ISSN 0111-1760

A Problem with Some Estimations and Interpretations of the Mark-up in Manufacturing Industry

J Felipe & JSL McCombie

No. 0117

Jesus Felipe is at the Georgia Institute of Technology and JSL McCombie is Fellow and Director of Studies of Economics, Downing College, Cambridge. We are grateful, as usual, for the comments and encouragement of Geoff Harcourt. We are also grateful for the comments received at seminars given at the University of Otago and the Victoria University of Wellington, New Zealand.
Email: jesus.felipe@inta.gatech.edu *and* jslm2@cam.ac.uk

September 2001

Department of Economics
School of Business
University of Otago
PO Box 56
Dunedin
NEW ZEALAND
Ph: 00 64 3 479 8725
Fax: 00 64 3 479 8174
Email: economics@otago.ac.nz

A Problem with Some Estimations and Interpretations of the Mark-up in Manufacturing Industry

ABSTRACT This paper evaluates the methodological foundations of some recent attempts to estimate econometrically the degree of market power and the degree of returns to scale in manufacturing. The method discussed is based on estimating the aggregate production function in growth rate form. It is argued, following an argument made in another context by Phelps Brown, Shaikh, and Simon, that as the data used in empirical analyses are in value terms (i.e., monetary values at constant prices), the parameter derived as a mark-up can be reinterpreted simply as a coefficient from the income accounting identity, which takes a value of unity subject to omitted variable bias. Thus, it cannot be unambiguously interpreted as a mark-up. It is also shown that the large estimates of the degree of increasing returns to scale are similarly flawed. The argument also has implications for understanding cyclical fluctuations of the Solow residual, which turns out to be largely the result of the procyclical fluctuations of the profit rate. We conclude by questioning whether the aggregate production function can ever be statistically tested, or, in other words, whether it is capable of being refuted, as opposed to its parameters being merely estimated.

1. Introduction

In a series of papers Hall (1986, 1987, 1988a, 1988b, 1990) has proposed an innovative method to estimate whether firms set prices above or equal to marginal costs, and hence whether or not they exhibit market power. The method consists in comparing movements in output and inputs through the aggregate production function. An extension of this approach also estimates the degree of returns to scale. Although several alternative procedures to Hall's method have been proposed subsequently, and it has been re-evaluated, it nevertheless represents the standard departure point for many analyses of market power.

Jesus Felipe, Georgia Institute of Technology, Atlanta, GA 30332-0610, USA. E-mail: jesus.felipe@inta.gatech.edu.and International University of Japan, Niigata 949-7277 Japan. E-mail: jfelipe@iuj.ac.jp J.S.L. McCombie, University of Cambridge, U.K. E-mail: jslm2@cam.ac.uk

The purpose of this paper is to provide an assessment of the methodological foundations of this approach, and to demonstrate an important limitation. In doing this, we elaborate and extend an important critique of the aggregate production function discussed, in another context, by Phelps Brown (1957), Shaikh (1974, 1980, 1987), Simon (1979) and McCombie (1987). Felipe and McCombie (2001, 2002) discuss the general issues involved. Given the importance of this critique, it is surprising that it has been largely ignored in the literature. The problem with Hall's method is that the parameter that is theoretically derived as the mark-up, and estimated as such, cannot be unambiguously interpreted in this manner.

The difficulty arises from the fact that the method used to derive an expression for the mark-up is based upon a transformation of the aggregate production function. This poses a problem because of the existence of an underlying accounting identity which defines the measure of output (whether value added or gross output) in terms of the total compensation of the factors of production. It is shown that the equation used to estimate the mark-up can be derived simply as an algebraic transformation of the accounting identity, which has no behavioural implications. In other words, all that is being estimated is an approximation to the identity and the estimates do not necessarily reflect either the underlying technology of the economy or the state of competition. The fact that often the supposed mark-up takes a value that is greater than unity, especially when value added data are used, is due merely to omitted variables bias and cannot necessarily be taken to indicate the existence of market power. The same argument also explains why the use of gross output leads to values of this parameter of around unity. Hall's model, it has been argued, also has important implications for understanding the causes of business cycles. We also question this interpretation.

The rest of the paper is structured as follows. In Section 2, Hall's approach is outlined and there is a brief summary of the recent literature.[1] In Section 3, we show why this method to estimate the mark-up is subject to an insoluble interpretation problem. As a consequence, we argue that there is no reason for necessarily interpreting the Solow residual as a measure of technical change. Section 4 demonstrates that the reason why the Solow residual varies procyclically is due simply to variations in the rate of profit derived from the identity, most likely as a result of variations in capacity utilization over the cycle.

Section 5 considers why regressions where factor cost, as opposed to revenue, shares are used to weight the growth of inputs cannot provide evidence of the degree of returns to scale. Section 6 considers some possible objections to our critique and also poses the question as to whether or not it is possible to statistically test the aggregate production *per se*, as opposed to merely assuming it exists and estimating its putative parameters. Finally, Section 7 draws some conclusions.

2. A Brief Survey of the Recent Literature

This section provides an overview of the methodology introduced by Hall.[2] Discussion here is limited to key features of the model and a summary of the findings is presented in Table 1. In order to study whether or nor firms equate price to marginal cost, Hall estimated econometrically the degree of market power, namely, the ratio of price to marginal cost, denoted by μ, under the assumption that this mark-up is constant. He compared actual changes in costs to changes in output. Hall's (1988a) estimating equation is essentially derived from an aggregate production function and is:

$$(q_t - k_t) = \varphi + \mu\left(a_t n_t\right) + u_t \tag{1}$$

where $(q_t - k_t)$ is the growth of the output-capital ratio, $a_t n_t$ is the product of the labour share in value added, or the revenue share, a_t, multiplied by the growth of the labour-capital ratio, denoted by n_t (i.e., $n_t = \ell_t - k_t$, where ℓ_t is the growth of employment). φ_t is the rate of technical progress and is modelled as a constant, φ, plus a random element u_t, the latter being the error term of equation (1). Therefore, the relation between price and marginal cost can be found by comparing the actual growth in $(q_t - k_t)$ with the growth that would be expected given the rate of technical progress and $a_t n_t$.

The relationship between equation (1) and the Solow residual is a fundamental aspect of Hall's methodology. Solow's (1957) method for measuring total factor productivity growth is derived from the production function $Q_t = A_t F(L_t, K_t)$ where Q, L, K, and A are value added, the labour input, the stock of capital, and the level of technology. Solow calculated the rate of technical progress (or the rate of growth of total factor productivity) as $\varphi_t \equiv (q_t - k_t) - (a_t n_t)$. From a comparison with equation (1), it can be seen that Solow implicitly assumed that $\mu = 1$. Hall showed that when firms charge prices above marginal costs, the share of profits in revenue exceeds capital's output elasticity because the former includes monopoly profits. A corollary is that labour's share will be less than labour's output elasticity. It therefore follows that the estimate of μ in equation (1) will exceed unity in the presence of market power. The standard calculations of total factor productivity, following Solow (1957), assume constant returns to scale and perfect competition and, consequently, that the output elasticities equal the relevant factor shares in total revenue. Thus, in the presence of market power, the value of the Solow residual will be biased. A finding that the mark-up exceeds unity, therefore, is sufficient to reject the joint hypotheses that firms operate under constant returns to scale and are perfectly competitive.

Hall (1988a, 1990) used US industry data at the one- and two-digit SIC level. He applied instrumental variable (IV) estimation, as in the presence of market power the Solow residual is correlated with $a_t n_t$. The instruments used were variables that affect demand but that should be uncorrelated with technical change. They were the growth rate of the oil price, the growth rate of military expenditures, and a dummy variable for whether the President was a Democrat or a Republican.[3] However, for empirical purposes, Hall did not estimate equation (1). Instead, he provided the inverse estimate of the instrumental variable regression (i.e. $(a_t n_t) = c + b_1(q_t - k_t)$). The reciprocal $\hat{\mu} = 1/\hat{b}_1$ maps all mark-ups greater than unity into the interval from zero to one. The rationale for estimating the inverse regression is that when overhead labour and labour hoarding are high, the growth of labour is only likely to be weakly correlated with the instruments, even though the growth of output is highly correlated (Hall, 1988a p.934). Under these circumstances, the estimated mark-up and its variance are large.[4]

The regressions yielded, in general, relatively high and statistically significant estimates of μ, suggesting either that firms fail to maximise profits, or that they possess substantial market power. Therefore, the results were taken to refute the oft-made assumptions of constant returns to scale and perfectly competitive markets.[5]

Waldman (1991) noted that exceptionally high mark-ups were found by Hall (1988a) in some non-manufacturing industries. Waldman argued that this was caused by the procedures used by the Bureau of Economic Analysis (BEA) to estimate real value added in non-manufacturing industries (the data used by Hall). He argued, in particular, that the deflation method adopted, whether it was double deflation, direct deflation, extrapolation, or some mixture of these methods, was crucial in order to explain Hall's results for the non-manufacturing industries. The essence of the problem lies in the procedures used by the BEA to estimate real value added in non-manufacturing industries. The defects in the deflation method used by the BEA for those industries biased upward Hall's estimates of the mark-ups. In those cases where the BEA used double deflation, the estimates of value added have no immediate bias. But for those industries where the BEA used direct deflation, or extrapolation, real value added was underestimated during years of upward oil price shocks. Waldman, nevertheless, concluded that his critique did not invalidate Hall's overall method of estimating the mark-up, as it was only concerned with the measurement of data.

Domowitz *et al*. (1988) and Norrbin (1993) also adopted Hall's method, but with the modification of introducing intermediate inputs in the analysis.[6] Their findings were different. Although Domowitz *et al*. (1988) rejected the null hypothesis that price equals marginal cost in US manufacturing, their estimates were much lower than Hall's. In the case of Norrbin, his mark-ups were relatively small and insignificantly different from 1. Both argued that Hall's estimates were subject to a bias from the use of value added rather than gross output, which they argued was a preferable measure of output for estimating the mark-up. Norrbin, following Hall, derived a similar expression to provide an estimate of the mark-up, but obtained from the gross output production function $Y_t = A'_t F(L_t\ K_t, M_t)$, where Y_t and M_t denote gross output and intermediate materials and A'_t is the Hicks-neutral rate of technical progress. The equation is:

$$(y_t - k_t) = \varphi' + \mu\left(s_{Lt}n_t + s_{Mt}(m_t - k_t)\right) + u_t \qquad (2)$$

where $(y_t - k_t)$ denotes the growth of the gross output-capital ratio, $(m_t - k_t)$ is the growth of the intermediate materials-capital ratio, $s_{Lt} = w_t L_t / Y_t$ and $s_{Mt} = p_{Mt} M_t / Y_t$ are the shares of labour and intermediate materials in gross output respectively. w, L, Y, M, and p_M are the wage rate, employment, gross output, intermediate material inputs and their price respectively. φ' is the constant rate of technical progress. When Norrbin reproduced Hall's tests (slightly modified) but introducing intermediate inputs in the analysis, he found: (i) that no significant correlations existed between the instruments and the Solow residual, i.e., the latter is orthogonal to the instruments selected by Hall; and (ii) in contrast to Hall's estimates of large mark-ups, the mark-ups were relatively small and insignificant (i.e., $\hat{\mu}$ was approximately equal to 1).

<Table 1 here>

3. The Accounting Identity and the Problem of the Interpretation of the Mark-up

The purpose of this section is to provide an alternative explanation for the results obtained using the method introduced by Hall. It will be shown that equations identical to equations (1) and (2) above can be obtained as simple algebraic transformations of the income accounting identity, namely, value added (gross output) equals the sum of wages plus profits (plus intermediate inputs). In fact, we shall show that the parameter of the equations estimated, namely μ, can simply be reinterpreted as a biased estimate -- due to misspecification -- of a coefficient of the accounting identity. The true, or unbiased, value of μ must, by definition, be equal to unity regardless of the state of competition.

Naturally, the econometric estimation of this equation, an identity, cannot provide any *independent* evidence that the coefficient estimated is the value of the mark-up parameter. It follows that, under the interpretation presented here, although the use of gross output, instead of value added, leads to different estimates of μ, this is not the central problem. In fact, a consideration of the identity suggests on *a priori* grounds that the degree of bias is likely to be smaller when gross output, rather than value added, is used and this is confirmed empirically. Likewise, data problems (such as those introduced by the deflation method chosen) may also lead to biases, but these are of secondary importance. Bresnahan (1989) has also argued that Hall's model has problems because average incremental cost will not provide and accurate approximation to marginal cost if the latter is not constant, and intertemporal aspects of firm behaviour and differences between variable and quasi-fixed factors are not considered systematically. These are valid criticisms, but again they are of second-order importance.

3.1 The Use of Value-Added Data

The income identity for value added in real terms is defined as:

$$Q_t \equiv w_t L_t + r_t K_t \quad (3)$$

where *Q, w, r, L,* and *K* denote output (value added at constant prices), the average wage rate, the average profit rate, the level of employment, and the stock of capital, respectively. Equation (3) is an accounting identity that holds for every period of time at the level of the firm, sector, and the total economy, and is compatible with *any* degree of competition and *any* degree of returns to scale. As it is an identity, there are no behavioural assumptions necessarily underlying it (e.g. that economic profits are zero). It simply expresses how value added is distributed between labour and capital, without any causal implications. As Samuelson (1979, p. 932) put it: "No one can stop us from labeling this last vector [residually computed profit returns to "property" or to the nonlabor factor] as *rK* as J.B. Clark's model would permit – even though we have no warrant for believing that noncompetitive industries have a common profit rate *r* and use leets capital *K* in proportion to the *Q* - *wL* elements!"[7,8]

Expressing equation (3) in growth rates, the following is obtained:

$$q_t \equiv \varphi_t + a_t \ell_t + (1 - a_t) k_t \quad (4)$$

where:

$$\varphi_t \equiv a_t \varphi_{wt} + (1 - a_t) \varphi_{rt} \quad (5)$$

The variables φ_{wt} and φ_{rt} are the growth rates of the wage and profit rates and a_t and $(1-a_t)$ represent the labour and capital (revenue) shares in value. These three equations form the basis of our argument. It must be stressed again that all we have done is to transform an accounting identity, and that no assumptions have been made about the structure of the economy or industries to which the identity pertains. Equation (4) can be rewritten as:

$$q_t - k_t \equiv \varphi_t + (a_t n_t) \qquad (6a)$$

If equation (6a) is estimated as:

$$q_t - k_t \equiv \tau \varphi_t + \mu^* (a_t n_t) \qquad (6b)$$

it is evident that τ and μ^* must both equal unity. Because this derivation follows from an identity, the parameters τ and μ^* do not have, *per se*, any necessary interpretation as structural or economic parameters.[9] The question that arises, in the light of these arguments, is why the estimate of μ in the studies surveyed above generally differs from unity.

This is because equation (6b), or a variant of it, is estimated, but with φ_t replaced by a constant. The equation estimated is thus formally equivalent to equation (1) above. The estimate of μ in equation (1) could simply be interpreted as a biased estimate of $\mu^* = 1$ in equation (6b). The bias in equation (1) is due to the misspecification of the term φ_t and is given by:

$$E(\hat{\mu}) = \mu^* + \tau \frac{cov(\varphi_t, a_t n_t)}{var(a_t n_t)} = 1 + \frac{cov(\varphi_t, a_t n_t)}{var(a_t n_t)} \qquad (7)$$

as $\mu^* = \tau = 1$. Thus, the magnitude of the parameter μ in equation (7) will be determined by the ratio of the covariance of the omitted and included variables to the variance of the latter, i.e., one plus this bias. The latter can of course be negative and larger than one, thus leading to a parameter below 1. But, in terms of our argument, this has no bearing on the issue of market structure and returns to scale. Note that μ will be *1* if, and only if, $cov(\varphi_t, a_t n_t) = 0$. This condition, in the context of our derivation (from an identity), does not necessarily imply that markets are competitive. And likewise, $\mu > 1$ will be the result of $cov(\varphi_t, a_t n_t) > 0$. This, again, cannot be interpreted as evidence of market power. These arguments explain the results obtained by Hall and others in their mark-up regressions. If these regressions had included the correct specification of φ_t (rather than substituting φ_t by $\varphi + u_t$), they would have found μ equals unity. This is also made readily apparent by considering Hall's (1988a, p.926) definition of the mark-up. This, when it varies over time and there is no error, is defined by Hall as $\mu_t = \frac{(q_t - k_t) - \varphi_t}{a_t n_t}$.

But we know from a consideration of the identity that $(q_t - k_t) - \varphi_t \equiv a_t n_t$ and, hence, substituting this into Hall's equation, once again we arrive at the result that $\mu = 1$. The only reason why Hall does not find that this is the case econometrically is that, as we have seen above, he assumes that φ_t is a constant plus a random term, whereas empirically φ_t derived from the identity varies over time. Moreover, φ_t could be calculated directly from the data using equation (5) and hence included explicitly in the regression.

Although we find that the estimate of the mark-up is equal to unity plus the bias due to the omission of φ_t from the identity (or rather due to the proxying of it by a constant), it should be emphasised that the central tenet of our critique does not rest on this empirical finding. (The reasons for the bias are discussed in Section 4 below.) For example, suppose that the wage and the profit rate are constant over time in our data set. (In practice, this may be true of a number of developing countries.) In these circumstances, the identity becomes simply $q_t \equiv a_t \ell_t + (1-a_t)k_t$ or $(q_t - k_t) \equiv a_t n_t$.

Consequently, estimating the equation $(q_t - k_t) = c + \mu(a_t n_t)$ would, because of the identity, give an estimate of μ of unity, as there is no bias. (We would also get the same result if $a_t\varphi_{wt} + (1-a_t)\varphi_{rt}$ were orthogonal to $a_t n_t$). The orthodox approach would interpret this as implying that markets were competitive. There is, of course, nothing to production theory to suggest that the fact that wages and the profit rate are constant over time will ensure perfect competition.

In practice, the weighted average growth of the wage and profit rates show a procyclical variation around its mean. (This is confirmed econometrically in Section 4.) This will be the error term in the regression and leads to a bias in the estimates of the putative "output elasticities" (which may consequently sum to greater or less than unity and may differ from the factor shares). But this, likewise, cannot be assumed to be an independent test of the degree of returns to scale and the marginal productivity theory of factor pricing. We can always improve the goodness of fit by including a flexible non-linear time trend as, after all, there is nothing in neoclassical production theory that requires technical progress to occur at a constant rate. Alternatively, this could be accomplished by including a variable that is closely correlated with the deviations around the mean of φ_t. We would thus more closely approximate the accounting identity and so the estimates of the output elasticities would tend to the values of the observed factor shares (see Section 6). The problem is that by using constant-price value data, all the regression is doing is in effect tracking the accounting data.

It must also be pointed out that in the context of an identity the use of an "inverse" estimate and instrumental variable (IV) estimation must be seen as irrelevant, since there is no endogeneity problem, and it may even be a source of error. It can be shown that the IV estimator of the direct regression can be written as

$\mu^{IV} = 1 + \frac{cov(\varphi_t, z_t)}{cov(a_t n_t, z_t)}$ and that of the inverse regression as $\mu^{IV} = 1 - \frac{cov(\varphi_t, z_t)}{cov[(q_t - k_t), z_t]}$,

where z_t are the instruments. If the instruments are not highly correlated with $(a_t n_t)$, the estimates will diverge substantially from those provided by OLS.

Table 2 shows the estimates of equation (1) for the manufacturing sector of the US over the period 1958-91. Using OLS, the estimate of the mark-up is large, statistically significant, and takes a value of 3.573. From the data we also find that $cov(\varphi_t, a_t n_t) = 5.353\times10^{-5}$ and $var(a_t n_t) = 2.077\times10^{-5}$. The ratio of these two expressions is an exact measure of the bias of the OLS regression given by equation (7) and is 2.57. (The estimate reported in Table 2 is 1 plus this bias.) Table 2 also shows the IV estimates of equation (1) as well as the inverse IV results, as specified by Hall. The same argument follows through in these cases.

<Table 2 here>

3.2 The Use of Gross Output

What is the effect of using gross output instead of value added? Not surprisingly, the argument outlined above still holds. As discussed in Section 2, several authors have argued that the correct specification of the production function to estimate the mark-up requires the use of gross output as the output measure and the inclusion of intermediate inputs as a regressor and that this significantly affects the estimate of the mark-up. Nevertheless, the relationship between the gross output production function and the underlying income identity in terms of total sales follows equally. The difference is that the identity is now:

$$Y_t \equiv w_t L_t + p_{M\,t} M_t + v_t K_t \qquad (8)$$

where Y_t and M_t denote gross output and the constant price value of intermediate inputs respectively (using double deflation); v_t is the profit rate for gross output; and p_{Mt} is the relative price of materials. In growth rates:

$$y_t \equiv \varphi_t' + s_{Lt}\ell_t + s_{Mt}m_t + (1 - s_{Lt} - s_{Mt})k_t \qquad (9)$$

where, it will be recalled, s_{Lt} and s_{Mt} are the shares of labour and intermediate materials in gross output. Moreover,

$$\varphi_t' \equiv s_{Lt}\varphi_{wt} + s_{Mt}\varphi_{p_M t} + (1 - s_{Lt} - s_{Mt})\varphi_{vt} \qquad (10)$$

The variables φ_{wt}, $\varphi_{p_M t}$ and φ_{vt} are the growth rates of wages, the relative price of intermediate materials, and profit rate in gross output. Equation (9) can be rewritten as:

$$(y_t - k_t) = \varphi'_t + s_{Lt} n_t + s_{Mt} (m_t - k_t) \qquad (11)$$

It is readily apparent that if equation (11) is estimated econometrically with τ and μ^* as the coefficients of φ_t' and $(s_{Lt}\, n_t + s_{Mt}\, (m_t - k_t))$, respectively, τ and μ^* must be equal to 1. As in the case of value added data, suppose equation (11) is estimated, but with the term φ'_t replaced by a constant. This would be equation (2), and the estimate of μ can be once again reinterpreted as a biased estimate of $\mu^* = 1$ in equation (11). The bias is given now by:

$$E(\hat{\mu}) = \mu^* + \tau \frac{cov(\varphi'_t,\ s_{Lt} n_t + s_{Mt}(m_t - k_t)\ )}{var(s_{Lt} n_t + s_{Mt}(m_t - k_t))} = 1 + \frac{cov(\varphi'_t,\ s_{Lt} n_t + s_{Mt}(m_t - k_t))}{var(s_{Lt} n_t + s_{Mt}(m_t - k_t))} \qquad (12)$$

The results will be, in general, different from those obtained using value added, but the argument remains the same. Norrbin's (1993) finding of small mark-ups (some below unity) should be simply interpreted as a low (or negative) value of $cov[\varphi'_t, s_{Lt}\, n_t + s_{Mt}(m_t - k_t)] / var(s_{Lt}\, n_t + s_{Mt}(m_t - k_t))$. Table 2 reports the estimates corresponding to equation (2). In our data set, $cov[\varphi'_t, s_{Lt}\, n_t + s_{Mt}(m_t - k_t)] = 4.734 \times 10^{-5}$ and $var(s_{Lt}\, n_t + s_{Mt}(m_t - k_t)) = 1.549 \times 10^{-4}$ for the manufacturing sector. The ratio is again exactly equal to the bias of the OLS regression coefficient (see equation (12)), namely 0.42.

As noted in Section 2, in general, the use of gross output, as opposed to value added, will produce a lower estimate of μ (compare left and right hand sides of Table 2). This can also be explained in terms of the accounting identity and the biases given by equations (7) and (12). From equation (11), the closer the weighted average of the growth of the wage and profit rates is to a constant, the lower the bias. Generally speaking, the growth of materials shows a much higher procyclical variation than either the growth of labour or capital, and it is highly correlated with the growth of output. Similarly, the revenue shares show little cyclical variation, so they may be treated as a constant. The weighted average of the growth rates of the factor prices when the price of intermediate goods is included shows a procyclical variation around its mean . However, this is smaller than that of the weighted growth of the wage and profit rates. Thus, the inclusion of the weighted growth of materials in the sum of the weighted growth of inputs causes a greater cyclical variation compared with the case when value added data is used (compare Figures 1 and 2).[10] Consequently, the degree of fluctuation in $s_{Lt}n_t + s_{Mt}(m_t - k_t)$ exceeds that of $a_t n_t$.

This, combined with the stability of the factor shares, means that the denominator of the bias in equation (12) will be much larger than that of equation (7), i.e., ($var(s_{Lt}n_t + s_{Mt}(m_t - k_t)) > var(a_t n_t)$). The covariances in the numerators in both expressions for the degree of bias are of a similar order of magnitude. Therefore, the degree of omitted-variable bias will be considerably reduced with the use of gross output, and the goodness of fit will improve simply because we are more closely approximating the accounting identity.[11]

4. Profit Rates, Cyclical Fluctuations, and the Mark-up

There has been a long debate as to the causes of cyclical fluctuations in the growth of labour productivity and total factor productivity. The standard argument for a long time was that such procyclicality was the result of labour hoarding (Oi, 1962). Productivity declines in a temporary slump because under-utilised workers are kept on by firms in anticipation that once the recession is over they will be productive again. It is costly to lay off and then re-hire workers. This view was challenged by the work of the real business cycle school, according to which economic fluctuations are driven by exogenous technological shocks (Prescott, 1986). More recently, Hall (1988a, 1990) has argued that perfect competition rules out procyclicality, and, therefore, the observed procyclicality of the Solow residual is the result of imperfect competition. However, by examining the sources of bias in the accounting identity, we can show that the observed procyclicality of the Solow residual is mostly due to the procyclical fluctuations of the observed rate of profit. It will be recalled that we are skeptical as to the interpretation of the Solow residual as “technical change”.

The bias in equation (6a) can be estimated through the auxiliary regression: $\varphi_t = c + \phi\,(a_t n_t)$, where the estimated parameter ϕ equals $cov(\varphi_t, a_t n_t)/var\,(a_t n_t)$. As from equation (5) we know the two components that make up φ_t, we can estimate the two auxiliary equations:

$$a_t \varphi_{wt} = c + \phi_1 (a_t n_t) \tag{13a}$$

$$(1 - a_t)\,\varphi_{rt} = c + \phi_2\,(a_t n_t) \tag{13b}$$

The results are reported in the upper half of Table 3, where it can be seen that most of the bias is due to equation (13b), i.e. it is caused by the cyclical fluctuation in the weighted growth of the rate of profit.

A consideration of the graphs of the growth of output and labour over the period shows that they both exhibit cyclical fluctuations with troughs and peaks in roughly the same years, although the fluctuations are more severe in the case of output. The growth of capital shows virtually no cyclical fluctuations.

[Table 3 about here]

The growth of the wage rate is mildly procyclical (wages are sticky) whereas that of the rate of profit is markedly so. The most likely explanation of this observation is that, because of excess capacity in the downturn, profits fall in a recession (for a recent assessment of the link between profits and cycles see Zarnowitz, 1999). Thus, the procyclical movement of the rate of profit is not a surprising result. The shares of labour and capital show a slight secular trend over the period but no noticeable cyclical fluctuations. Consequently, φ_{rt} is highly correlated with the growth of output, q, and the growth of labour ℓ, and this explains the results of the second auxiliary regression, and the substantial degree of omitted variable bias leading to the implausible results of the mark-up.

Two further points are worth noting. First, if we were to adjust the capital stock and measure the growth of the capital services by the rate of change of the flow of capital services, this would show a pronounced cyclical fluctuation and the implied rate of profit would exhibit very little variation, as capital's share is roughly constant. This would reduce the degree of bias of the estimate of μ. Secondly, as we noted above, if there were no growth in the wage rate or in the rate of profit, the estimate of μ would by definition be unity, although this would not imply that the economy had suddenly become more competitive.

There is a similar argument concerning the use of gross output. While φ'_t shows a procylical fluctuation, the degree is less than in the case of value added. Consequently, the bias in proxying φ'_t by a constant is less than in the case of value added. The bias in estimating equation (11) when φ'_t is replaced by a constant term can be estimated using the auxiliary regression: $\varphi'_t = c + \delta(s_{tL}n_t + s_{Mt}(m_t - k_t))$, where the estimated δ equals $cov(\varphi'_t, s_{Lt}n_t + s_{Mt}(m_t - k_t))/var(s_{Lt}n_t + s_{Mt}(m_t - k_t))$. Since from equation (10) we know the three components of φ'_t, we can estimate the three auxiliary equations:

$$s_{Lt}\varphi_{wt} = c + \delta_1 \left(s_{Lt}n_t + s_{Mt}(m_t - k_t)\right) \quad (14a)$$

$$\left(1 - s_{Lt} - s_{mt}\right)\varphi_{vt} = c + \delta_2\left(s_{Lt}n_t + s_{Mt}(m_t - k_t)\right) \quad (14b)$$

$$s_{Mt}\varphi_{p_M t} = c + \delta_3\left(s_{Lt}n_t + s_{Mt}(m_t - k_t)\right) \quad (14c)$$

The regression results shown in the bottom half of Table 3 confirm the above arguments. In particular, it can be seen that the growth of the relative price of materials shows no significant correlation with the growth of the weighted factor inputs. The important conclusion of the previous analysis is that the procyclicality of the Solow residual, and, therefore, economic fluctuations, is a result of the procyclicality of the profit rate derived from equation (3), acting through the accounting identity. This is not to say, though, that there is an implied direction of causality here. The association is most likely due to the well-established fact of variations of capacity utilisation, together with a constant profit share, and does not necessarily reflect variations in technical change.

Surprisingly, none of the studies that have estimated the mark-up specified the regression in the unconstrained form, *viz.*, for value added:

$$(q_t - k_t) = \varphi + \mu_1 (a_t \ell_t) - \mu_2 (a_t k_t) \qquad (15a)$$

and for gross output:

$$(y_t - k_t) = \varphi + \mu_1 (s_{Lt} \ell_t) - \mu_2 (s_{Lt} k_t) + \mu_3 (s_{Mt} m_t) - \mu_4 (s_{Mt} k_t) \qquad (15b)$$

The results are reported in Table 4 for value added and gross output. We cannot reject the null hypothesis that $\mu_1 = \mu_2$ for value added and $\mu_1 = \mu_2 = \mu_3 = \mu_4$ for gross output. However, for gross output (which as we saw in Section 1 has been argued to be the preferable specification) the coefficient on the weighted growth of capital ($-\mu_2$) takes the wrong sign, regardless of whether OLS or IV estimation is used. It is only the large associated standard errors the prevents the null hypotheses from being rejected. This casts serious doubts on the specification of the model, even if we accept all of Hall's assumptions.

<Table 4 here>

But the problems do not end here. The use of value added data give the result that the supposed output elasticity of capital is negative and the use of gross output that it is implausibly small.

Considering first the case of value added, it is possible to calculate two alternative estimates of labour's output elasticity. Under the usual neoclassical assumptions, labour's output elasticity is defined as $\alpha_t = \mu(wL/pQ) = (p/x)(wL/pQ) = (wL/xQ)$ where α, p and x are labour's output elasticity, the price of output, and the marginal cost. From equation (15a), labour's output elasticity may be defined as either $\alpha_t = \mu_1 a_t$ or $\alpha_t = \mu_2 a_t$. From these we may derive two alternative estimates of capital's avaerage output elasticity, namely, $(1-\alpha) = 1 - \mu_1 \bar{a}$ and $(1-\alpha) = 1 - \mu_2 \bar{a}$, where $\bar{a}$ is the average share of labour over the period (which does not show very much annual variation). The output elasticity of capital is defined as one minus labour's output elasticity as Hall assumes constant returns to scale in the theoretical derivation of the estimating equation.

From the estimates of μ reported in Table 4 and the calculated average value of labour's share, the values of labour's output elasticity given by $\mu_1\bar{a}$ are 1.527 (OLS) and 1.949 (IV) respectively. These results give values of the output elasticity of capital of –0.527 and –0.949. Labour's output elasticities given by $\mu_2\bar{a}$ are 2.058 (OLS) and 2.985 (IV), implying that the alternative estimates of the output elasticities of capital are also negative, namely, -1.058 and
-1.985. In other words, a faster growth of the capital stock, if these results are to be believed, actually reduces the growth of output.

The implied values of the output elasticity for capital from the restricted specification reported in Table 2 are equally implausible and are –0.518 and –0.076. For the inverse regression, the value added output elasticity of labour is not significantly different from zero and hence little credence can be placed on the value of capital's output elasticity calculated from the results of this specification.

In the case of gross output, the output elasticity of labour is given by $\mu_1\bar{s}_L$ and $\mu_2\bar{s}_L$ and the output elasticity of materials is given by $\mu_3\bar{s}_M$ and $\mu_4\bar{s}_M$. In the unrestricted regression, as we have noted, the coefficient of s_Lk takes the wrong sign, although it is statistically insignificant, and hence there is little point in calculating capital's output elasticity. In the restricted regression (Table 2), the estimates of capital's output elasticity are again highly implausible, varying from 0.063 (OLS) to 0.004 (IV).

Hall's (1988a) results also produce implied negative values for capital's output elasticity. If we take labour's share of 0.5 as the value which produces the lower limit for the estimate of capital's output elasticity, only one out of his estimates for the 7 one-digit SIC industries (services) gives a positive value for capital's output elasticity, and even here it is an implausible 0.07.[12] (See Hall, 1988a, Table 4, p.940.)

Consequently, even if one accepts all the usual neoclassical assumptions, this alone should raise questions as to whether the correct specification of the production function is being estimated and whether the results are reliable. This problem, as far as we are aware, has been ignored in the literature as all the studies, including Hall's, estimate the constrained form for both value added and gross output, which conceals this implausible result.

5. The Solow Residual, Increasing Returns to Scale, and Revenue and Cost Shares

Hall (1988a, p.922) states that “the test developed in this paper rests on the assumption of constant returns to scale. That is, the hypothesis being tested is the joint hypothesis of competition and constant returns to scale”. Ignoring for the moment the implausible estimates obtained for the output elasticity of capital, a significant mark-up is found which rejects the null hypothesis of perfectly competitive markets. Consequently, it is not clear why Hall estimates the value of the mark-up by a method that *assumes* constant returns to scale. There seems to be an internal contradiction in this procedure.

It is noticeable that Hall’s (1988b, 1990) later work extends his analysis to allow for the possibility of increasing returns to scale. The degree of returns to scale is given by Hall as $\gamma = (\partial F/\partial L)(L/Q) + (\partial F/\partial K)(K/Q)$. As $\mu\bar{a}$ is taken to be labour’s output elasticity when value added data are used, it follows that $\beta = \gamma - \mu\bar{a}$, where β is the output elasticity of capital. It is a straightforward matter to show that, using revenue shares, the specification allowing for increasing returns to scale becomes:

$$(q_t - k_t) = \varphi + \mu a_t(\ell_t - k_t) + (\gamma - 1)k_t \qquad (16)$$

Equation (16) is identical to Hall’s (1990) equation (5.26).

This equation was estimated using our data set. Because of the significant autocorrelation, the exact AR(1) ML method was used. The results for value added were:

$$\underset{}{(q_t - k_t)} = \underset{(5.07)}{0.079} + \underset{(13.13)}{3.529}(\ell_t - k_t) \underset{(-3.39)}{- 1.675}k_t \qquad \bar{R}^2 = 0.837 \qquad DW = 1.99$$

This gives an implausible value of the degree of returns to scale of –0.675 (with a t-value of –1.36). Using the IV approach (with an AR(1) specification) does not improve the results:

$$(q_t - k_t) = 0.120 + 4.102(\ell_t - k_t) - 2.970\, k_t$$
$$(1.92) \quad (3.04) \qquad\qquad (-1.67)$$

The estimate of γ is -1.970 (and the t-value equals -1.11). The $\overline{R}^2$ and the DW statistic are not reported as they do not have their usual interpretation when IV estimation is used.

Hall (1988b, 1990) uses cost, rather than revenue factor shares in his regression analysis. The advantage of this is that no assumption about the state of competition is needed when the cost shares are used. The cost share of labour is calculated as $a_c = wL/(wL + r_c K)$, where r_c is the shadow or rental price of capital, calculated under a number of what can be best described as heroic assumptions, but which will not be considered here.

Hall's methodology using cost shares is similar to that using revenue shares, but with some important differences. First, when cost shares are used and there are constant returns to scale, there should be no correlation between the residual and exogenously determined movements in output and input growth. On the other hand, with increasing returns to scale, the Solow residual will be positively correlated when output growth increases, even though there has been no shift in the production function. Hall's estimating equation now becomes:

$$q_t = \varphi + \gamma(a_{ct}\ell_t + (1 - a_{ct})k_t) \qquad (17)$$

The focus of interest is on the degree of returns to scale as a direct estimate of the mark-up may be calculated as the ratio of labour's revenue to cost share, $a_r/a_c = \mu = p/x$. Hall (1988a, Table 1 and 1988b, Table 1) reports data for the nondurable goods industry. Using these data, we calculated that the mark-up for this industry is on average 1.10, which is significantly less than the value of 2.06 obtained from the regression analysis using revenue shares as weights. Hall (1988b, p.4) now concludes: “As a practical matter, it makes almost no difference whether cost or revenue shares appear in the productivity measure, because pure profit is sufficiently small that cost and revenue are the same”. This is at somewhat at variance with the results of his previous approach, namely the estimates of the mark-up obtained from the regression analysis and discussed above, to the extent that the latter implied substantial market power. However, Hall (1988b, 1990) finds significant estimates of increasing returns to scale at both the one-digit and the two-digit SIC level.

We confirmed this for the nondurable goods industries, using Hall’s data. When the inverse IV regression $(a_{ct}\,\ell_t + (1\text{-}a_{ct})k_t) = -\varphi/\gamma + (1/\gamma)q_t$ is estimated, it is found that $\hat{\gamma}$ is 3.731 (with a t-value of 2.10), which is close to Hall’s estimate of 3.107.[13] Regressing the growth of output directly on the weighted growth of the factor inputs using the IV method gives a smaller, but still substantial, value of increasing returns to scale of 2.658 (t-value: 2.74).

However, the reason for these results is similar to the one we have discussed above, namely that all that is being estimated is a misspecified identity. As the cost and revenue shares are very close in value, the accounting identity is given by approximately:

$$q_t \cong a_c\varphi_{wt} + (1\text{-}a_c)\varphi_{rt} + a_c\ell + (1\text{-}a_c)k \qquad \text{(18a)}$$

or, alternatively,

$$q_t \cong \varphi_t + \gamma(a_c\ell + (1\text{-}a_c)k) \qquad \text{(18b)}$$

where $\gamma = 1.0$. The fact that empirically γ exceeds unity is because once again φ_t is proxied erroneously by a constant. However, the argument is a little more complex than this. In the neoclassical analysis, even though in the presence of market power the appropriate values of the output elasticities are the cost shares, the measure of the "volume" of output is still constant-price value added. This is measured as $VA \equiv wL + r_cK + \Pi$, where Π is total monopoly profits. This last term may be written as r_mK where r_m is the monopoly component of the rate of return derived from the accounting identity, r, i.e., $rK = r_cK + r_mK$. Thus, it could be legitimately argued that the monopoly profits should be excluded from the definition of the "volume" of output (i.e. value added should be calculated using marginal costs rather than market prices) so that the residual does not include the rate of change in monopoly profits. The latter, of course, has nothing to do with the rate of technical change. Hence, if this procedure is followed and the adjusted value added is given by $VA^* \equiv wL + r_cK$, the arguments above concerning the identity follow through exactly.

Finally, even if we were to assume an underlying aggregate production function together with the standard neoclassical assumptions, Hall's specification of equation (17) conceals the evidence of a serious misspecification error, similar to that found above with the use of revenue shares. When equation (17) is estimated by IV in the unrestricted form of $q_t = \varphi + \gamma(a_{ct}\ell_t) + \gamma((1\text{-}a_{ct})k_t)$ the following results are obtained, which include a negative output elasticity for capital:

$$q_t = \underset{(2.61)}{0.083} + \underset{(2.58)}{1.950}(a_{ct}\ell_t) - \underset{(-1.82)}{8.389}((1\text{-}a_{ct})k_t)$$

Using OLS gives comparable results:

$$q_t = \underset{(2.61)}{0.051} + \underset{(8.26)}{1.410}(a_{ct}\ell_t) - \underset{(-2.64)}{3.203}((1\text{-}a_{ct})k_t) \quad \overline{R}^2 = 0.726 \quad DW = 1.531$$

Thus, even granted the usual neoclassical assumptions, no reliance can be placed on Hall's results as correctly measuring the degree of returns to scale or that a correctly specified production function is being estimated.

6. Testing the Aggregate Production Function

It is useful to deal briefly with possible misinterpretations of our critique and to consider whether one can actually statistically test for the existence of a well-defined aggregate production function. The latter is certainly the view of Solow (1974, p. 121) who considers that: "when someone claims that aggregate production functions work, he means (a) that they give a good fit to input-output data without the intervention of factor shares and (b) that the function so fitted has partial derivatives that closely mimic observed factor shares".

It could be argued that the second part of expression (7), namely the degree to which μ is biased away from unity, has, in effect, an economic interpretation in terms of reflecting the production technology. This is because equation (5), $\varphi_t \equiv a_t \varphi_{wt} + (1 - a_t)\varphi_{rt}$, under the usual neoclassical assumptions can be interpreted as the dual to the Solow residual and equals the rate of technical progress. However, this reasoning overlooks the fact that for what we referred to as the bias in equation (7) to have a technological interpretation requires the existence of an aggregate production function together with the conditions for producer equilibrium (as is implicit in the studies discussed in Section 2 above). Only in this case could one argue that the Solow residual measures technical progress, and that the growth of each factor multiplied by its factor share measures the contribution of the input to overall output growth in a causal sense. (In the case of increasing returns to scale the factor share would be multiplied by γ.) However, there may well not be any well-defined or well-behaved aggregate production function.

The aggregate production function has been subject to a number of serious well known, although widely ignored, criticisms both in terms of aggregation problems (Fisher 1993) and the Cambridge Capital Theory Controversies (Harcourt 1972). The standard defence of the aggregate production function has been along instrumentalist lines in that these criticisms are irrelevant since the production function "works", in the sense that it gives good statistical predictions and a close statistical fit to the input-output data. But this raises a further empirical issue. The arguments and results in Sections 3 and 4 have been derived exclusively from an accounting identity, without any reference to the state of competition, the aggregate production function, and optimization conditions. This is not a trivial issue as it makes the production function a non-refutable hypothesis. The better the goodness-of-fit of the putative production function, the closer the estimates of the elasticities will be to the values of the factor shares, without necessarily implying competitive markets or indeed the existence of a well-behaved production function.

Furthermore, our critique does *not* imply that the neoclassical approach ignores the presence of the accounting identity (see, for example, Jorgenson and Griliches, 1967). However, it does raise important methodological issues relating to the testing and estimation of production functions. The orthodox approach assumes that there is a technological relationship between *Q*, *K,* and *L* and that one can estimate such technological parameters as the elasticity of substitution, the degree of returns to scale, and the output elasticities. It also shows that if perfect competition and the marginal productivity theory of factor pricing prevail, then factor shares will equal the output elasticities. Thus, from the production function together with the assumptions of profit maximisation and competitive markets, we can derive the expression $q_t = \varphi_t + a_t\ell_t + (1-a_t)k_t$. But, from the identity, we have seen that $q_t = a_t\varphi_{wt} + (1-a_t)\varphi_{rt} + a_t\ell_t + (1-a_t)k_t$.

The orthodox approach to estimating production functions specifies some particular form of the production function and usually (but not always) specifies a linear time trend to capture technical change. To the extent that the actual specification of the production function tracks over time the changes in factor shares and in the weighted average of the growth rates of the wage and profit rates, then we will find, by virtue of the identity, that the estimated output elasticities do, in fact, equal the factor shares. In practice, the shares are often roughly constant and a Cobb-Douglas gives a good fit, but this is not essential to the argument.

Consequently, if we were to estimate the identity as $q_t = b_2(a_t\varphi_{wt}) + b_3((1-a_t)\varphi_{rt}) + b_4(a_t\ell_t) + b_5((1-a_t)k_t)$ or, equivalently, as $q_t - k_t = b_6(a_t\varphi_{wt} + (1-a_t)\varphi_{rt}) + b_7(a_t(\ell_t - k_t))$ all the estimated coefficients would be unity. If in the last equation we replace $(a_t\varphi_{wt} + (1-a_t)\varphi_{rt})$ by a constant, then the estimate of b_7 will biased to the extent that the weighted growth of wages and the rate of profit are not orthogonal to $(\ell_t - k_t)$. It turns out, as we have seen, that the two variables are positively correlated – both vary procyclically – so that the biased estimate of b_7 exceeds unity. But since we are merely dealing with an identity, this does not imply anything about the structure of production, the mark-up or the degree of returns to scale.

What happens if the supposed aggregate production function is freely estimated econometrically (i.e. without making use of the factor shares)? After all, many studies find that the estimated elasticities do not equal the factor shares. In this case, a particular parameterization of the production function, such as the Cobb-Douglas relationship, would have been estimated. But we can also compare this functional form with the accounting identity and infer *a priori* the results that would be obtained. If factor shares in this economy are relatively constant ($a_t \approx a$), and the weighted growth of the wage and the profit rate grow at a constant rate, then the estimated output elasticities of labour and capital (*viz.* α and β) would be close to the relevant shares in the national income accounts ($\alpha = a$ and $\beta = 1-a$). But would that represent a failure to refute the assumption of competitive markets? The answer is that it would not. Factor shares can be constant for a variety of reasons that have nothing to do with a Cobb-Douglas production function (a constant mark-up pricing policy, for example, or the Kaldorian macroeconomic theory of distribution (Kaldor, 1956)), and thus all that would be estimated is the accounting identity.

However, the estimated parameters of the Cobb-Douglas function are often found to differ greatly from the observed factor shares. In fact, it is not unknown for the estimate of the output elasticity of capital to be negative (see Lucas, 1970 and Tatom, 1980). Empirically, we saw that this was true of the estimation of Hall's specification, but this also can be shown to occur when the production function is freely estimated. First, consider the full identity, $q_t = a_t\varphi_{wt} + (1-a_t)\varphi_{rt} + a_t\ell_t + (1-a_t)k_t$. Since the factor shares do not vary greatly, we estimated them statistically as $q_t = \psi_1\varphi_{wt} + \psi_2\varphi_{rt} + \psi_3\,\ell_t + \psi_4\,k_t$ and also in the log-level form. The degree to which these estimates are well determined provides an indication of their constancy. The OLS estimates for the value-added identity are shown in Table 5.[14] (The results for gross output are similarly well determined. They are available upon request.)

<Table 5 here>

These results indicate that, indeed, the factor shares must be relatively constant for the values of the estimated parameters to be so close to them. In the light of this result we estimated the Cobb-Douglas "production function" with a linear time trend of the form $Q = A_0 e^{\lambda t} L^{\alpha} K^{\beta}$. The OLS estimates are reported in Table 6 in logarithmic and exponential growth rate form.

<Table 6 here>

It can be seen that the "output elasticities" greatly differ from their factor shares and the output elasticity of capital is indeed negative. The question arises as to why the parameters in Table 5 are close to the factor shares, while those in Table 6 are so different. Note that the Cobb-Douglas production function in Table 6 would simply be the income identity if the factor shares and φ_t were constant. If this were the case, the Cobb-Douglas relationship would have yielded elasticities that were equal to the shares, and an R^2 equal to 1, as it would be merely capturing the identity. We saw that in Table 5 that factor shares are indeed roughly constant. Therefore, the reason for the observed estimates of the coefficients is the result of approximating φ_t by a constant in the specification of the Cobb-Douglas. The results reported in Table 6 indicate that this is a very poor approximation for the data used.

But this result should not come as any surprise in view of the discussion above concerning the estimate of the mark-up, as the path of φ_t is already known. As indicated above, given the weak procyclicality of the wage rate, and that factor shares are roughly constant, the fluctuations in φ_t over time are largely due those of the growth of the profit rate. It is apparent that a constant does not provide a good approximation.[15] We conjecture that we need a rather complex functional form to approximate correctly φ_t (perhaps a function of sines and cosines). Finding this exact form might not be easy, but it certainly exists, and it is the one that would take us back to the identity. The estimated production function would look like a Cobb-Douglas in *L* and *K* (with the elasticities very close to the factor shares), and would have this added complex term.[16]

The negative elasticity of capital may now be readily explained. As noted above, this is not an unusual result. It is because the rate of return, which varies procyclically, is omitted from, or wrongly approximated in, the regression.[17] The effect of this omission may be readily calculated. The OLS estimator of α in the Cobb-Douglas is (in deviations from the mean) is given by:

$$\hat{\alpha} = \frac{(\Sigma q_t k_t)(\Sigma \ell_t^2) - (\Sigma q_t \ell_t)(\Sigma \ell_t k_t)}{(\Sigma \ell_t^2)(\Sigma k_t^2) - (\Sigma \ell_t k_t)^2} \qquad (19)$$

To calculate the bias, substitute $a_t\varphi_{wt} + (1-a_t)\, \varphi_{rt} + a_t\ell_t + (1-a_t)k_t$, derived from the identities given by equations (4) and (5), for q_t in equation (19). Taking expectations gives:

$$E(\hat{\alpha}) = (1-a) + a\frac{[(\Sigma k_t\varphi_{wt})(\Sigma \ell_t^2) - (\Sigma \ell_t\varphi_{wt})(\Sigma \ell_t k_t)]}{(\Sigma \ell_t^2)(\Sigma k_t^2) - (\Sigma \ell_t k_t)^2} +$$

$$+ (1-a)\frac{[(\Sigma k_t\varphi_{rt})(\Sigma \ell_t^2) - (\Sigma \ell_t\varphi_{rt})(\Sigma \ell_t k_t)]}{(\Sigma \ell_t^2)(\Sigma k_t^2) - (\Sigma \ell_t k_t)^2} \qquad (20)$$

The covariances between the stock of capital and the growth rates of the wage and profit rates are negative; while those between the growth rates of employment and the growth rates of the wage rate, profit rate, and stock of capital are positive. This indicates that the second and third terms in (20) will be negative. A similar argument shows why the estimated coefficient of labor in the Cobb-Douglas production function is often well above the value of labour's share.

In many production function studies it has been found that adjusting the growth of capital for changes in capacity utilization improves the goodness of fit, and the coefficient on the growth of capital becomes positive and close to capital's factor share. The reason for this is that the adjusted-for-utilization-capacity stock of capital is procyclical, and given that capital's share is constant, this implies that the derived rate of profit will exhibit less fluctuation. Consequently, the specification of the putative production function more closely approximates the underlying identity (Felipe and Holz, 2001, McCombie, 2001).

There is a further issue that needs to be addressed. It could be held that our argument could be applied to *any* regression analysis, and thus it would be a pointless exercise.[18] For example, suppose the true relationship between two variables, y and x, is $y = c + b_8x$. It is possible, according to this argument, to define a third variable as $z \equiv y - x$ and that we show erroneously that the supposedly true model is always actually $y = b_9x + b_{10}z$, where b_9 and b_{10} are, by definition, unity. The estimate of b_8 is simply that of b_9 (i.e. unity) biased by the omission of z. This, it could be argued, is all the critique amounts to. Of course, there are many behavioural relationships where z, so defined, would have no meaning, if only because y and x are variables measured in different units and which it is not possible to add or subtract. In fact, this is likely to be true of the vast majority of regressions. To take a simple example. Suppose the regression was trying to explain differences the observed rate of profit (y) in terms of the industry concentration ratios (x). z has no meaningful interpretation in this case. One could easily find other examples. However, this counter-argument to the critique presented in this paper is not convincing, even when we can calculate z.

To see this, let us take a case where an identity does not pose any problems. Consider the case of the consumption function (excluding, for expositional purposes, the government sector and assuming a closed economy). With the usual notation, the specification of the simplest function is $C = c + b_{11}Y$, where b_{11}, the parameter of interest, is the marginal propensity to consume. In other words, it implicitly tests whether there is a stable relationship between changes in consumption and income, allowing savings also to vary. There is, of course, an underlying identity $Y \equiv C + S$ (and $S \equiv I$), which implies that, ignoring questions of exogeneity, the relationship can be expressed variously as $S = -c + (1-b_8)Y$ or as the auxiliary equation, $C = c/(1-b_{11}) + b_{11}/(1-b_{11})S$, from which estimates of the marginal propensity to consume may be derived. There would indeed be little point in estimating $C = b_{12}Y + b_{13}S$, since all b_{12} shows is how consumption varies with a change in income, holding savings constant. The estimate of b_{12} *must* by definition take a value of unity. This simple example merely shows that the existence of an identity does not necessarily preclude the testing of refutable hypotheses. The consumption function does provide some meaningful economic information (an estimate of the marginal propensity to consume and its standard error).

However, this argument is qualitatively different from the one with regard to the production function.[19] The best statistical fit of the putative production function is given by the *whole* identity. There is no way that the estimation of the identity, or of any auxiliary equations, pertaining to the identity can necessarily tell us anything about the technical conditions of production. Moreover, as we have argued above, if the weighted growth of the wage and the profit rate is constant then there is no bias and μ will equal unity, but this cannot be interpreted as implying that perfect competition prevails. The critique of Hall does not depend *solely* on the existence of a bias. It rests on the fact that the best fit to any production function is merely reflecting an underlying identity and the estimate of μ must equal unity. If this identity is not correctly specified then the estimate of μ will differ from unity.

Of course, one could always start by *assuming* the existence of an aggregate production function, as does Hall, and interpret the results accordingly. But as Simon (1979) has persuasively argued, the principle of parsimony, or Occam's razor, suggests that all the estimates are picking up are the underlying accounting identity. As we have noted above, there may not even be any well-defined underlying aggregate production function because of the existence of the well-known aggregation problems and/or the problems thrown up by the Cambridge Capital Theory Controversies. Nevertheless, we would still get the same statistical results and we would also find that the data give a good statistical fit to a freely estimated aggregate production function (i.e., one that does not make use of the observed factor shares as in Hall's method).

7. Conclusions

This paper has shown that some recent attempts to estimate econometrically the degree of market power and the degree of returns to scale are problematical. The method pioneered by Hall is based on a comparison of rates of change of output and inputs based on the usual neoclassical assumptions and the existence of a well-behaved production function. However, there is a problem in that there is also a relationship between the growth of output in value terms and that of inputs (together with factor prices) given by the underlying accounting identity. Because of this, it has been shown that the estimate of the putative mark-up is also the same as unity plus the size of the omitted variable bias inherent in estimating the (misspecified) identity. It turns out that the fact that the estimate of the coefficient of the growth of the labor-capital ratio, weighted by its revenue share, differs from unity is simply due to the fact that the weighted growth of factor prices varies procyclically. This is also the reason why estimates of the supposed degree of returns to scale find such large magnitudes. There are a number of reasons why this procyclical fluctuation may occur (e.g., cyclical variation in capacity utilization rates) that have nothing to do with the degree of competition. There is no way to identify Hall's model (as there are not two behavioural equations) and to show unambiguously that what he (and others) have estimated is the value of the mark-up. Indeed, as has been noted above, Shaikh (1974, 1980) and Simon (1979) have pointed out that, for reasons of parsimony or Occam's razor, the data are more likely to be only reflecting the identity. Whatever view is taken, Hall's procedure must be viewed with a great deal of caution.[20]

Table 1. Estimates of the Mark-up: Summary of Findings and Methods

	Size of the mark-up μ	Output Measure	Estimation Method
Hall (1986)	Large	Value Added	IV
Hall (1987)	Large	Value Added	IV
Domowitz et al. (1988)	Small	Gross Output	IV
Hall (1988a, 1990)	Large	Value Added	IV (Inverse regression)
Waldman (1991)	Not applicable	Data construction and instruments problems	Not applicable
Norrbin (1993)	Small	Gross Output	IV
Basu (1996)	Small	Gross Output	SUR

Notes: The classification of the mark-ups as "large" or "small" refers to how much they depart from μ=1, and is relative to Hall's findings. Hall (1986) used as an instrument the growth rate of real GNP. Hall (1987) used five sets of instruments: (i) oil, oil lagged, and three military variables; (ii) three military variables; (iii) oil, military variables, and political dummy; (iv) military variables and political dummy; (v) rate of growth of real GNP. Hall (1988a &b, 1990) used military expenditures, oil price (both in growth rates), and a political party dummy. Domowitz et al. (1988) ran the regressions with two sets of instruments: one was output, and the other one were military expenditures and the import price. Norrbin (1993) used the same three instruments as Hall (1988). Basu (1996) did not directly estimate the mark-up. He inferred this result from the rest of his work. Since he estimated approximately constant returns to scale, and in practice we do not observe large pure profits, it must be the case that mark-ups of price over marginal cost must also be small. He used Seemingly Unrelated Regression (SUR) estimation.

Table 2 Value Added and Gross Output Mark-ups

Value added: $(q_t - k_t) = \varphi + \mu(a_t n_t)$

$a_t n_t = -1/\varphi + (1/\mu)(q_t - k_t)$

Gross output: $(y_t - k_t) = \varphi' + \mu(s_{Lt}n_t + s_{Mt}(m_t - k_t))$

$(s_{Lt}n_t + s_{Mt}(m_t - k_t) = -1/\varphi' + (1/\mu)(y_t - k_t)$

	Value Added			Gross Output		
	OLS	IV	Inverse IV[a]	OLS	IV	Inverse IV[a]
φ (or φ')	0.038	0.029	0.090	0.009	0.008	-0.009
	(6.42)	(1.65)	(1.05)	(7.23)	(5.37)	(-5.72)
μ	3.573	2.747	8.992	1.425	1.346	1.380
	(10.54)	(1.62)	(1.05)	(37.53)	(9.00)	(9.52)
$\overline{R}^2$	0.775	n.a.	n.a.	0.978	n.a.	n.a.
DW	1.300	n.a	n.a.	1.925	n.a.	n.a.
S.CORR χ^2_1	5.471	0.447	0.550	0.036	0.001	0.003
RESET χ^2_1	0.073	0.752	0.103	0.755	0.400	1.163
NORM χ^2_2	0.130	1.711	2.681	0.901	0.667	0.953
HET χ^2_1	0.001	1.184	0.751	0.065	0.109	0.348
INST χ^2_2	n.a.	5.956	2.511	n.a.	2.045	2.286
$H_0{:}\mu = 1$ χ^2_1	57.632	1.059	0.872	125.37	5.358	6.871

http://www.qaa.ac.uk/revreps/subjrev/Law/Law%20Index.htm

Notes:
t-statistics are in parentheses.

S.CORR χ^2_1 is the Lagrange Multiplier for serial correlation.
RESET χ^2_1 is Ramsey's RESET test for the functional form.
NORM χ^2_2 is the normality test based on the skewness and kurtosis of the residuals.
HET χ^2_1 is the heteroscedasticity test based on the regression of the squared residuals on squared fitted values.
INST χ^2_2 is Sargan's test is the test for the validity of the instruments.

Critical values for chi-square tests: χ^2_1=3.84; χ^2_2=5.99. These are values for a 5% confidence level.

[a] The reported coefficients φ (φ') and μ are derived from the inverse regression coefficients.

n.a. denotes not applicable

Source: Data for U.S. aggregate manufacturing for the period 1958-1991 derived from the NBER productivity database

Table 3 OLS Auxiliary Regressions

Equation	Regressand	Constant	Slope Coefficient	$\overline{R}^2$	DW
Value Added					
(13a)	$a_t \varphi_{w\,t}$	0.008 (4.06)	0.251 (2.05)	0.091	1.069
(13b)	$(1\text{-}a_t)\varphi_{r\,t}$	0.029 (6.64)	2.327 (9.12)	0.720	1.600
Gross Output					
(14a)	$\alpha_t \varphi_{w\,t}$	0.003 (4.61)	0.078 (3.69)	0.283	1.153
(14b)	$(1\text{-}\alpha_t\text{-}b_t) \varphi_{v\,t}$	0.005 (4.72)	0.387 (11.78)	0.811	1.925
(14c)	$b_t \varphi_{p_{Mt}}$	0.001 (0.87)	-0.042 (-1.47)	0.035	2.111

Note: t-statistics are in parentheses.
Source: See Table 2

Table 4 Unrestricted Mark-up Regressions

Value Added: $(q_t - k_t) = \varphi + \mu_1 (a_t \ell_t) - \mu_2 (a_t k_t) + u_t$

Gross Output: $(y_t - k_t) = \varphi' + \mu_1 (s_{Lt} \ell_t) - \mu_2 (s_{Lt} k_t) + \mu_3 (s_{Mt} m_t) - \mu_4 (s_{Mt} k_t) + u_t$

	Value Added		Gross Output	
	OLS	IV	OLS	IV[a]
φ (φ')	0.051	0.074	0.013	0.044
	(5.19)	(1.73)	(3.43)	(1.20)
μ_1	3.510	4.481	1.141	4.789
	(10.55)	(1.92)	(2.94)	(1.20)
$-\mu_2$	-4.732	-6.862	2.699	15.621
	(-6.07)	(-1.73)	(1.34)	(0.72)
μ_3	–	–	1.453	0.440
			(14.95)	(0.38)
$-\mu_4$	–	–	-3.252	-9.581
			(-3.80)	(-1.05)
$\overline{R}^2$	0.786	n.a.	0.981	n.a.
DW	1.190	1.020	2.370	1.610
$S.CORR$	7.144	3.748	1.336	0.111
H_0^a (χ_1^2)	2.67	1.35	n.a.	n.a.
H_0^b (χ_3^2)	n.a.	n.a.	7.74	1.01

Notes: [a] Since there are five regressors, at least one additional instrument must be added. The growth of military expenditures and the growth of oil price, both lagged one period, were also included.

n.a. denotes not applicable.

S.CORR χ^2_1 is the Lagrange Multiplier for serial correlation

$H_0^a : \mu_1 = \mu_2$; $H_0^b : \mu_1 = \mu_2 = \mu_3 = \mu_4$. Critical values (5%): χ_1^2 =3.84; χ_3^2 =7.81.

Source: See Table 2

Table 5 The Value Added Accounting Identity

Equation	ψ_1	ψ_2	ψ_3	ψ_4	
(i)	0.415 (22.71)	0.562 (60.75)	0.447 (26.58)	0.557 (26.92)	$\overline{R}^2 = 0.9989$;
(ii)	0.415 (30.19)	0.574 (82.63)	0.434 (35.05)	0.572 (58.86)	$\overline{R}^2 = 0.99997$

Notes:

Equations (i) and (ii) are the estimates using growth rates and the logarithms of the levels respectively. The constant is not reported for equation (ii).
Equation (i) was estimated using the Exact ML AR(1) method and equation (ii) by the Exact ML AR(2) method.
n.a. denotes not applicable.
t-statistics are in parentheses.

The average labour and capital share for the period is 0.435 and 0.565. The standard deviation of labour's share is 0.049.

Source: See Table 2

Table 6 The Cobb-Douglas Production Function

Equation	λ	α	β	
(i)	0.076 (5.52)	1.524 (13.04)	-2.041 (-4.44)	$\overline{R}^2$ = 0.831; DW = 1
(ii)	0.052 (5.36)	1.535 (11.12)	-1.142 (-3.22)	$\overline{R}^2$ = 0.991; DW = 1.

Notes:

Equations (i) and (ii) are the estimates using growth rates and the logarithms of the levels respectively. The constant is not reported for equation (ii).
Equation (i) was estimated using the Exact ML AR(1) method and equation (ii) by the Exact ML AR(2) method.
n.a. denotes not applicable.
t-statistics are in parentheses.

Source: See Table 2

REFERENCES

Abbott, T.A., Griliches, Z. & Hausman, J.A. (1998) Short Run Movements in Productivity: Market Power versus Capacity Utilization, in: Z. Griliches (Ed.) **Practicing Econometrics: Essays in Method and Application** (Edward Elgar).

Barsky, R., Bergen, M., Dutta, S. & Levy, D. (2000) What Can the Price Gap Between Branded and Generic Tell US about Mark-ups? Paper presented at the NBER Conference on Income and Wealth, Scanner Data and Price Indexes (September 2000).

Basu, S. & Fernald, J. G. (1995) Are Apparent Spillovers a Figment of Specification Error? **Journal of Monetary Economics**, 36, pp. 165-188.

Basu, S. and Fernald, J. (1997) Returns to Scale in Production: Estimates and Implications. **Journal of Political Economy**, 105, pp. 249-283.

Bresnahan, T. F. (1989) Empirical Studies of Industries with Market Power, in: R. Schmalensee & Willig, R.D. (Eds.) **Handbook of Industrial Organization**, Vol.II (Elsevier Science).

Domowitz, I., Hubbard, R. G., & Petersen, B. C. (1988) Market Structure and Cyclical Fluctuations in U.S. Manufacturing, **Review of Economics and Statistics**, 70, pp. 55-66.

Eden, B., & Griliches, Z. (1993) Productivity, Market Power and Capacity Utilization when Spot Markets are Complete, **American Economic Review, Papers and Proceedings**, 83, pp. 219-223.

Felipe, J. (2001) Endogenous Growth, Increasing Returns, and Externalities: An Alternative Interpretation of the Evidence, **Metroeconomica**, 54, 391-427.

Felipe, J. & C. Holz (2001) Why Do Aggregate Production Functions Work? Fisher's Simulations, Shaikh's Identity, and Some New Results, **International Review of Applied Economics**,15, 261-285.

Felipe, J and J.S.L. McCombie (2001), How Sound are the Foundations of the Aggregate Production Function? Georgia Institute of Technology and Downing College, Cambridge (mimeo).

Felipe, J. & J.S.L. McCombie (2002), Methodological Problems with Neoclassical Analyses of the East Asian Miracle, **Cambridge Journal of Economics** (forthcoming).

Fisher, F. (1971) Aggregate Production Functions and the Explanation of Wages: A Simulation Experiment, **The Review of Economics and Statistics**, LIII, pp. 305-325.

________ (1993) **Aggregation. Aggregate Production Functions and Related Topics** J. Monz (Ed.) Cambridge, MA, The MIT Press.

Hall, R. E. (1986) Market Structure and Macroeconomic Fluctuations, **Brookings Papers on Economic Activity**, 2, pp. 285-322.

________ (1987) Productivity and the Business Cycle, **Carnegie-Rochester Conference Series on Public Policy**, 27, pp. 421-444.

________ (1988a) The Relation between Price and Marginal Cost in U.S. Industry, **Journal of Political Economy**, 96, pp. 921-947.

________ (1988b) Increasing Returns: Theory and Measurement with Industry Data, National Bureau of Economic Research.

________ (1990) Invariance Properties of Solow's Productivity Residual, in:
P. Diamond (Ed.) **Growth/Productivity/Employment** (Cambridge, MA, The MIT Press).

Harcourt. G.C. (1972) **Some Cambridge Controversies in the Theory of Capital** (Cambridge, MA Cambridge University Press).

Jorgenson, D. W. & Griliches, Z. (1967) The Explanation of Productivity Change, **Review of Economic Studies**, 34, pp. 249-283.

Jorgenson, D.W., Gollop, F. & Fraumeni, B. (1987) **Productivity and U.S. Economic Growth**. (Cambridge, MA, Harvard University Press).

Kaldor, N. (1956) Alternative Theories of Distribution, **Review of Economic Studies**, XXIII, pp. 83-100.

Lucas, R. E. (1970) Capacity, Overtime, and Empirical Production Functions, **American Economic Review. Papers and Proceedings**, 60, pp. 23-27.

McCombie J.S.L. (1987) Does the Aggregate Production Function Imply Anything about the Laws of Production? A Note on the Simon and Shaikh Critiques. **Applied Economics**, 19, pp. 1121-36.

McCombie, J.S.L. (2000-2001), The Solow Residual, Technical Change and Aggregate Production Functions, **Journal of Post Keynesian Economics,** 23, pp.267-297 (erratum, 23 no.3 p.544)

McCombie, J.S.L. (2001) What do Aggregate Production Functions Show? Second Thoughts on Solow's "Second Thoughts on Growth Theory" **Journal of Post Keynesian Economics, (**forthcoming).

McCombie, J.S.L. & Dixon, R. (1991) Estimating Technical Change in Aggregate Production Functions: A Critique, **International Review of Applied Economics**, 5, pp. 24-46.

Norrbin, S. C. (1993) The Relation between Price and Marginal Cost in U.S. Industry: A Contradiction, **Journal of Political Economy**, 101, pp. 1149-1164.

Oi, W. (1962) Labour as a Quasi-Fixed Input, **Journal of Political Economy**, 70, pp. 538-555.

Phelps-Brown, E.H. (1957) The Meaning of the Fitted Cobb-Douglas Function, **Quarterly Journal of Economics,** 71, pp.546-60.

Prescott, E.C. (1986) Theory Ahead of Business Cycle Measurement, **Federal Reserve Bank of Minneapolis Quarterly Review**, 10, pp.9-22.

Samuelson, P. (1979) Paul Douglas's Measurement of Production Functions and Marginal Productivities, **Journal of Political Economy**, 87, pp. 923-939.

Shaikh, A. (1974) Laws of Production and Laws of Algebra: The Humbug Production Function, **The Review of Economics and Statistics**, 56, pp. 115-120.

________ (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** (Cambridge, MA, Cambridge University Press).

Shaikh, A. (1987) Humbug Production Function in Eatwell, J., Milgate, M. and Newman, P. (Eds) **The New Palgrave. A Dictionary of Economic Theory and Doctrine**. (London: Macmillan).

Shapiro, M. D. (1987) Are Cyclical Fluctuation in Productivity due more to Supply Shocks or Demand Shocks?, **American Economic Review Papers and Proceedings**, 77, pp. 118-124.

Shea, J. (1999) What Do Technology Shocks Do?, in: B. S. Bernanke & Rotemberg, J. (Eds.) **NBER Macroeconomics Annual 1998** (Cambridge, MA, The MIT Press).

Simon, H. A. (1979) On Parsimonious Explanations of Production Relations, **Scandinavian Journal of Economics**, 81, pp. 459-474.

Sims, C. (1969) Theoretical Basis for a Double Deflated Index of Real Value Added, **The Review of Economics and Statistics**, LI, pp. 470-471.

Solow, R. M. (1957) Technical Change and Aggregate Production Function, **Review of Economics and Statistics**, 39, pp. 312-320.

Solow, R.M (1974) "Laws of Production and Laws of Algebra: The Humbug Production Function: A Comment", **Review of Economics and Statistics**, LVI, p.121.

Tatom, J. A. (1980) The "Problem" of Procyclical Real Wages and Productivity, **Journal of Political Economy**, Vol.88, pp. 385-394.

Waldman, R. J. (1991) Implausible Results or Implausible Data? Anomalies in the Construction of Value-Added Data and Implications for Estimates of Price-Costs Mark-ups, **Journal of Political Economy**, 99, pp. 1315-28.

Zarnowitz, V. (1999) Theory and History Behind Business Cycles: Are the 1990s the Onset of a Golden Age?, **Journal of Economic Perspectives**, 13, pp. 69-90

Acknowledgements We are grateful to the participants in the Economics Seminars of the Centro de Estudios Monetarios y Financieros, CEMFI (Madrid, Spain), Universidad Autonoma de Madrid, Georgia Institute of Technology, Manchester Metropolitan University, and the University of Cambridge, as well as to the participants in the ASSA Meetings in New York, January 1999 (session on Productivity chaired by Dale W. Jorgenson), for their valuable comments. We have received useful suggestions from Robert Chirinko, Paul David, Javier Diaz, Nazrul Islam, Pete Klenow, Ishaq Nadiri, and Zhenhui Xu on different versions of the paper, although this does not imply that they all agree with our arguments. We are also thankful to Ludy Pardo for excellent research assistance. Jesus Felipe acknowledges financial assistance from the Georgia Institute of Technology Foundation. The usual disclaimer applies.

Notes

[1] There is a related literature on returns to scale and externalities. This is analyzed in Felipe (2001) and McCombie (2000-2001).

[2] Bresnahan (1989) provides a comprehensive survey of empirical studies on the estimation of market power. He concluded that industry case studies for some concentrated industries tend to indicate the existence of substantial market power. The main difference between this literature and that pioneered by Hall is that the former uses case studies, and the mark-up is calculated by estimating the slope of the demand curve.

[3] Abbott, Griliches, and Hausman (1998) and Eden and Griliches (1993) raise questions about the validity of Hall's instruments.

[4] The direct instrumental variable estimator is $\hat{\mu} = \frac{Cov(q_t, z_t)}{Cov(n_t, z_t)}$, where z_t are the instruments. Under the circumstances described in the text, the denominator becomes an artificially small number and the numerator a high number.

[5] Hall did not test for either lags or the stationarity of the data. As this does not affect our critique, we have followed Hall's procedure for comparability.

[6] Domowitz *et al*. (1988) used a slightly different procedure. Instead of estimating $\mu = p/x$ (where p is the price and x is the marginal cost), they estimated $\rho = (p-x)/p$. This led to an estimating equation slightly different from Hall's. They regressed the Solow residual on the growth of the output/capital ratio.

[7] The notation has been changed to make it the same as that in the rest of the text.

[8] In Hall's theoretical derivation of the estimating equation, he assumes that r is the user cost of capital to a firm and not the rate of profit implicit in the identity. The problem is that the user cost of capital is not generally observable and so the resulting estimating equation does not contain it. Our argument is that because of the accounting identity, this equation does implicitly contain, or can be interpreted as containing, the rate of profit implicit in the identity, as will be seen below.

[9] It should be noted that under the assumptions of perfect competition and the marginal productivity theory of factor pricing the output elasticities will equal the factor shares and φ_t is defined as the Solow residual and commonly interpreted as a measure of exogenous technical progress. However, if all that it is being captured is the underlying identity, this interpretation becomes problematic. This is discussed further in Section 6 below.

[10] When the weighted growth of the gross output inputs is regressed, using OLS, on the weighted growth of the value added inputs, the regression coefficient is 2.11 (t-statistic = 12.30).

[11] The results for 20 two-digit SIC manufacturing industries, which are not reported here but which are available on request from the authors, confirm the results for total manufacturing.

[12] In the case of nondurables, we can be more specific as from Hall (1988, Table 1) the average labour share is 0.72. This, together with an estimated value of the mark-up ratio of 3.096 gives capital's output elasticity of -1.23.

[13] The reason for the difference is that we only used the two instruments for which Hall reports the data, namely the rate of growth of oil prices and military expenditure.

[14] Correcting for the presence of serial correlation does not significantly affect the results.

[15] The analysis, not surprisingly, applies to the case where factor shares do show significant variation over time. If factor shares vary sufficiently, the Cobb-Douglas will not give the best statistical fit, and thus one would have to fit other functional forms, such a translog. It can be shown that the translog can also be considered a more sophisticated transformation of the accounting identity (see Felipe and McCombie, 2002). Thus, all the arguments here follow through for other more complex specifications of production functions.

[16] See Shaikh (1980) for a persuasive discussion of this point, and more recently Felipe and Holz (2001) and Felipe and McCombie (2002). For this particular data set, we fitted the expression $lnQ_t = 0.029 lnA(t) + 0.303 lnL_t + 0.782 lnK_t$ (with t-values 2.29, 4.30, and 15.48 respectively. $R^2 = 0.919$, D.W.= 0.50), where $A(t) = sin\ t + cos$

t^2. (*Sin* is the sine function, *cos* is the cosine function, and *t* is a time trend.) Although this result does not give an exact fit with the estimated coefficients of *lnL* and *lnK* equalling the factor shares, it shows that finding the right approximation is simply a matter of persistence.

[17] If the weighted average of the growth rates of the wage and profit rates is constant, it will be orthogonal to the growth of the factor inputs. Consequently, if the latter vary procyclically, it is likely that the growth rates of the wage and profit rates will be constant (factor shares are roughly constant) and the coefficients unbiased. As the growth of capital shows relatively little variation but the growth of the profit rate varies procyclically, if the latter is omitted from the regression, it will bias the coefficients of labour and the stock of capital. Conventional production function studies often adjust the capital stock for changes in capacity utilisation which induces a procyclical pattern in the growth of the capital stock and brings the values of the estimated coefficients closer to the relevant factor shares.

[18] This possibility was pointed out by a referee.

[19] Output in the production function should theoretically be terms of physical quantities, but these are proxied by constant price value data. In these circumstances, the fact that labour is paid, say, a constant fraction *m* of the homogeneous output and capital is paid *(1-m),* does not in any way *cause* the output elasticities of labour and capital to take the values of *m* and *(1-m).* If the factors are paid their physical marginal products, then this is the only time factor shares will equal the output elasticities. In practice, the use of value data means that this is not the case: the factor shares will always equal the output elasticities and specification of the production function is isomorphous with the full accounting identity. The causation runs, in this sense, from the values of the factor shares to the putative output elasticities. See

Felipe and McCombie (2001) and McCombie (2001) for a further discussion of this issue.

[20] It should be noted that Barsky *et al*. (2000) have proposed a purely micro-economic procedure to estimate mark-ups that yields very interesting results and avoids the problems discussed in this paper.