_t$, denoted *TFPG*, 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 v_t.

FIGURE 2 ABOUT HERE

Figure 2 also shows the three components of v_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 movements 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 $\hat{r}_t$.[16] By introducing in the regression $b_t\hat{r}_t$ (regression (6)), results already improve substantially.

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 a value-added to simplify), the labor share can be written as $1 - a_i = (w_i L_i / Y_i)$; and similarly the capital share as $a_i = (r_i K_i / Y_i)$ (where i denotes the units of the cross section). For a low dispersion in factor shares, the

---

[16] 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).

approximation $1 - \overline{a} = (\overline{w}\,\overline{L}/\overline{Y})$, where a bar denotes the average value of the variable, holds. Then the following also holds for the labor share[17]

$$(1-a_i)/(1-\overline{a}) \cong (w_i/\overline{w})(L_i/\overline{L})/(Y_i/\overline{Y}) \qquad (14)$$

and a similar expression follows for the capital share

$$(a_i/\overline{a}) \cong (r_i/\overline{r})(K_i/\overline{K})/(Y_i/\overline{Y}) \qquad (15)$$

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

$$\ln(w_i/\overline{w}) + \ln(L_i/\overline{L}) - \ln(Y_i/\overline{Y}) \cong [(1-a_i)/(a_i/\overline{a})] - 1 \qquad (16)$$

and

$$\ln(r_i/\overline{r}) + \ln(K_i/\overline{K}) - \ln(Y_i/\overline{Y}) \cong (a_i/\overline{a}) - 1 \qquad (17)$$

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

$$\ln Y_i \cong B + (1-\overline{a})\ln w_i + \overline{a}\ln r_i + (1-\overline{a})\ln L_i + \overline{a}\ln K_i =$$

$$= B + A(i) + (1-\overline{a})\ln L_i + \overline{a}\ln K_i \qquad (18)$$

where $B = (\ln\overline{Y} - (1-\overline{a})\ln\overline{w} - \overline{a}\ln\overline{r} - (1-\overline{a})\ln\overline{L} - \overline{a}\ln\overline{K}) = 0.$

As before, the implication of the derivation is that equation (18) may be mistaken for a production function. 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 *A(i)* in equation (18) will be well approximated by the constant term, so that *A(i)* will be, effectively, a constant. This means that the cross-sectional regression $Y_i = C_0 L_i^{\alpha} K_i^{\beta}$ should work very well provided only that factor shares in the cross-section do not vary excessively.

Table 3 provides the results for a cross-section of 48 Indian firms for 1980.[18] As before, we show first the full approximation to the identity, equation (18) in regression (1), and then the production function, regression (2). The results for the

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

[18] These are firms producing textiles, the industry with the largest number of firms in the data available to us.

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 works very well, with estimates that anyone would take as plausible. The reason *must be* that the term $A(i) = (1 - \bar{a}) ln\, w_i + \bar{a}\, ln\, r_i$ is minimally and sufficiently (though not perfectly) well approximated by the constant. Indeed, Figure 3 shows that $A(i)$ does not vary excessively with respect to the average.[19]

TABLE 3 ABOUT HERE

FIGURE 3 ABOUT HERE

4. Conclusions

This paper has shown that the true cause of the alleged 'endogeneity bias' that allegedly appears when estimating 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 labor (the wage bill) and capital (total profits) (plus materials in the case of gross output). As a result we have questioned recent attempts to solve the problem by developing new estimators.

We emphasize that we are not saying that it is absolutely impossible to estimate a production function. The argument in this paper does not apply when all variables entering the production function are 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 and test the marginal productivity theory of distribution. It must be noted, however, that even with physical data, there is no guarantee that that this will be the case, since the true production function could only be estimated if we knew what the true path of technical change is, which of course is not possible.

Despite these problems, this indicates that efforts must be made at collecting data for homogeneous products, the only ones for which physical quantities exist, rather than spending efforts at 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

[19] The regression of $A(i)$ on a constant yields a value of 7.31, the mean, statistically significant at 1%. The maximum value of the series is 8.99 and the minimum is 6.05.

physical quantities the problems that these new estimators try to address do apply and their attempt at correcting for endogeneity bias is valid.[20]

[20] 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.

TABLE 1. TIME SERIES. VALUE ADDED REGRESSION AND THE ACCOUNTING IDENTITY. 1976-1989. OLS ESTIMATES.

	LOG LEVELS		**GROWTH RATES**			
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Constant	0.588	14.716	0.000	0.035	-0.003	0.009
	(12.59)***	(4.60)***	(0.53)	(0.96)	(3.57)***	(1.38)
Trend		0.080				
		(3.78)***				
w	0.289		0.281			
	(43.19)***		(25.35)***			
r	0.709		0.706			
	(168.00)***		(145.88)***			
L	0.292	-0.017	0.278	0.146	0.308	0.246
	(57.57)***	(0.07)	(35.78)***	(0.40)	(36.94)***	(4.03)***
K	0.710	-0.816	0.705	-0.202	0.744	0.697
	(127.30)***	(2.17)*	(69.02)***	(0.48)	(64.85)***	(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
$\overline{R}^2$	0.999	0.83	0.999	-0.16	0.999	0.97

Notes: Absolute value of t statistics in parentheses; * significant at 10%; ** significant at 5%; *** significant at 1%.

The average factor shares are =0.70 (capital) and (1-0.70)=0.30 (labor); with ranges 0.67-0.74 (capital) and 0.26-0.33 (labor).

TABLE 2. TIME SERIES. GROSS OUTPUT REGRESSION AND THE ACCOUNTING IDENTITY. 1976-1989. OLS ESTIMATES

	LOG LEVELS		**GROWTH RATES**			
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Constant	1.071	8.674	-0.001	0.019	-0.002	0.009
	(19.30)***	(4.86)***	(0.90)	(1.07)	(1.97)*	(1.01)
Trend		0.052				
		(4.42)***				
w	0.162		0.166			
	(14.72)***		(12.93)***			
r	0.328		0.330			
	(59.98)***		(52.20)***			
z	0.518		0.508			
	(27.26)***		(25.59)***			
L	0.142	-0.038	0.135	0.021	0.158	0.152
	(21.33)***	(0.24)	(17.46)***	(0.11)	(16.50)***	(1.59)
K	0.315	-0.513	0.329	-0.124	0.349	0.388
	(32.78)***	(2.32)**	(28.42)***	(0.61)	(26.20)***	(2.69)**
M	0.535	0.547	0.532	0.673	0.535	0.429
	(65.28)***	(2.97)**	(64.49)***	(4.67)***	(68.31)***	(4.89)***
TFPGR						1.076
						(5.14)***
TFPG					1.040	
					(57.98)***	
No. Obser.	14	14	13	13	13	13
D.W. stat.	2.406	2.255	1.874	1.777	2.324	2.330
$\overline{R}^2$	0.99	0.95	0.99	0.63	0.99	0.90

Notes: Absolute value of t statistics in parentheses; * significant at 10%; ** significant at 5%; *** significant at 1%.

The average factor shares are =0.33 (capital), =0.53 (materials) and (1-0.33-0.53)=0.14 (labor); with ranges 0.30-0.38 (capital), 0.47-0.56 (materials) and 0.13-0.16 (labor).

TABLE 3. CROSS-SECTION. VALUE ADDED REGRESSION AND THE ACCOUNTING IDENTITY. OLS ESTIMATES

	(1)	**(2)**
Constant	0.578	0.282
	(2.94)***	(1.37)
w	0.257	
	(1.52)	
r	0.736	
	(71.41)***	
L	0.281	0.529
	(44.12)***	(9.15)***
K	0.720	0.440
	(105.31)***	(7.27)***
No. Obser.	48	48
$\overline{R}^2$	0.999	0.96

Notes: Absolute value of t statistics in parentheses; * significant at 10%; ** significant at 5%; *** significant at 1%

The average factor shares are =0.69 (capital) and (1-0.60)=0.31 (labor); with ranges 0.57-0.85 (capital) and 0.14-0.42 (labor).

Figure 1. Value Added. Total Factor Productivity Growth and its Components

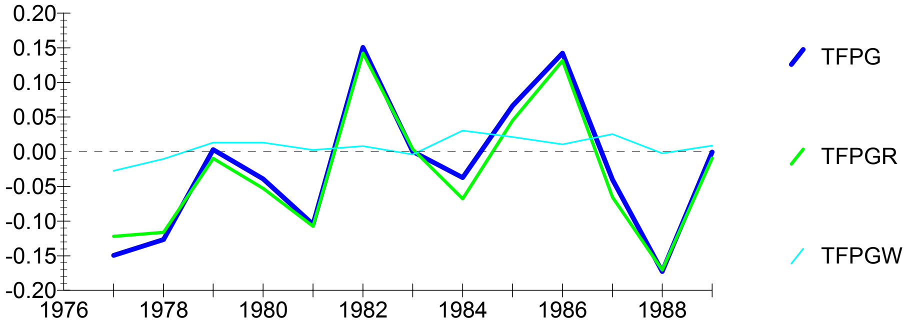


Figure 2. Gross Output. Total Factor Productivity Growth and its Components

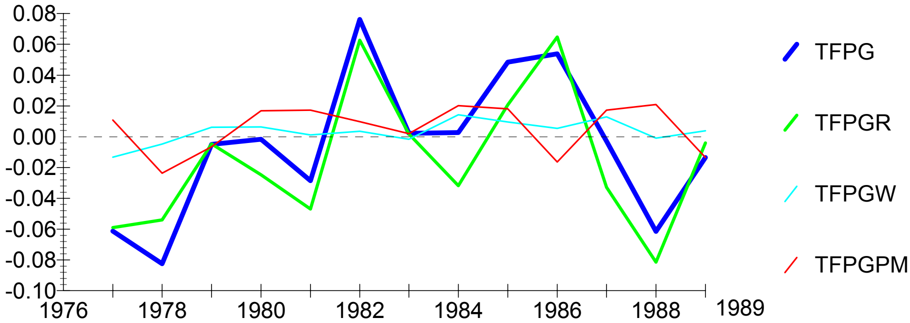
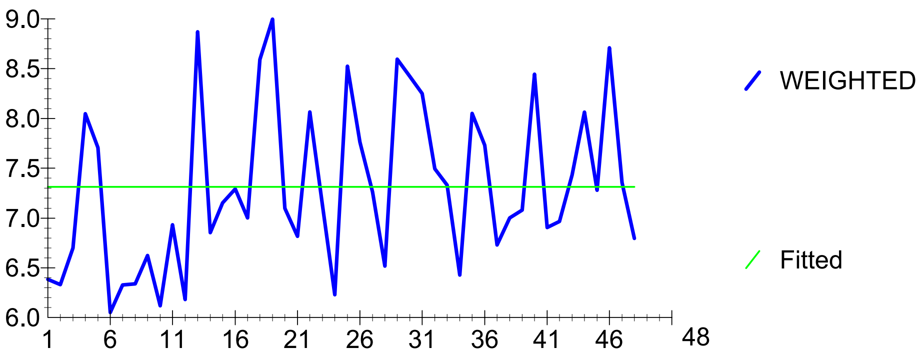


Figure 3. Cross Section. Weighted Average Factor Prices



References:

Barro, Robert. 1999. "Notes on Growth Accounting." Journal of Economic Growth, Vol.64, No.2 (June): 119-137.

Blundell, R., and S. Bond. 2000. "GMM Estimation with Persistent Panel Data: An Application to Production Functions." Econometric Reviews, 19 (3): 321-340.

Bronfenbrenner, M. 1944. "Production Functions: Cobb-Douglas, Interfirm, Intrafirm." Econometrica, 12 (January): 35-42.

________. 1971. *Income Distribution Theory*. Theory. Macmillan: London.

Cohen, A., and G.C. Harcourt. 2003. "Whatever Happened to the Capital Controversies?" Journal of Economic Perspectives, Vol.17, No.1 (Winter): pp.199-214.

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

Douglas, P. 1948. "Are There Laws of Production?" American Economic Review, Vol.38 March): 1-41.

Felipe, Jesus, and F.M. Fisher. 2003. "Aggregation in Production Functions: What Applied Economists Should Know." Metroeconomica, Vol.54, No.2-3 (May-September): 208-262.

Felipe, Jesus, and JSL McCombie. 2003. "Whatever Happened to the Cambridge Capital Theory Controversies? A Comment." Journal of Economic Perspectives, Vol.17, No.4 (Fall): 229-231.

Fisher, Franklin M. 1971. "Aggregate Production Functions and the Explanation of Wages: A Simulation Experiment." The Review of Economics and Statistics, Vol.53, No.4 (November): 305-325.

Griliches, Z., and J. Mairesse. 1998. "Production Functions: The Search for Identification." In Steinar Strøm (ed.), Econometrics and Economic Theory in the 20th Century, pp.169-203. Cambridge; Cambridge University Press.

Intriligator, M.D. 1978. *Econometric Models, techniques and applications*. Prentice Hall: Englewood Cliffs, N.J.

Katayama, H., S. Lu, and J.R. Tybout. 2003. "Why Plant-Level Productivity Studies are often misleading, and an Alternative Approach to Inference." NBER Working Paper 9617.

Levinsohn, J., and A. Petrin. 2000. "Estimating Production Functions Using Inputs to Control for Unobservables." NBER Working Paper 7819.

Levinsohn, J., and A. Petrin. 2003. "Estimating Production Functions Using Inputs to Control for Unobservables." Review of Economic Studies, 70: 317-341.

Lucas, R. 1970. "Capacity, Overtime, and Empirical Production Functions." American Economic Review, Papers and Proceedings, 60 (2), May: 23-27.

Marschak, J., and W.H. Andrews. 1944. "Random Simultaneous Equations and the Theory of Production." Econometrica, 12 (32-4) (July-October): 143-205.

Olley, S., and A. Pakes. 1996. "The Dynamics of Productivity in the Telecommunications Equipment Industry." Econometrica, 64 (6): 1263-1298.

Phelps Brown, E.H. 1957. "The Meaning of the Fitted Cobb-Douglas Function." Quarterly Journal of Economics, 71: 546-560.

Samuelson, Paul. 1979. "Paul Douglas's Measurement of Production Functions and Marginal Productivities." Journal of Political Economy, 87, 5 (October): 923-939.

Simon, H., and F. Levy. 1963. "A Note on the Cobb-Douglas Function." Review of Economic Studies, 30(2), 83: 93-94.

Simon, H. 1979a. "On Parsimonious Explanations of Production Relations." The Scandinavian Journal of Economics, 81: 459-474.

Simon, H. 1979b. "Rational Decision Making in Business Organizations." American Economic Review, 69: 493-513.

Shaikh, Anwar. 1980. "Laws of Production and Laws of Algebra: Humbug II." In Edward J. Nell (ed.), *Growth, Profits and Property. Essays in the Revival of Political Economy,* pp. 80-95, Cambridge: Cambridge University Press.

Wallis, K.F. 1980. *Topics in Applied Econometrics*. Second edition revised. The University of Minnesota Press: Minneapolis.

Wibe, Sören. 1984. "Engineering Production Functions: A Survey." Economica, 51: 401-411.

Wold, H., and R. Faxer. 1957. "On the Specification Error in Regression Analysis." Annals of Mathematical Statistics, 28: 265-267.

Zellner, A., J. Kmenta, and J. Drèze. 1966. "Specification and Estimation of Cobb-Douglas Production Function Models." Econometrica, 34: 784-795.