ISSN 0265-8003

AGGREGATE PRODUCTION FUNCTIONS AND PRODUCTIVITY MEASUREMENT: A NEW LOOK

John Muellbauer

Discussion Paper No. 34
November 1984

Centre for Economic Policy Research
6 Duke of York Street
London SW1Y 6LA

Tel: 01 930 2963

CEPR Discussion Paper No. 34
November 1984

Aggregate Production Functions and Productivity Measurement: A New Look*

ABSTRACT

Many economists currently take a somewhat jaundiced view of the estimation of aggregate production functions. Three problems seem particularly troublesome; the unobservables problem, especially with regard to utilization, the aggregation problem and the simultaneous equation problems. This paper presents theoretical arguments and empirical evidence from British manufacturing for the view that the first of these is the most serious with important dimensions in the measurement of capital and output as well as that of utilization.

New light is shed on two classic questions. One was first raised by Feldstein (1967) who observed in a cross-section context that the elasticity of output w.r.t. average observed hours of work significantly exceeded the elasticity w.r.t. employment. Craine (1973) observed a similar result for time-series data. The other question is one with which most researchers on productivity have struggled: how to correct productivity for cyclical variations in the utilization of inputs. A novel answer based on the use of overtime hours data is found to give excellent empirical results.

JEL classification: 226, 631, 641.

Keywords: British manufacturing productivity, aggregate production functions, unobservables.

John Muellbauer
Nuffield College
Oxford OX1 1NF
0865 248041

*I am grateful to Richard Layard for the use of facilities at the Centre for Labour Economics at the London School of Economics; to the ESRC for finance under grants HR6235, HR8745 and B00220012; and to Dr Lionel Mendis for his skilled programming. Particularly helpful comments on earlier versions were received from the CSO, Angus Deaton, Catherine Elwes, Ian Gazeley, Terence Gorman, David Grubb, David Heathfield, David Hendry, Lionel Mendis, Doug McWilliams, Jim Poterba, Maurice Scott. Responsibility for errors remains mine.

NON-TECHNICAL SUMMARY

One of the main reasons for estimating an aggregate production function is to measure productivity. But three problems seem particularly troublesome: the unobservables problem, especially with regard to utilization, the aggregation problem and the simultaneous equation problem. This paper presents theoretical arguments and empirical evidence on quarterly data for British manufacturing over 1956-1973 for the view that the first of these is the most serious. There are important measurement issues for capital and output as well as labour utilization.

The new proposal for the measurement of labour utilization rests on the idea that high utilization rates are reflected in high aggregate weekly overtime hours being observed. But since most workers are paid for a standard work week even when underutilized, no good direct measure of underutilization exists. The unobserved mean of the whole distribution of utilization rates can, however, be inferred from the observed truncated upper tail as a non-linear function of overtime hours relative to normal hours. With an allowance for the British institutional feature of 'systematic' overtime, such a measure gives excellent empirical results. It also solves the problem of how to correct labour productivity for variations in cyclical utilization to which it is notoriously sensitive. Furthermore, it provides a solution to a question first raised by Feldstein (1967). Feldstein on cross-section data and Craine (1973) for time-series found that the elasticity of output w.r.t. average observed hours of work significantly exceeds that w.r.t. employment. Feldstein suggested fixed costs as a potential explanation. The alternative or additional explanation put forward here relies on the fact that because of the non-linearity of the labour utilization-overtime relationship, the elasticity of output w.r.t. average paid for hours, given normal hours, exceeds that of employment.

The paper also finds circumstantial evidence for large measurement errors in the British official capital stock figures and for smaller measurement errors in the output figures which can, however, significantly distort the short run picture of productivity changes and of activity levels.

1. Introduction

One of the main applications of estimated aggregate production functions is to the measurement of productivity. This paper takes a new look at the underlying methodology. Though there have been waves of enthusiasm in the past for the estimation of production functions on aggregate time series data, many economists currently take a somewhat jaundiced view of such activity. Three potential problems could be particularly troublesome. These are the unobservables problem, especially with regard to utilization, the aggregation problem and the simultaneous equations problem. The view taken below is that the first of these is by far the most important though aggregation plays an important role in coping with it. The theoretical arguments which are presented are supported by a substantive piece of empirical work on quarterly British manufacturing data for 1956–83.

New light is shed on two classic questions. One was first raised by Feldstein (1967) who observed in a cross-section context that the elasticity of output with respect to average hours of work significantly exceeded the elasticity w.r.t. employment. Feldstein suggested fixed costs as an explanation. He argued plausibly that part of the daily or weekly time input is taken up with starting up or winding down production. A similar effect arises from an increase in hours increasing the intensity of utilization of the capital stock without adding to fixed interest costs. Craine (1973) and others have supported Feldstein's cross-section results with aggregate time series evidence.

The other question is one with which most researchers on productivity have struggled at some time. This is how to correct productivity for the pronounced cyclical variations, or to put it another way, the variations in utilization of inputs, to which it is subject. Among the solutions which have been proposed are to use general distributed lags in estimating employment or production functions to pick up short term fluctuations in utilization, to use the unemployment rate or the rate of profit as a cyclical indicator, to survey firms on whether they are working at full capacity and use the survey mean as an indicator and to measure electricity consumption as a percentage of the installed wattage to indicate the level of utilization. There are difficulties with all these proposals and I believe that the arguments against the solution proposed below are less severe.

The paper is structured as follows. Section 2 deals with aggregation problems abstracting from the utilization issue. Section 3 is devoted to measurement problems for labour input with the main emphasis on the measurement of labour utilization. I argue that there are no good direct measures of 'underutilization' but that overtime hours of operatives are a good indicator of 'overutilization'. In aggregate, underutilization should be low when overutilization is high and vice versa. By a statistical aggregation argument I show how to construct a measure of average utilization from a non-linear function of average overtime hours per operative scaled by normal hours. Appendix 1 derives an alternative utilization measure.

Section 4 discusses measurement problems for capital. There is a widely held position that the capital stock should not be adjusted for utilization. One potential source of information on capital utilization is a survey of firms' opinions on capacity utilization and this suggests a test of the hypothesis that such data contain no information additional to that in labour utilization. Appendix 2 shows how a capacity utilizaton measure can be derived from such data. The main focus is on the fact that in most countries gross capital stock data are constructed from gross investment data under fixed service lives assumptions. These assumptions fly in the face of the responses economic theory would predict for retirement decisions when prices, wages, taxes and demand conditions vary. Without data on retirement decisions, there seems little prospect in applying theories of scrapping to derive proxies for unobserved scrapping. The proposal to deal with this problem through the use of shifts in time trends is thus inevitably crude.

Section 5 discusses measurement problems in output. Methods of measuring output differ somewhat in different countries but British practice is probably quite representative. Particular emphasis is given to four problems. The first arises from approximating changes in real value added by applying fixed value added weights to changes in gross output volumes. This results potentially in what I call the 'gross output bias'. The other three biases arise because of problems with the price deflators which are used to deflate the current price data which are the major source for the output index. Thus there is potentially an 'domestic price bias' because, in the absence of reliable export price indices, the CSO uses domestic wholesale prices to deflate the exported component of output. There is potentially a 'list-price bias' because the price data which are collected may not fully reflect transactions prices. Finally, there is potentially a 'price control bias' because of the incentives faced by firms in periods of price controls to distort the prices or specifications of goods in order to bypass these controls. Darby (1984) has argued that such an effect was important for the U.S. Observable proxies are proposed for each of these biases.

Section 6 describes the empirical application to quarterly data for 1955-83 for British manufacturing. The basic equation which best embodies these ideas is subjected to a battery of econometric tests and comparisons with alternatives. These include tests of parameter stability and of whether the residuals are well behaved, tests of the exogeneity of employment and overtime hours and a test of the hypothesis that the utilization of capital as well as that of labour is important. In addition, an explanation is given of how the Feldstein-Craine result arises for such data.

Section 7 summarises the conclusions for the methodology of aggregate production function estimation. The substantive implications of this approach for measuring and understanding British manufacturing productivity are discussed in Mendis and Muellbauer (1984).

2. Aggregation in the Absence of Variations in Utilization

This section pursues what is essentially a Divisia index approach (see Divisia (1952) and, for example, Jorgenson and Griliches (1967)) to measuring the relation between changes in outputs and inputs. The analysis of aggregation is fairly standard, following the pioneering work of Theil (1954). For firm i, I assume there exists a constant returns production function linking value added output, labour and capital. I abstract from aggregation problems over types of labour and types of capital within the firm, though the techniques used can in principle easily be extended to deal with them. The production function is time dependent reflecting the state of technology, established practices governing the allocation of labour within firms and is meant to hold at a normal rate of utilization:

$$q_{it} = F_i\,(\bar{h}_{it}\,\ell_{it},\, K_{it},\, t) \tag{1}$$

where q_i = real value added, ℓ_i = employment, K_i = gross capital stock, and $\bar{h}_i$ = normal hours of work. Thus, in rates of change, suppressing time subscripts and assuming $\bar{h}_i$ constant,

$$d\ell n\, q_i = \alpha_i\, d\ell n\, \ell_i + (1-\alpha_i)\, d\ell n\, K_i + \theta_i\, dt \tag{2}$$

where θ_i reflects changes in technology and work practices and α_i is the elasticity of output with respect to employment. For cost minimizing firms, α_i is the share of labour in factor payments. The fact that the weights on $d\ell n\, \ell_i$ and $d\ell n\, K_i$ add to unity reflects the constant returns to scale assumption. Aggregating across firms with the same value of α, one can define

$$d\ell n\, \bar{q} \equiv \sum_i w_i^q\, d\ell n\, q_i \tag{3}$$

where w_i^q is the share of the ith output in total output for the aggregate in question. Then, suppressing time and sector subscripts,

$$d\ell n\ \bar{q} = \alpha d\ell n\ \bar{\ell} + (1-\alpha)\ d\ell n\ \bar{K} + (\Sigma\ w_i^q\ \theta_i)\ dt$$

$$+ \alpha\ \Sigma(w_i^q - w_i^\ell)\ d\ell n\ \ell_i + (1-\alpha)\ \Sigma(w_i^q - w_i^k)\ d\ell n\ K_i \qquad (4)$$

where w_i^ℓ is the share of the ith employment level in total employment and w_i^k is the analogous capital share. Identifying these aggregates with observable sectoral aggregates within manufacturing, it is clear that the last two terms in (4) are aggregation biases about which one can do nothing without access to individual firm data. For sector I, let us therefore write (4) as

$$d\ell n\ \bar{q}_I = \alpha_I\ d\ell n\ \bar{\ell}_I + (1-\alpha_I)\ d\ell n\ \bar{K}_I + \theta_I\ dt \qquad (5)$$

where the last term in (5) is the sum of the last three terms in (4).

Aggregating across sectors with different α_I:

$$d\ell n\ \bar{q} = (\Sigma\ w_I^q\ \alpha_I)\ d\ell n\ \bar{\ell} + (1-\Sigma\ w_I^q\ \alpha_I)\ d\ell n\ \bar{K} + \Sigma\ w_I^q\ \theta_I\ dt$$

$$+ \Sigma\ w_I^q\ \alpha_I\ \ (d\ell n\ \bar{\ell}_I - d\ell n\ \bar{\ell}) + \Sigma\ w_I^q\ (1-\alpha_I)(d\ell n\ \bar{K}_I - d\ell n\ \bar{K}) \qquad (6)$$

The last two terms are aggregation biases which are measurable. The first is a value added weighted covariance between labour shares in value added α_I and the sectoral growth rates of employment and the second term an analogous expression for capital. Therefore, these aggregation biases are zero when the sectoral rates of change are identical or, more generally, when deviations in sectoral rates of change are distributed independently from sectoral factor shares.

With constant utilization rates and cost minimizing firms, (6) is the basis for accounting for growth by chaining together period to period changes. This is essentially the kind of technique recommended and used by Jorgenson and Griliches (1967). The residual $\Sigma\ w_I^q\ \theta_I$ is a weighted average of the firm θ_i's and of the inescapable aggregation biases represented by the last two terms in (4). A major reason for deriving (6) is to understand how far one would be likely to go wrong if utilization rates were constant in fitting a Cobb-Douglas production function in the form

$$\ln \bar{q} = \text{const.} + \alpha \ln \bar{\ell} + (1-\alpha) \ln \bar{K} + \theta t \qquad (7)$$

Empirically, it would appear that apart from cyclical fluctuations associated with variations in utilization rates, the α_I's for British manufacturing are remarkably stable over time. Since, by the construction of the aggregate output index, the w_I^q are constant weights, this would suggest that the constancy of $\alpha \approx \Sigma w_I^q \alpha_I$ would be a good approximation. The evidence for British manufacturing suggests that the aggregation biases over sectors are small.* Thus there are good reasons for believing that the aggregation problems in fitting (7) to aggregate data are unlikely to be resolved by using industry group data.

Of course, if utilization rates were constant there would be no difficulty about measuring labour productivity by chaining together $d\ln \bar{q} - d\ln \bar{\ell}$, perhaps taking the aggregation bias over sectors, $\Sigma w_I^q \alpha_I (d\ln \bar{\ell}_I - d\ln \bar{\ell})$ into account. In fact, the main argument for production function estimation as opposed to growth accounting via (6) is precisely that it offers a way of finding an econometric model which will pick up varying utilization rates and systematic biases in capital and output measurement. Once one has found such a model one can measure changes in productivity correcting for changes in rates of utilization.

---

*An approximation to the aggregation bias for labour can be obtained by taking $\Sigma w_I^q \alpha_I (\Delta\ln \bar{\ell}_I - \Delta\ln \bar{\ell})$ where w_I^q and α_I are averages of the beginning of period and end of period weights. This gives less than one half of one percent respectively over the entire periods 1955-70 and 1970-83.

3. Utilization and the Measurement of Labout Input

Most writers on productivity discuss the measurement of quality and compositional changes in the labour force. I have nothing new to say here. Effective hours of work per unit time (e.g. one quarter), distinguished from paid for hours, consist of effective hours per week h and working weeks per quarter WW, the latter depending on paid holiday arrangements. Write labour input in (1) as $WW_{it}h_{it}\ell_{it}$.*

Labour utilization, measured as the proportional deviation of effective weekly hours from normal hours, is defined by

$$u_{it} = \ell n\ h_{it} - \ell n\ \bar{h}_{it} \qquad (8)$$

Various ways of proxying utilization have been used in the past. Denison (1979) used the cyclical deviation in the share of profits but this is sensitive to short run movements in input and output prices which may not be reflected in utilization rates. Sometimes unemployment rates of workers have been used, see e.g. Baily (1981), but this seems inappropriate for various reasons including the confusion of supply side effects arising from demographics and supply incentives and the difficulties in defining a sectoral

---

*It should be noted that there is another aspect of production, multiple shift working, which (1) thus amended may not represent well. One can argue that adding a night shift to an existing day shift is like replicating the plant without altering the capital stock. Where d and n superscripts refer to day and night shift magnitudes, this suggests in place of (1),

$$q_{it} = F_i(WW^d_{it}\ h^d_{it}\ \ell^d_{it},\ K_{it},\ t) + \lambda_i\ F_i\ (WW^n_{it}\ h^n_{it}\ \ell^n_{it},\ K_{it},\ t)$$

and λ_i takes account of the possibility that night workers are less productive. For British manufacturing, observations on the proportion of shift workers ranging from 13% to 25% are available only for three dates pre-1973, and there is no option but to ignore this complication. With better data, labour utilization could be defined separately for each shift and aggregated by the proportion of workers in each.

unemployment rate. The Wharton approach of measuring deviations from trend output is little help since it begs the question of what determines the trend. The Jorgenson and Griliches (1967) technique of taking electricity use as a proxy seems sensible but was heavily criticised by Denison (1969) though it has been used with some success by Heathfield (1972, 1983) in a study of the British engineering industry.* Baily (1981, 1982) also tried lay offs, deviations of the rate of change of employment from trend and survey based indices of capacity utilization. Another technique, see Chatterji and Wickens (1982), is to estimate production or employment functions where distributed lags pick up utilization effects, and use the steady state solution of such an equation to measure cyclically corrected productivity growth. However, my experience is that the steady state solution is typically not precisely enough determined to detect trend shifts before 3 or 4 years have elapsed.

The utilization measure proposed in this paper has a stronger theoretical foundation than most of the above. The method involves an old idea, that of 'smoothing by aggregation'. In the aggregate production function we need a concept of aggregate labour utilization, u_t so that:

$$\alpha u_t = \Sigma w_i^q \alpha_i u_{it} \qquad (9)$$

where w_i^q is the share of the i^{th} plant in reference year output. With w_i the wage, $\alpha \approx \Sigma w_i \ell_i / \Sigma p_i q_i$, $\alpha_i \approx w_i \ell_i / p_i q_i$, $w_i^q = p_i q_i / \Sigma p_i q_i$ it follows that

$$u_t \approx \Sigma w_i \ell_i u_{it} / \Sigma w_i \ell_i$$

Since $w_i \ell_i / \Sigma w_i \ell_i$ is a close approximation to the share of the i^{th} plant in the total number of workers, u_t is a close approximation to the average utilization rate averaging over all workers.

---

*Electricity consumption as a percentage of installed wattage is the utilization variable here. One problem with it is that it may be sensitive to variations in electricity prices relative to other inputs.

Let u_j now refer to the j^{th} worker. If such a worker is working overtime $u_j = \ln(\bar{h}_j + \text{overtime hours}_j) - \ln \bar{h}_j$. What cannot usually be observed is below normal utilization when 'undertime' is being worked but the worker is still paid for a normal week. Thus we observe u_j when $u_j \geq 0$ but not when $u_j < 0$. Approximating the distribution of u_j with a continuous density function $\phi(u)$, define

$$u^* = \int_{u>0} u\,\phi(u)\,du \tag{10}$$

Since u_j for $u_j \geq 0$ is overtime proportional to normal hours for the j^{th} worker, think of u^* as proportional overtime averaging over all workers whether working overtime or not.

Analogously, define

$$\hat{u} = \int_{u<0} (-u)\phi(u)\,du \tag{11}$$

and think of $\hat{u}$ as unobserved proportional 'undertime' again averaging over all workers. One can also think of u^* and $\hat{u}$ as truncated means weighted by Prob $(u>0)$ and Prob $(u<0)$. Figure 1 illustrates the distribution of u.

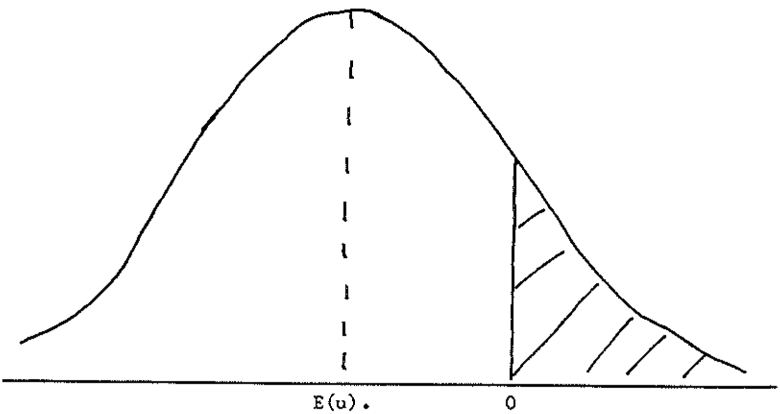

Figure 1: the distribution of proportional deviations of utilization rates

By definition, mean utilization

$$E(u) = u^* - \hat{u}\,. \tag{12}$$

Now imagine the distribution shifting horizontally but with its spread constant. As it moves to the right, u_t^* increases and $\hat{u}_t$ declines. This traces out a smooth trade off between u_t^* and $\hat{u}_t$. For each type of distribution a particular type of trade off will be associated.

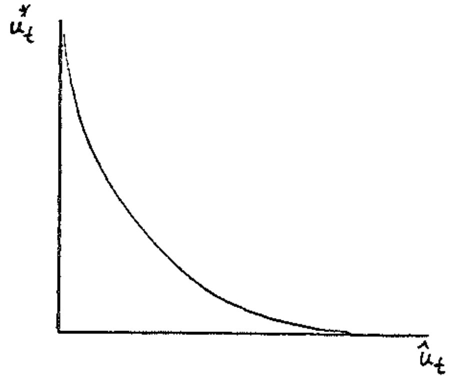

Figure 2: the u_t^*, $\hat{u}_t$ trade off

For example, a symmetric distribution always implies a symmetric trade off. Since the cost per effective hours of paying for a normal week tends to infinity as h_i tends to zero and since h_i has a finite upper limit, it seems likely that the distribution is bounded. A simple functional form which implies both a bounded distribution and which can allow for potential asymmetry is $(u_t^* + c_1)(\hat{u}_t + c_2) = c$ where $c_1 > 0$, $c_2 > 0$ and defined for $u_t^* \geq 0$, $\hat{u}_t \geq 0$. Then, applying (12), mean utilization is*

$$E(u_t) = c_2 + u_t^* - \frac{c}{u_t^* + c_1} \tag{13}$$

*Similar ideas have been used in the context of measuring aggregate excess supply Z from data on unemployment U but with data on vacancies V missing. Assuming UV = c, see Hansen (1970), $Z=U-V=U-cU^{-1}$.

It is an interesting curiosity demonstrated to me in a 1980 correspondence by Angus Deaton that, if $c_1 = c_2 = 0$, this trade-off corresponds to a t-distribution with 2 degrees of freedom. Such a distribution has tails so fat that the variance is unbounded and is thus the assumption implicit in Hansen (1970).

c reflects the spread of the distribution in Figure 1. If this is constant over time, we can 'observe' $E(u_t)$ from (13) even though 'undertime' $\hat{u}_t$ is not observable.

In practice, there is another complication. It appears that for some firms a part of overtime is regarded by the workers as normal or systematic. Indeed, the reductions in normal hours in Britain in the late 1950's and the early 1960's seem to have been accompanied by an increase in systematic overtime. This suggests replacing u_t^* by the "true" concept $u_t^* - s_t$ where s_t refers to the systematic overtime component and where s_t increases as normal hours fall. Thus

$$E(u_t) = u_t^* + c_2 - s_t - \frac{c}{u_t^* + c_1 - s_t} \qquad (14)$$

Appendix 1 discusses how mean utilization can be derived when u_t follows a Normal distribution and shows that the empirical results are remarkably similar to those discussed in the body of the paper below.

4. Capital Measurement, Capacity Utilization and Unobserved Scrapping

It can be argued that a separate concept of capital utilization has cutting power; for example, the tonnage of shipping laid up ought to be subtracted from the total tonnage in a shipping production function. In manufacturing, were the data available, one might want to subtract unused buildings from the total stock of buildings. There is a widespread view, however, see Kendrick (1973), Denison (1974) and Gollop and Jorgenson (1980), in which it is argued that capital is a kind of overhead concept and that a separate concept of capital utilization has no place in a production function, given that the other inputs are correctly measured.

One might ask then what surveys of firms' capacity utilization measure. Price (1977) reports that in a special inquiry into answering practices in the CBI survey* two thirds of respondents measured current output against the capacity only of buildings and plant while one fifth included, in addition, the availability of labour. This sort of information is not decisive on the issue since a relatively high level of output may simply reflect a relatively high level of employment and overtime work. But since one might want to entertain the hypothesis that something special to capital is being measured by such surveys, Appendix 2 below is devoted to the construction of a capacity utilization variable based on the proportion of firms (weighted by size) who report below capacity operation. Section 6 contains an empirical test of whether such a variable has any explanatory power given that the labour utilization measure is incorporated in the production function.

As in most countries, the CSO figures for gross capital stock are derived by a perpetual inventory method from gross investment flows and assumed retirements. The figures on retirements are based on assumptions about

---

*The Confederation of British Industry (CBI) Industrial Trends Survey is similar to surveys in many other countries and asks: "Is your present level of output below capacity (i.e. are you working below a satisfactory or full rate of operation)?"

service lives of different kinds of assets which have remained unchanged over a long period but were revised in 1983. In fact, the assumption is that for each type of asset there is a smooth distribution of service lives around a central estimate, see Griffin (1976). The assumed service lives do not vary cyclically nor respond to relative factor prices.

However, this does not mean that economic scrapping is entirely ignored in the capital stock figures. When a firm purchases new equipment, the second hand market value of equipment it scraps is subtracted from the purchases to give a net investment figure. There is a conceptual problem here since the gross capital stock figures are meant to reflect productive capacity rather than a market asset valuation. Thus, particularly at times of low profitability, the second hand market value may under-estimate the productive capacity of equipment being scrapped. Furthermore, when companies become bankrupt and assets are sold abroad or simply taken out of use, it appears that no allowance whatever for scrapping is made by this procedure. Thus it seems likely that scrapping is understated when economic recession or relative price shifts raise economic scrapping. A similar point is made by Baily (1981) and Raasche and Tatom (1981) to help explain the slowdown in productivity growth which occurred in U.S. manufacturing from 1973.

The question is what observables can be taken as proxies for scrapping not directly observed. Economic theory would suggest estimating a vintage production model in which the capital stock as such does not appear but which exploits the optimality conditions that govern which vintages remain in use Malcolmson and Prior (1979), Malcolmson (1980) and Mizon and Nickell (1983) have estimated such models and report considerable empirical success. However, the results are not informative on questions of productivity change. Shifts in the parameter which measures technical progress in the path-breaking paper by

Malcolmson and Prior and in Malcolmson cannot be identified.* Mizon and Nickell report some difficulty in their model in finding sensible estimates of the technical progress parameter. More work is clearly needed on what can be learned about productivity change from vintage models when no observations on scrapping are available.

Baily (1981) favours a stock market valuation of the capital stock and takes a weighted average of it and the gross capital stock as his measure of the true capital stock. As some of his discussants remark, this may not be ideal since other influences such as variations in interest rates that can have an exogenous source will influence the stock market valuation. Scott (1976, 1981) indeed argues that no observable capital stock concept is meaningful though gross investment is. Apart from unobserved scrapping, he notes that maintenance is not properly measured and questions the assumption made in defining the gross capital stock that each asset remains as productive when new until the end of its life.

The practical response taken below to these difficulties is to fit time trends with linear splines allowing slope changes to occur at times when, on a priori grounds, one would expect a great deal of unobserved scrapping.

---

*Prior restrictions on several of the other parameters are required for identification.

5. Measurement Biases in Output

In most countries, indices of aggregate output are constructed from a mixture of basic indicators, some of them measuring output in physical units and others in current prices deflated by price indices. Both raise issues of how to incorporate quality changes. A great deal has been written on the theory and practice of quality measurement and the quality correction of price indices, see for example, Deaton and Muellbauer (1980), ch.8 and Triplett (1983) for recent overviews.* This paper does not contribute to this literature. As far as estimating aggregate production functions is concerned, measurement biases in output quality are likely to be rather trend like and will show up in the trend coefficients. This paper is concerned with other sources of bias. Although the context of the discussion below is British manufacturing, there are parallels in other countries for all four of the biases to be discussed.

The index of output at constant factor cost aims to measure movements in real value added i.e., to separate out the contribution to final output of labour and capital from that of other inputs such as raw materials and imported intermediate manufactured goods. This separation is not easy to make, especially at an industry level. To do it properly requires frequent input-output tables derived from censuses of production whose costs have been thought prohibitive compared with the value of the information yielded. Instead, the CSO approximates changes in value added for each of a quite disaggregated list of goods by changes in gross output. These are weighted by value added weights from a quinquennial census of production. This is fine as long as value added and gross output change in the same proportion. However, when raw

---

*Though curiously little applied research on the topic has been undertaken for British output statistics.

material prices or prices of imported intermediate manufactured goods change relatively to those of labour and capital, firms have an incentive to substitute.

Substitution can be of two types, the first and probably the larger results in changes in the pattern of output towards goods less intensive in the inputs whose relative prices have increased. This results in no bias since the weights applied to the gross output changes are fixed and so unaffected by relative input price changes. However, the second, substitution for each type of good between labour and capital on the one hand and other inputs on the other does give a bias. Value added increases faster (slower) than gross output when the relative prices of raw materials or imported intermediates increases (decreases) so that the output index understates (overstates) the true increase. Since substitution takes time, one would expect this measurement bias to be negatively correlated with the lagged (log) ratio of raw material prices to domestic wholesale prices of output PR and with a (log) index of foreign competitors' wholesale output prices in Sterling relative to domestic ones, PW.

The theory can be more formally explained as follows. Given substitution possibilities in production between raw materials and other inputs, the analogy of (2), again abstracting from utilization, is

$$d\ell n\, q_i^* = \alpha_i\, d\ell n\, \ell_i + \beta_i\, d\ell n\, K_i + (1-\alpha_i-\beta_i)\, d\ell n\, m_i + \theta_i\, dt \qquad (15)$$

Here q_i^* is gross output. Having assumed constant returns, the derived demand for raw material input takes the form

$$d\ell n\, m_i = d\ell n\, q_i^* + d\ell n\, g_i \text{ (relative factor prices)} \qquad (16)$$

Substituting (16) into (15), gives

$$d\ell n\, q_i^* = \frac{\alpha_i}{\alpha_i+\beta_i}\, d\ell n\, \ell_i + \frac{\beta_i}{\alpha_i+\beta_i}\, d\ell n\, K_i + \left(\frac{1-\alpha_i-\beta_i}{\alpha_i+\beta_i}\right) d\ell n\, g_i$$

$$+ \frac{\theta_i}{\alpha_i+\beta_i}\, dt \tag{17}$$

Note that the weights on $d\ell n\, \ell_i$, $d\ell n\, K_i$ sum to unity and that since g_i ( ) is a decreasing function of PR given some substitution possibilities, a negative response of gross output to PR is implied. These points generalize easily when there is a vector of other inputs.

Bruno (1984) and Bruno and Sachs (1982) suggest that this bias, the 'gross output bias', is an important part of the explanation of the slowdown in productivity growth in industrial countries after 1973. However, as Grubb (1984) has argued, the empirical magnitudes are unlikely to make this a major part of the story.

About two thirds of the British manufactured output index in recent years has been based on value deflated data. A further bias to be considered arises because no satisfactory export price deflators exist for most of output which is exported. The CSO therefore uses domestic wholesale price indices instead. However, because of exchange rate movements, it is likely that in the short run there can be significant divergencies in these domestic prices from the unobserved export prices so that a 'domestic price index bias' in output results. An observable indicator of this bias is the ratio PW of foreign to domestic wholesale prices. Much of this effect operates immediately when exchange rates change but then tends to unwind as competitive and cost pressures act on export prices. This predicts a positive coefficient on PW_t in an equation for measured output with a smaller or zero coefficient in the long run. It is likely that movements in the unobserved export prices are somewhat more attenuated than in foreign wholesale prices but this will be reflected in the estimated coefficients on PW.

A third source of bias arises from another problem with the deflators. Although these aim to capture transactions prices, it is probable that they are partly based on list prices, hence giving rise to a 'list price bias'. Discounts measure the gap between list and transactions prices and are likely to be sensitive to changes in competitive pressure and to changes in underlying costs. An increase in PW reflects a reduction in competitive pressure and so a reduced gap between list and transactions prices. Then measured price indices will tend to understate true price increases and measured output increases overstate true ones. This would imply a positive PW effect, though one that eventually unwinds at least partially, in an equation for measured output. One expects similar effects for increases in cost pressures. Since transactions prices are likely to be more flexible than list prices, an increase in costs will be associated in the short run with measured price increases understating true ones. This would imply positive short run effects for PR and PW, the latter representing imported intermediates, in an equation for measured output. In the long run, these effects should be zero, as list prices adjust fully.

Finally, consider the effects of price controls. Darby (1984) argues that price controls instituted in the U.S. in 1971 were widely evaded, for example, by firms claiming spurious quality improvements or simply relabelling goods. The increase in the official price indices was therefore understated and in output overstated. These biases reversed in 1974 as price controls were taken off and reported output and so productivity fell by more than the true figures. In Britain, price controls were introduced in April 1973, slightly relaxed in December 1974 and August 1976 and replaced in August 1977 by the much weaker Price Code. The Price Commission which operated these policies was finally abolished in 1979. The Price Commission's quarterly reports give figures which measure intervention both in number of cases and by the money value of sales affected. These make it possible to measure roughly the intensity of the controls. The hypothesis is that the more intense the controls, the greater the incentive of firms to evade them and the greater the bias in the official price indices and so in measured output.

6. Empirical Results

(a) The Data

I begin by briefly* describing the data and their sources, whose names are abbreviated as follows: ET is Economic Trends, HABLS is the Historical Abstract of British Labour Statistics, DE is the Department of Employment and DEG is the DE Gazette. Manufacturing is defined by the 1968 SIC.

q_t = index of manufacturing output at constant factor prices, seasonally adjusted and stock adjusted from 1970. This is the only seasonally adjusted variable in the data set. Source: ET, CSO. Range: 65.3 in 1955.1, 115.3 in 1974.2.

ℓ_t = employment in manufacturing, an average of 3 monthly figures and refers to all employees, part time and full time. Source: HABLS, DEG and DE. Range: 5.347 million in 1983.4, 8.491 million in 1965.4.

K_t = gross capital stock in 1980 prices. This includes assets leased to the manufacturing sector and is based on service life assumptions newly introduced in 1983. Source: 1983 National Income and Expenditure and CSO. Range: 88.4 billion in 1955.1, 211.1 billion in 1983.4.

$\bar{h}_t$ = normal hours = 0.4425 x NH_t where NH_t = index of normal hours per week. Source: HABLS, DEG. Range of NH: 100 in 1955.1, 88.6 in 1983.4 and 90.4 from 1968 to 1979, ie. 40 hours per week.

$$OH_t = \frac{\text{weekly overtime hours per operative on overtime x fraction of operatives on overtime}}{\text{normal hours}}$$

an average of 3 monthly observations from 1961 and a mid quarter observation before 1961. Source: HABLS, DEG. Range: 0.0400 in 1958.3, 0.0876 in 1973.4.

---

*A fuller description is in Mendis and Muellbauer (1984).

$PR_t = \ln \frac{\text{wholesale price index for raw materials purchased by manufacturing}}{\text{wholesale price index for home sales of manufacturing}}$

Source: ET. Range: -0.352 in 1972.2, 0.128 in 1974.1.

$PRD_{t-3} = PR_{t-3} - PR_{1969.2}$ from 1970.1 and 0 before 1970.1.

$PW_t = \ln \frac{\text{wholesale price index for foreign competitors}}{\text{wholesale price index for home sales of manufacturing}}$

Source: ET and earlier figures from U.N. Monthly Bulletin of Statistics.

Range: -0.346 in 1981.1, 0.119 in 1976.4.

$PWD_t = PW_t - PW_{1970.1}$ before 1970.1 and 0 from 1970.1.

$PC_t = \frac{\text{Price Commission intervention in £ terms}}{\text{Wholesale price index for home sales of manufacturing}}$

Source: Price Commission Reports. Range: 0 up to 1973.1 and from 1977.4, 7.74 in 1974.1.

TRJ = 0 before observation J, 1 at J, 2 at J+1, 3 at J+2 etc.

Si = 1 for ith quarter, 0 otherwise.

Si TR = Si x trend.

EX1 = Excess of average January and February temperature over 1941-70 mean, in Centigrade and defined for the 1st quarter only. Source: Annual Abstract of Statistics and Monthly Digest of Statistics. Range: -3.0 in 1963.1, 2.35 in 1957.1.

EX2 = Excess of preceding December temperature over 1941-70 mean, in Centigrade and defined for 1st quarter. Source: as for EX1. Range: -3.7 in 1982.1, 3.2 in 1974.1.

π = Proportion of firms operating below full capacity reported by CBI Industrial Trends Survey. Triannual 1958-72, then quarterly. Quarterly interpolation centred at mid-quarter. Source: CBI. Range: 0.38 in 1965.1, 0.84 in 1980.4.

ww = ℓn (52 - 1.2 - average weeks annual holiday entitlement), linear quarterly interpolation. This is an indicatior of the number of weeks in a normal working year, assuming 1.2 weeks of public holidays. The coverage is all manual workers in national collective agreements or Wages Councils orders. Source: HABLS, DEG. Range: 3.887 in 1955.1, 3.827 in 1983.4.

PO = proportion of employees who are operatives, linear quarterly interpolation. Data are biannual from 1963-1974 and otherwise annual. Source: HABLS, DEG. Range: 0.700 in 1980.4, 0.801 in 1955.1.

(b) Derivation of a parsimonious specification

The process which led to the final equation for manufacturing output can be explained as follows. Given the theoretical discussion in Sections 2-5, write an aggregate Cobb-Douglas production function, imposing constant returns, in the form

$$\ell n\ (q_t/K_t) = \alpha_0 + \alpha(\ell n\ (\ell_t\ \bar{h}_t/K_t) + ww_t + u_t)$$

$$+ \text{ output measurement bias effects} + \text{trend effects} \quad (18)$$

where u_t represents the average proportionate deviation from normal of weekly labour utilization. Section 3 suggests how u_t should be measured. In equation (14), take $s_t = s_0 - s_1\ (\dot{NH}_t - 90.4)$ where $s_0>0$, $s_1>0$ and $NH_t - 90.4$ is positive before 1968, zero for 1968-79 and negative after 1979. Since $u_t^* \approx OH_t$, mean utilization is

$$u_t = OH_t + c_2 - s_0 + s_1 (NH_t - 90.4) - c(OH_t + c_1 - s_0 + s_1(NH_t - 90.4))^{-1} \quad (19)$$

I approximate (19) by the expression

$$u_t = \text{const.} + OH_t - c \; OH_t^{-1} - c_0 \; (NH_t - 90.4) \; OH_t^{-1} \quad (20)$$

which avoids the use of non-linear estimation.

To allow for the possibility that the proportion π of firms reported in the CBI Industrial Trends Survey to be operating below full capacity contains information additional to that in the overtime data, I define the variable $CU = (\pi/1-\pi)^{0.4}$, see Appendix 2. cu can then be included in (18) as an additional regressor.

The measurement biases in output discussed in Section 5 are proxied through the variables PR , PW and PC . q_t in (18) refers to measured output so when this exceeds true output there are positive measurement biases. Four biases were considered: the gross output bias, the domestic price index bias, the list price bias and the price control bias. There are good reasons for allowing for dynamic effects in the variables used to proxy these biases. The intensity of price controls PC almost certainly has a lagged effect. The Price Commission reports that firms were changing prices of individual goods about 2 or 3 times per annum. The PC data refer to quarters which are one month in advance of the conventional definition. Thus PC_{t-1} implies an average lag of 2 months and PC_{t-2} one of 5 months. For reasons discussed in Section 5, the lag responses to PR and PW are likely to be more complicated. In practice, rate of change effects ΔPR_{t-j}, ΔPW_{t-j} with $j \leqslant 4$, were included in addition to the levels effects.

The trend effects are considered to be of two kinds. The first represents slow changes in technology and work practices, in aggregation biases and in unmeasured changes in output quality, labour force composition, shift work and paid holidays. The second represents the unmeasured scrapping (or loss of productivity) of capital that would have followed the 1973 oil crisis and the collapse of manufacturing output in 1979-80. Both types are represented by linear splines of the form $\Sigma \beta_i TRJ_i$. This yields a continuous line made of straight line segments which change slope at observations J_1, J_2 etc. To select dates for changes in trend slopes corresponding to the first type of trend effects, cyclical peaks were chosen as candidates on the grounds that these were relatively frequent and to reduce the risks of confusing the effects of trend and cycle. Since the share of labour in manufactured value added is of the order of 0.7, $\ln q_t - 0.7 \ln \ell_t - 0.3 \ln K_t$ was plotted against time to pick out these cyclical peaks. This gave the following dates: 1959.4, 1964.4, 1968.3, 1973.1 and 1979.2. Slope changes reflecting unobserved capital scrapping were investigated between 1974.1 and 1974.4 and between 1979.2 and 1980.3.

The equations estimated also allowed for seasonal dummies and for unusual weather affecting output in quarter 1 through the excess temperature variables EX1 and EX2. Finally, special dummies for strikes and other unusual events were included.

In the initial estimation, the sample was split pre and post-1970. There were two reasons for this. From 1968 and particularly in 1970-1972 there was a major change in the basic sources of output data with physical measures of production increasingly replaced by deflated values from quarterly sales inquiries carried out by the Business Statistics Office.* This should, by the arguments of Section 5, have led to a break in the coefficients of the

*See, CSO (1976). CSO (1959) suggests that about 31% of the data was based on deflated values in the 1954 based index and about 33% for the 1958 based index. CSO (1970) suggests a figure of 40% for the 1963 based index while CSO (1976) suggests 66% for the 1970 based index.

variables corresponding to the output measurement biases. The second reason is that with the moves to flexible exchange rates in 1971 and 1972, it seems likely that a structural break would have occurred in the relationships between price forecasts and past data. To the extent that price expectations play a role in, for example, the substitution effect entailed in the gross output bias, one might therefore expect parameter shifts in the effects of these price variables on measured output. Given that these arguments suggest parameter shifts between 1968 and 1972, so that 1970.1 is a mid-point and given the shift to a stock-adjusted definition of output from 1970.1, this date was chosen for the sample split.

The CBI capacity utilization variable CU is defined only from 1958.3 so that the two periods initially considered were 1958.3 - 1969.4 and 1970.1 - 1983.4. For both periods the coefficient on CU_t was positive, against the prediction of theory, though insignificantly different from zero in both cases. Defining CU for different values of θ in the plausible range made only slight differences to the t-ratios on CU. In contrast, the employment and overtime variables had sensible and significant coefficients in both periods. This suggested that CU is dominated by the overtime hours based concept of utilization. Dropping CU from the equation, it was possible to extend the first period back to 1956.1.

The next step was to search for a parsimonious representation in each period of the general distributed lags in PR and PW. This suggested PW_t, $\Delta_3 PR_t$ and $\Delta_4 PW_t$ in the first period and PR_t, $\Delta_3 PR_t$ and $\Delta_4 PW_t$ in the second. F-tests for the 7 restrictions respectively gave $F_{7,26} = 0.56$ (2.39) and $F_{7,21} = 0.32$ (2.49), where the critical values at the 5% level are given in parenthesis. Having now more parsimonious equations it was thought necessary to go back and check that CU_t was still insignificant in case the earlier finding had been due to overfitting. CU_t proved insignificant again in both periods with the sign in the later period still positive.

The two equations for 1956.1–1969.4 and 1970.1–1983.4 were now simplified further by omitting three of the shifting trends, two trending seasonals, PC_{t-1} (t ratio = 0.6) and refining slightly the (0, 1) dummies. With 6 restrictions for 1956.1–1969.4 and 4 for 1970.1–1983.4, the F-tests are $F_{6,33} = 1.89$ (2.40) and $F_{4,28} = 2.27$ (2.70). The resulting equations are shown as R.1(a) and (b) in which PR_t and PW_t have been replaced by PRD_{t-3} and PWD_t from which the 1970.1 values of PR_t and PW_t have been subtracted.

These equations suggest that pooling might well be an acceptable restriction. There are 13 restrictions: on the intercept, the trend, the responses of output to employment, the two overtime variables, $\Delta_3 PR_t$, $\Delta_4 PW_t$, three seasonals, one trending seasonal and the two excess temperature variables. The F-test is $F_{13,71} = 1.61$ (1.90).

The resulting equation is R.2. The obvious naive alternative hypothesis is $\Delta \ell n(q_t/\ell_t)$ = constant which has a standard error of 0.01936 compared with R.2's of 0.007457. $\ell n(q_t/\ell_t)$ itself is, of course, heavily trended and has a standard deviation of 0.2322. The estimated elasticity of output w.r.t. employment at 0.681 is plausible. OH^{-1} and its interaction with normal hours are both highly significant with the anticipated signs. The latter term suggests that part of the increased overtime following reductions in normal hours in the 1950's and 60's itself became normal and thus should not be included in cyclical overtime. The cumulative trend effect, in annualized terms is 1.8% up to 1959.3, then 2.5% up to 1972.4, dropping to 0.8% from 1973.1 to 1979.2, −2.3% from 1979.3 to 1980.2 and back to 2.5% from 1980.3. This is consistent with prior expectations of higher rates of unobserved scrapping in the periods 1973.1 to 1979.2 and 1979.3 to 1980.2.

The price control variable is significant, with a plausible lag and size of coefficient. This result parallels that of Darby (1984) for the U.S. The relative price of raw materials to output has a significant post-1970 levels effect consistent with a rather small gross output bias.* The rate of change effect like that of the relative price of foreign manufactures is consistent with the 'domestic price bias' and the 'list price bias' discussed in Section 5. The pre-1970 levels effect of the relative price of foreign manufactures suggests that export prices were able to diverge more permanently from foreign wholesale prices in that period. Finally, the excess temperature variables suggest a significant first quarter output effect as the result of unusual weather in December, January and February.

(c) Further tests and comparisons

The main features of the extensive tests and comparisons of R.2 with alternative specifications are discussed next. Since there is inevitably some arbitrariness about some of the 0,1 dummies which were included, R.3 shows the effect of excluding all except for the 1972.1 miners' strike and the 1974.1 'three-day week' energy crisis dummies over which there can be no argument. The results show that no coefficient in R.3 differs by more than one estimated standard error from the value estimated in R.2 which is reassuring for R.2.

Next consider the results of tests of structural stability and lack of residual autocorrelation. A Lagrange multiplier test of this hypothesis against the alternative of up to fourth order residual autocorrelation and shifts in the parameters between 1956.1 to 1969.4 and 1970.1 to 1983.4 gives $F_{21,63} = 1.38$ (1.74) while the test of structural stability alone had given $F_{13,71} = 1.53$ (1.90). A structural stability or forecast test for 1980.1 to

---

*At its peak in 1974 the bias is about 2% on a base of 1970. This supports Grubb's (1984) argument that the effect posited by Bruno (1984) and Bruno and Sachs (1982) should not be exaggerated.

1983.4 compared with 1956.1 to 1979.4 gives $F_{16,69} = 1.33$ (1.73). A structural stability test for 1973.2 to 1980.2 compared with the pooled sample 1956.1 to 1973.1, 1980.3 to 1983.4 gives $F_{12,71} = 0.85$ (1.89). This is interesting because the intervening period is that of the two oil shocks.

Let us now turn to the question of whether this estimated production function suffers from bias because of the possible endogeneity of employment and overtime hours. The first step is to re-estimate R.2 by instrumental variables. To do this $(\ell n\ \ell_t + OH_t)$ and OH_t^{-1} are the two variables to be instrumented, though the instrumented value of OH_t^{-1} also enters in interaction with normal hours. The instrumenting equations were estimated for 1955.3 to 1969.4 and 1970.1 to 1983.4. The list of instruments includes lags of the following variables: $\ell n\ \ell$, OH, ℓn q, ℓn NH, PR, PW, the national vacancy rate, ℓn (world industrial production), ℓn (world exports of manufactures) and a real interest rate term. Relatively parsimonious forms of these equations with sensible long run values of the coefficients were selected. The results of thus instrumenting the production function as specified in R.2 are shown in R.4. The standard error of the equation increases from 0.007457 to 0.008305 and the parameter estimates are all very close to those in R.2. This provides informal support for the proposition that the endogeneity bias can be ignored.

The validity of the over-identifying restrictions entailed in the instrumentation is tested by comparing the likelihoods of the unrestricted and restricted reduced forms. This gives a chi-squared statistic of 19.4 With 30 degrees of freedom the critical value is 43.8 at the 5% level. To test the hypothesis of zero endogeneity bias the Revankar and Hartley (1973) test was used, this being also interpretable on the Lagrange multiplier test principle as discussed, e.g. by Engle (1982). Under the null hypothesis, the residuals from regressions of the potentially endogenous variables on instruments independent of the disturbances in R.2 should have zero coefficients when included as additional regressors in R.2. The resulting F-test gives

$F_{3,81}$ = 1.44 (2.73) where the variables and instruments are as described in connection with R.4. It seems, therefore, that we need not worry about endogeneity bias.

In contrast, omitting labour utilization as represented by the overtime variable OH_t from the regression produces all the symptoms of gross mis-specification. As can be seen from R.5, the residual standard error doubles, the Durbin Watson statistic falls to 0.87 and the elasticity of output w.r.t. employment is estimated at 1.70. Such absurd returns to labour are sometimes associated with "Verdoorn's Law". I would interpret this "Law" merely as a cyclical measurement error phenomenon: the result of omitting the utilization of labour.* As remarked in the introduction, there are standard cost of adjustment arguments to explain why output expands faster than employment in the upswing and contracts faster than employment in the downswing, thus giving rise to apparently large returns to labour. Further evidence of the mis-specification resulting from the omission of labour utilization can be found in the substantial alterations in many of the other coefficients in R.5 compared with R.2.

The non-linearity entering through OH_t^{-1} seems to be important. Omitting the OH_t^{-1} terms, unrestricting the OH_t coefficient and including an interaction between NH_t and OH_t raises the standard error to 0.008143 compared with R.2's 0.007457. On the other hand, (19) above suggested a more sophisticated non-linear specification. Estimating (19) by least squares gives a t-ratio for $c_1 - s_0$ of 1.0. Setting $c_1 - s_0 = 0$ gives $\hat{c}$ = 0.01226 (9.3), $\hat{s}_1$ = 0.002015 (6.2), where t-ratios are in parenthesis. But, although the other parameter estimates are very close, the standard error is slightly higher at 0.007799 compared with 0.007457 in R.2 which is meant to be an

---

*Verdoorn's Law says that the rate of growth of output per head is an increasing function of the rate of growth of employment. Chatterji and Wickens (1982) also provide evidence consistent with the interpretation of Verdoorn's Law as a cyclical phenomenon.

approximation to this specification. The Durbin-Watson statistic is 1.97 compared with 2.09 in R.2. This supports the more convenient R.2 specification.

(d) The Feldstein-Craine result

In more standard specifications of the production function, investigators sometimes define labour input as the product of the number of employees and the paid for hours per employee h^0.* Sometimes a test is carried out of the hypothesis that in a Cobb-Douglas context the elasticity of output w.r.t. hours is the same as that w.r.t. employment. As noted in the Introduction above, Feldstein (1967) argued that the hours elasticity should exceed the employment elasticity. Averaging his cross-section estimates for British manufacturing gives $\ell n\ q = \text{const.} + 0.773\ \ell n\ \ell + 2.046\ \ell n\ h^0 + .210\ \ell n\ K$. Craine (1973) examined Feldstein's hypothesis for time-series data on U.S. manufacturing. Imposing constant returns to scale he obtained for 1949.2 to 1967.4:

$$\underset{}{\ell n\ (q/K)} = \text{const.} + \underset{(23.2)}{0.007t} + \underset{(17.3)}{0.789}\ \ell n\ (\ell/K) + \underset{(14.8)}{2.177}\ \ell n\ h^0 \qquad (21)$$

s.e. = 0.012, DW = 0.87, d.f. = 71.

In the context of the current paper, average hours per operative h^0 is normal hours plus the difference between average overtime hours and average short time. But short-time is quantitatively unimportant being typically less than 10% of overtime. Thus we can approximate $\ell n\ h$ by $\ell n\ (\bar{h}\ (1 + \dot{O}H))$. Re-specifying R.2 in the Craine manner leads to the fourth to seventh terms in R.2, again estimating for 1956.1-83.4, being replaced by

$$\underset{(23.2)}{1.207}\ \ell n\ (\ell/K) + \underset{(14.3)}{2.176}\ \ell n\ (\bar{h}(1+\dot{O}H)) + \underset{(14.6)}{0.00934t} \qquad (22)$$

---

*This is observed average hours as conventionally understood rather than effective hours and will be referred to as 'average hours' or simply 'hours' in what follows.

and s.e. = 0.00967, DW = 1.47. All other variables are as in R.2. There is a remarkable similarity in the elasticities of output w.r.t. average hours from these three very different data sets.

My interpretation of these results, given the theory presented in Section 3 above, is in terms of the correlation between ℓn (average hours) and the overtime based utilization measure. However, since this correlation is disturbed by variations in normal hours, we ought to find that entering normal hours as an additional variable to remove this source of variation ought to improve the results. This indeed is what happens. Instead of (22), again for 1956.1-83.4 we now find

$$\underset{(7.9)}{0.697}\ \ell n\,(\ell/K) + \underset{(14.8)}{3.479}\ \ell n\,(\bar{h}(1+OH)) - \underset{(6.6)}{1.852}\ \ell n\,\bar{h} + \underset{(5.0)}{0.00449}t \qquad (23)$$

s.e. = 0.007910, DW = 2.14. The fit is much improved; there is now no sign of first order residual autocorrelation and the elasticity of output w.r.t. employment is much more satisfactory. The non-linearity in response implied by the theory suggests that adding a quadratic term in $\ell n\,\bar{h}(1+OH)$ might improve the fit further. The s.e. now is 0.00765, DW = 2.19 and the elasticity of output w.r.t. employment is estimated at 0.775 (8.5).

Equation (23) suggests that one can obtain results almost as good as R.2 without the effort of constructing a special overtime series. Instead, readily available average hours data can be used. However, if normal hours went through substantial changes, as they did in the 1950's and 1960's, it is essential that a normal hours variable be included.

Overall, these results strongly support the hypothesis that the Feldstein-Craine finding of an elasticity of output w.r.t. average hours considerable in excess of the elasticity w.r.t. employment is the result of the correlation between average hours and an omitted labour utilization variable. Leslie and Wise (1980), who call this the labour hoarding explanation, reject this hypothesis on the basis of a time-series/cross-section study on annual British data for 28 industries for 1948-1968. They find that the inclusion of industry specific dummies and trends in a pooled cross-section reduces the hours elasticity to a value close to the employment elasticity. Hence they argue that there is an upward bias in the hours elasticity caused by omitted industry-specific efficiency effects. This is a serious challenge to the interpretation of the Feldstein-Craine result given here and deserves comment.

One problem with the study is that the hours data, though used to explain annual output, are based on hours observed over only two weeks. Thus there is likely to be quite a serious random measurement error in these data which will bias downwards the hours coefficient. Note that the cross-section variation in hours is already being picked up by the industry specific coefficients and it is the measurement error in hours relatively to the cyclical variation in annual hours which matters. Secondly, a great deal of the variation in average hours over 1948-68 is due to the considerable reduction in normal hours in the 1950's and early 1960's. As we have seen, this tends to reduce the correlation between average hours and labour utilization.* In this respect, it is noteworthy that Leslie's (1984) study of annual data for 20 U.S. industries for 1948-76 shows the hours elasticities to be significantly higher than the employment elasticities despite the inclusion of both industry specific dummies and a Wharton capacity utilization index.

(e) Further aspects of labour input

Another aspect of hours is paid holiday entitlements. Unfortunately the Department of Employment does not publish information which relates to manufactur-

---

*My explanation of the Feldstein-Craine result thus predicts three conditions under each of which the hours elasticity would increase in a study of the Leslie and Wise type: firstly, use hours data which are more annually representative, secondly, include industry-specific normal hours as regressors or thirdly, estimate over a period in which there is little variation in normal hours.

ing but gives annual figures on holiday entitlements in all national agreements covering manual workers. Of these workers, the proportion in manufacturing is probably now a little under one half. Assuming that the national figures are representative of manufacturing permits the construction of the logarithmic weeks worked measure ww described in the data section above. This assumes 5 working days per week so that 6 public holidays per annum corresponds to 1.2 weeks and shows a fall of 6% from 1955 to 1983. Specifying labour input as $(\ln \ell_t \bar{h}_t + ww_t + u_t)$ gives an equation standard error of 0.007549 and the other coefficients virtually unaltered even for the trend coefficients. The largest changes for these are 8% reductions in the absolute size of the coefficients for TR79.3 and TR80.3 compared with R.2 and a 16% increase in that for TR59.4.

Another trend like change which has occurred is in the incidence of shift work and is no doubt reflected in the estimated trend effects. Before 1973 the only official figures are very sparse. These show a percentage of workers in manufacturing on some kind of shift work of 12.5% in 1954, 20% in 1964 and 24.9% in 1968, see National Board for Prices and Incomes (1970), p.65.*

Yet another aspect of labour input is in PO the proportion of employees who are operatives. Data on this are biannual from 1963 to 1974 and otherwise annual. Using a linear quarterly interpolation of PO, one might replace the labour input term in the production function (18) by

$$\alpha_1 \ln (\ell_t PO_t \bar{h}_t/K_t) + \alpha_2 \ln \{\ell_t(1-PO_t)\bar{h}_t/K_t\} + (\alpha_1+\alpha_2) u_t \qquad (24)$$

The assumption here is that the same utilization factor applies to non-operatives as to operatives. Although the equation s.e. improves to 0.007346, the estimated $\alpha_2 = 0.044$ (0.6) is implausibly low. Applying the utilization term to operatives only raises the equation standard error marginally. Given a share of operatives in the wage bill of roughly 0.7, it is interesting to test the hypothesis $\alpha_2/\alpha_1 = 3/7$. This is easily rejected, the t-test being 2.42. But clearly the

*An estimate by Fishwick (1980) for 1979 of 26% suggests a much slower rate of increase since 1968 and this is consistent with annual New Earnings Survey figures from 1973. Bosworth and Dawkins (1981) discuss comparability problems of these data.

implausible hypothesis that only the input of operatives matters is easily acceptable. Perhaps this is not altogether surprising, given the crudity of the data on the proportion of operatives but leaves one in a quandary over which definition of labour input to accept. Since the differences in fit are slight and the quarterly data on PO suspect, my own preference remains for R.2.

(f) The role of capital

Finally, the hypothesis of constant returns to scale was tested by adding $\ln K_t$ to the list of regressors in R.2. This produces a coefficient of -1.216 (3.2) and very little change in the other parameters except for the trends. The hypothesis that the capital stock plays no role in the model versus the unrestricted alternative is just rejected at $t = 2.33$ and gives an equation s.e. of 0.007258. These results should be taken not so much as a rejection of constant returns to scale but as a symptom of the measurement errors in the gross capital stock data, perhaps because unobserved scrapping is highest in downturns and lowest in upturns. An alternative hypothesis is that R.2 is cyclically mis-specified in some way: in cyclical behaviour, $\ln K$ lags a little over one year behind overtime hours and is negatively correlated with current overtime hours. However, one can easily accept the hypothesis of zero coefficients on $u_{t-1}, \ldots, u_{t-6}$ added as regressors to R.2 and, it should be recalled, for R.2 the Lagrange multiplier tests for residual autocorrelation are insignificant. It should also be noted that Craine (1973) too reports a better fit for U.S. manufacturing when capital is omitted from the production function.

The model selection question here is more one of the economic interpretation one wishes to put on the model than one of goodness of fit. If one wishes to interpret some of the trend shift effects as correction factors to apply to the observed capital stock, R.2 is clearly preferable to a specification where the capital stock is omitted entirely.

(g) The estimated equations

R.1(a): 1956.1-1969.4

$$\ell n(q_t/K_t) = \underset{(5.1)}{-7.113} + \underset{(5.0)}{0.752}\ (\ell n(\ell_t/K_t) + \ell n\ NH_t + OH_t) - \underset{(7.5)}{0.01318}\ OH_t^{-1}$$

$$+ \underset{(7.1)}{0.06062}\left(\frac{NH_t - 90.4}{100}\right) OH_t^{-1} + \underset{(3.6)}{0.00536}\ t + \underset{(2.9)}{0.00178}\ TR\ 59.4 + \underset{(2.5)}{0.0929}\ PWD_t$$

$$+ \underset{(1.5)}{0.0744}\ \Delta_4 PW_t + \underset{(1.9)}{0.0799}\ \Delta_3 PR_t + \underset{(1.0)}{0.0020}\ EX1_t + \underset{(2.2)}{0.0058}\ EX2_t + \text{terms in S1, S2,}$$

S3, S1TR and 2 (0, 1) dummies.

s.e. = 0.007447, SSE = 0.002163, $\bar{R}^2$ = 0.9453, DW = 2.38, n = 56, d.f. = 39

R.1(b): 1970.1-1983.4

$$\ell n(q_t/K_t) = \underset{(6.4)}{-7.444} + \underset{(6.4)}{0.768}\ (\ell n(\ell_t/K_t) + \ell n\ NH_t + OH_t) - \underset{(5.0)}{0.01035}\ OH_t^{-1}$$

$$+ \underset{(0.9)}{0.05998}\left(\frac{NH_t - 90.4}{100}\right) OH_t^{-1} + \underset{(4.8)}{0.00883}\ t - \underset{(5.2)}{0.00596}\ TR\ 73.1 - \underset{(4.2)}{0.00846}\ TR\ 79.3$$

$$+ \underset{(6.0)}{0.01403}\ TR\ 80.3 + \underset{(3.3)}{0.00434}\ PC_{t-2} - \underset{(3.4)}{0.0732}\ PRD_{t-3} + \underset{(2.1)}{0.0540}\ \Delta_4 PW_t$$

$$+ \underset{(2.6)}{0.0602}\ \Delta_3 PR_t + \underset{(2.3)}{0.0019}\ EX1_t + \underset{(1.0)}{0.0019}\ EX2_t + \text{terms in S1, S2, S3, SITR and}$$

6 (0, 1) dummies.

s.e. = 0.006817, SSE = 0.001487, $\bar{R}^2$ = 0.9974, DW = 1.82, n = 56, d.f. = 32

R.2 1956.1-1983.4

$$\ell n(q_t/K_t) = \underset{(8.3)}{-6.457} + \underset{(8.2)}{0.681}\ (\ell n\ \ell_t/K_t) + \ell n\ NH_t + OH_t) - \underset{(10.2)}{0.01207}\ OH_t^{-1}$$

$$+ \underset{(8.2)}{0.05140}\left(\frac{NH_t - 90.4}{100}\right) OH_t^{-1} + \underset{(4.2)}{0.0050}\ PC_{t-2} + \underset{(3.6)}{0.1015}\ PWD_t$$

$$- \underset{(3.1)}{0.0668}\ PRD_{t-3} + \underset{(2.7)}{0.0591}\ \Delta_4 PW_t + \underset{(3.7)}{0.0776}\ \Delta_3 PR_t + \underset{(5.2)}{0.00451}\ t$$

$$+ \underset{(4.0)}{0.00172}\ TR\ 59.4 - \underset{(8.2)}{0.00412}\ TR\ 73.1 - \underset{(4.4)}{0.00788}\ TR\ 79.3 + \underset{(6.0)}{0.01193}\ TR\ 80.3$$

$$+ \underset{(2.4)}{0.0040}\ EX1_t + \underset{(2.7)}{0.0035}\ EX2_t$$

+ terms in S1, S2, S3, SITR and 8 (0, 1) dummies.

s.e. = 0.007457, SSE = .004672, $\bar{R}^2$ = 0.9972, DW = 2.09, n = 112, d.f. = 84.

R.3 1956.1-1983.4

$$\ell n(q_t/K_t) = \underset{(6.9)}{-6.200} + \underset{(6.9)}{0.653}\ (\ell n(\ell_t/K_t) + \ell n\ NH_t + OH_t) - \underset{(9.0)}{0.01250}\ OH_t^{-1}$$

$$+ \underset{(7.4)}{0.05489}\left(\frac{NH_t - 90.4}{100}\right) OH_t^{-1} + \underset{(3.4)}{0.00472}\ PC_{t-2} + \underset{(2.3)}{0.0772}\ PWD_t - \underset{(2.2)}{0.0537}\ PRD_{t-3}$$

$$+ \underset{(2.4)}{0.0592}\ \Delta_4 PW_t + \underset{(3.3)}{0.0783}\ \Delta_3 PR_t + \underset{(4.1)}{0.00414}\ t + \underset{(3.4)}{0.00192}\ TR\ 59.4 - \underset{(7.8)}{0.00434}\ TR\ 73.1$$

$$- \underset{(3.5)}{0.00727}\ TR\ 79.3 + \underset{(4.7)}{0.01104}\ TR\ 80.3 + \underset{(4.1)}{0.0068}\ EX1_t + \underset{(2.1)}{0.0028}\ EX2_t + \text{terms in S1,}$$

S2, S3, SITR and 1972.1, 1974.1 dummies.

s.e. = 0.008863, SSE = 0.007070, $\bar{R}^2$ = 0.9960, DW = 1.83, n = 112, d.f. = 90.

R.4 1956.1-1983.4 estimation by instrumental variables

$$\ell n(q_t/K_t) = \underset{(7.5)}{-6.750} + \underset{(7.5)}{0.711}\ (\ell n(\ell_t/K_t) + \ell n\ NH_t + OH_t) - \underset{(8.1)}{0.01124}\ OH_t^{-1}$$

$$+ \underset{(5.6)}{0.04964}\left(\frac{NH_t - 90.4}{100}\right) OH_t^{-1} + \underset{(2.8)}{0.00466}\ PC_{t-2} + \underset{(1.9)}{0.0574}\ PWD_t - \underset{(3.5)}{0.0834}\ PRD_{t-3}$$

$$+ \underset{(2.8)}{0.0700}\ \Delta_4 PW_t + \underset{(3.6)}{0.0837}\ \Delta_3 PRP_t + \underset{(4.7)}{0.00463}\ t + \underset{(3.8)}{0.00203}\ TR\ 59.4 - \underset{(6.8)}{0.00372}\ TR\ 73.1$$

$$- \underset{(5.6)}{0.01077}\ TR\ 79.3 + \underset{(6.7)}{0.01484}\ TR\ 80.3 + \underset{(2.6)}{0.0047}\ EX1_t + \underset{(2.4)}{0.0034}\ EX2_t + \text{terms in}$$

S1, S2, S3, SITR and 8 dummies.

s.e. = 0.008305, SSE = 0.005800, $\bar{R}^2$ = 0.9954, DW = 2.02, n = 112, d.f. = 84.

R.5 1956.1-1983.4

$$\ell n(q_t/K_t) = \underset{(17.3)}{-15.909} + \underset{(16.9)}{1.702}\ (\ell n(\ell_t/K_t) + \ell n\ NH_t) + \underset{(1.3)}{0.00298}\ PC_{t-2} + \underset{(1.8)}{0.0846}\ PWD_t$$

$$- \underset{(1.2)}{0.0530}\ PRD_{t-3} + \underset{(5.3)}{0.2039}\ \Delta_4 PW_t + \underset{(2.5)}{0.1028}\ \Delta_3 PR_t + \underset{(11.5)}{0.01340}\ t + \underset{(5.5)}{0.00474}\ TR\ 59.4$$

$$- \underset{(7.0)}{.00527}\ TR73.1 - \underset{(3.7)}{0.01180}\ TR79.3 + \underset{(8.0)}{0.0270}\ TR80.3$$

$$+ \underset{(1.2)}{0.0040}\ EX1_t + \underset{(1.3)}{0.0033}\ EX2_{t-1} + \text{terms in S1, S2, S3, SITR and 8 (0, 1) dummies.}$$

s.e. = 0.01488, SSE = 0.01905, $\bar{R}^2$ = 0.9888, DW = 0.87, n = 112, d.f. = 86.

7. Conclusions

This paper has referred to three classic problems in estimating an aggregate production function on aggregate time series data. The first is the unobservables problem, particularly in regard to utilization and the measurement of capital and to a lesser extent the measurement of output and labour input. The second is the aggregation problem and the third the simultaneous equation bias problem. It was the hypothesis of this paper that the first of these problems is the most serious.

A new proposal for the measurement of labour input was put forward. This argued that high rates of labour utilization relative to a norm would be reflected in high aggregate weekly overtime hours being reported. However, since most workers are paid for a standard week even when being underutilized, there is no corresponding direct measure of below normal rates of labour utilization. This is a situation where we can observe the mean of the truncated upper tail of the distribution of utilization rates over firms. Given a reasonably constant spread the proposal was to derive an estimate of the mean of the whole distribution from the information about the truncated upper tail. This implies that the mean labour utilization rate is a non-linear function of overtime hours as a proportion of normal hours. With an additional allowance for a British institutional feature in which some of what is called overtime is in fact part of the usual work week for some employees, this measure of labour utilization gave excellent empirical results.

In particular, an elasticity of output w.r.t. employment of 0.681 is very reasonable in contrast to the huge elasticity estimated when labour utilization is omitted. Large elasticities are often interpreted as evidence of "Verdoorn's Law" which says that the rate of growth of output per head is an increasing function of the rate of growth of employment. On the current evidence, this "Law" is a cyclical measurement error phenomenon.

The theory put forward provides a convincing explanation of the Feldstein-Craine result that the elasticity of output w.r.t. average paid for hours of work substantially exceeds the elasticity w.r.t. employment. Since average paid for hours is approximately normal hours plus overtime hours averaged over operatives whether they work overtime or not, ℓn (average paid for hours) $\approx \ell n$ (normal hours) + OH , where OH is overtime hours averaged over all operatives and scaled by normal hours. According to the theory put forward, labour utilization as I have defined it, is a non-linear function of OH with a derivative greater than unity. Although the derivative of ℓn (output) w.r.t. labour utilization is the same as that w.r.t. ℓn (employment), the elasticity w.r.t. average paid for hours will therefore be greater. This gives the Feldstein-Craine result. Moreover, normal hours enters the relationship between average paid for hours and utilization. The theory therefore predicts and the empirical evidence agrees that for periods when normal hours are changing, normal hours should make a significant negative contribution in a production function in which ℓn (average paid for hours) enters as a regressor. An implication when normal hours is omitted is an unstable coefficient on ℓn (average paid for hours) for different samples containing different variations in normal hours.

The paper also considered biases in the measurement of capital and output. The former are likely to be large because the gross capital stock as usually measured assumes that assets have service lives and yield service flows which are invariant to changes in economic conditions. Although the device of proxying these measurement biases by shifts in the slopes of time trends is crude, the empirical evidence favoured the hypothesis of increased unrecorded scrapping and perhaps reduced service flows from 1973 with a particularly heavy incidence during the 1979.3 to 1980.3 period when British manufacturing experienced a crisis in which output fell by 16%. The finding, like Craine's (1973), that omitting the capital stock altogether from the production function improved the standard error, could be consistent with a badly measured concept. The hypothesis that a measure of capacity utilization based on a survey of firms' options on whether they were operating below full capacity contains no information additional to the overtime based labour utilization variable could be accepted. It is conceivable that this result is a consequence of the poor quality of the capital stock series which this concept might be adjusting. However, it is consistent with the widespread idea that the capital stock should not be utilization adjusted.

As far as output is concerned, four potential biases were considered and observable proxies constructed to measure their effects. The evidence is that for British manufacturing the 'gross output bias' that stems from the approximation in the construction of the output index of value added changes by gross output changes is small, probably about 2% at the peak of relative raw materials prices. The 'domestic price index' bias and the 'list price bias' cannot be fully disentangled empirically but have considerable short run significance, particularly when sharp changes in real exchange rates occur. The former comes from the CSO's approximation of unavailable export price deflators by domestic price deflators. The latter arises to the extent that the prices reported by firms are not transactions prices but list prices. Finally, there is strong evidence that, as Darby (1984) reports for the U.S., attempts by governments to control prices have a significant distortionary effect on price deflators and so on measured output as firms attempt to side-step these controls.

Appendix 1: an alternative measure of labour utilization

Apart from the overtime measure OH there are also data on the proportion p of operatives on overtime. If the distribution of $u = \ln h - \ln \bar{h}$ depends on two parameters, it is possible to deduce both parameters from OH and p. Suppose $u \sim N(\mu, \sigma^2)$. Then

$$OH = \int_0^\infty u\, f\left(\frac{u-\mu}{\sigma}\right) du, \quad p = \int_0^\infty f\left(\frac{u-\mu}{\sigma}\right) du = 1 - F\left(\frac{-\mu}{\sigma}\right)$$

As is well known (see Johnson and Kotz (1972), p.112-113),

$$\int_0^\infty u\, f\left(\frac{u-\mu}{\sigma}\right) du = \mu p + \sigma f\left(\frac{-\mu}{\sigma}\right) \quad (a.1)$$

Let $\frac{-\mu}{\sigma} = x$ where $x = F^{-1}(1-p)$. Then

$$OH = \mu p - \mu x^{-1} f(x) \quad (a.2)$$

so that $\mu = OH/(p - x^{-1} f(x))$ (a.3)

Thus we can derive an estimate of mean utilization μ from observing p and OH. Changes in systematic overtime as normal hours fell are allowed for by including an interaction term $(100 - NH)\mu$ which is zero in 1955.

A variation on R.2 is estimated in which the employment term is $\ln(\ell_t/K_t) + \ln NH_t$ and in which OH_t^{-1} and its interaction with normal hours are replaced by μ_t and $(100 - NH_t)\mu_t$. With $\hat{\alpha} = 0.684$ (6.6) and a coefficient on μ of 0.730 (9.0), the hypothesis that these two coefficients are equal is accepted. Imposing this restriction gives $\hat{\alpha} = 0.710$ (22.1) and a coefficient of -0.0158 (2.4) on $(100 - NH)\mu$. However, with s.e. = 0.008733, SSE = 0.006483, $\bar{R}^2 = 0.9962$, DW = 1.39, this equation is clearly inferior to R.2 even though the remaining parameter estimates are quite similar.

Given that both p and OH are subject to sampling variations it may be unrealistic to derive both μ and $\sigma = -\mu/x$ in this way. An obvious alternative is to treat σ as a constant. Then there are two alternative estimates of μ conditional on σ. The one based on OH, μ_{1t} is derived by solving the implicit function based on (a.1)

$$OH_t = \mu_{1t}\ (1-F(\frac{-\mu_{1t}}{\sigma})) + \sigma\ f(\frac{-\mu_{1t}}{\sigma})$$

The one based on p is

$$\mu_{2t} = -\ \sigma F^{-1}(1-p_t)$$

To investigate the empirical implications, R.2 was respecified with the employment term in the form $\ell n\ \ell_t/K_t + \ell n\ NH_t$ and OH_t^{-1} and its interaction with normal hours replaced by μ_{1t}, μ_{2t} and their respective interactions with normal hours. For a grid of values of σ, $\sigma = 0.7$ was found to give the lowest error sum of squares. Moreover, the hypothesis that the coefficients on μ_{2t} and its interaction with normal hours are zero could be easily accepted. Finally, the hypothesis that the coefficient on μ_{1t} is α could also be accepted. The resulting equation reads as follows:

R.6 1956.1-1983.4

$$\ell n\ (q_t/K_t) = \underset{(39.8)}{-6.284} + \underset{(37.8)}{0.682}\ (\ell n\ \ell_t/K_t + \ell n\ NH_t + \mu_{1t})$$

$$- \underset{(7.8)}{0.00921}\ (100-NH_t)\ \mu_{1t} + \underset{(4.4)}{0.00508}\ PC_{t-2} + \underset{(3.3)}{0.0947}\ PWD_t$$

$$- \underset{(3.4)}{0.0709}\ PRD_{t-3} + \underset{(3.4)}{0.0637}\ \Delta_4 PW_t + \underset{(3.3)}{0.0675}\ \Delta_3 PR_t + \underset{(10.6)}{0.00440}\ t$$

$$+ \underset{(4.8)}{0.00206}\ TR59.4 - \underset{(8.8)}{0.00429}\ TR73.1 - \underset{(5.2)}{0.0888}\ TR79.3 + \underset{(7.4)}{0.01341}\ TR80.3$$

$+ \underset{(2.5)}{0.00406}\ EX1_t + \underset{(2.7)}{0.00349}\ EX2_t$ + terms in S1, S2, S3, S1TR and 8 (0, 1) dummies.

s.e. = 0.007367, SSE = 0.004613, $\bar{R}^2$ = 0.9979, DW = 2.20, n = 112, d.f. = 85

The equation s.e. and the t-ratios are conditional upon $\sigma = 0.7$. Adjusting the equation s.e. as if the equation were linear in σ gives s.e. = 0.007410. This is very slightly better than R.2's s.e. of 0.007457. The similarity in the parameter estimates and fit of R.6 and R.2 is remarkable and reassuring for the robustness of R.2. There is virtually nothing to choose between them except convenience. Both are easily accepted against the maintained hypothesis which nests them both. R.6 requires the solution of a non-linear implicit function to generate estimates of average utilization and estimation for a grid of values of σ. R.2 can be estimated by OLS and linearity in the parameters has other convenient properties such as the ease of estimation by instrumental variables techniques.

Finally, it is worth noting that there are signs that the true underlying distribution of u is positively skewed. The value of μ from R.6 ranges from about −0.66 to −0.92. This is implausibly negative which suggests that the Normal distribution contains more mass in its lower tail than is plausible. Another way of seeing this is to note that using $p = 1 - F(\frac{-\mu}{\sigma})$, the values of p implied by these values of μ/σ turn out to be smaller than the observed values. In terms of (13), (14) or (19), positive skewness implies $c_2 > c_1$, though in fact c_2 cannot be separately identified. This is another piece of evidence which favours the representation of mean utilization in R.2 as an accurate as well as convenient approximation to the true relationship of mean utilization with OH_t.

Appendix 2: A method for measuring capacity utilization from surveys

In the CBI Industrial Trends Survey as in similar surveys, the question is asked "Is your present level of output below capacity?" (defined as a satisfactory or full rate of operation). Let q = output, $q(max)$ = maximum output. Let $-u_c = \ell n\ q(max) - \ell n\ q$ so that u_c is a proportionate deviation measure of capacity utilization. Suppose that different firms share the same view of what constitutes a satisfactory level of operation. It might, for example, be 80% of the maximum. Let

$$z = \ell n\ q(max) - \ell n\ q(satisfactory) \tag{a.5}$$

Now suppose that there is a distribution across firms of capacity utilization measured by $\ell n\ q(max) - \ell n\ q$ as illustrated in Figure 3. Imagine the distribution and so its mean, which measures the average degree of utilization that we want,

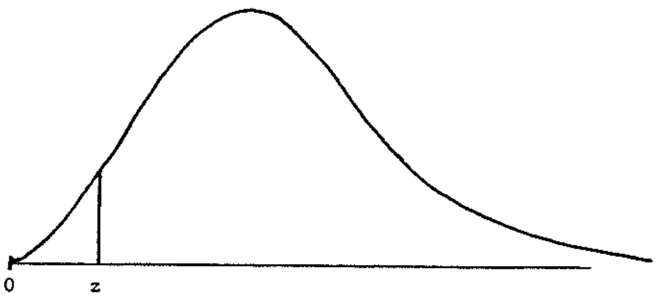

Figure 3: the distribution of (minus) capacity utilization across firms

shifting through time (though always with its lower limit fixed at zero). We will observe corresponding shifts in π, the proportion of firms (weighted by size) to the right of z.

The aim is to make deductions about the unobserved mean of the distribution from observations on π.* An obvious candidate for the distribution in Figure 3 is the lognormal, i.e. $\ell n\,(-u_c)$ is Normal with mean μ and variance σ^2. Then

$$\pi = 1 - F\left(\frac{\ell n\ z - \mu}{\sigma}\right) \tag{a.6}$$

where F( ) is the distribution function of the standard Normal distribution.

Then $\mu = \ell n\ z - \sigma\, F^{-1}\,(1-\pi)$ (a.7)

μ measures $E(\ell n(-u_c))$ but we want $E(-u_c)$. When $\ell n(-u_c)$ has a normal distribution,

$$\begin{aligned} E(-u_c) &= \exp \tfrac{1}{2}\sigma^2 .\ \exp \mu \\ &= \exp \tfrac{1}{2}\sigma^2\ \exp\{\ell n\ z - \sigma\, F^{-1}(1-\pi)\} \\ &= \text{const.}\ \exp\{-\sigma\, F^{-1}(1-\pi)\} \end{aligned} \tag{a.8}$$

where the constant of proportionality is $z \exp \tfrac{1}{2}\sigma^2$.

The Normal distribution can be quite well approximated by the logistic. As Amemiya (1981), p.1487 points out, if x is a standard Normal variable,

$$F(x) \approx \frac{e^{1.6x}}{1+e^{1.6x}} \tag{a.9}$$

where the RHS is the distribution function of a logistic variable with mean zero and variance $1.6^2/(\pi^2/3)$. Then (a.8) becomes

$$E(-u_c) = \text{const.}\ (\pi/1-\pi)^{\sigma/1.6}\ .$$

---

*Theil (1966) has one of the earlier discussions of this type of problem in an applied context.

Let us illustrate the empirical magnitudes this can take for some plausible parameter values. Suppose $z = 0.2$ which corresponds to satisfactory output being 82% of the maximum. Further, suppose that $\sigma = 0.4$. Then for the minimum observed value of $\pi = 0.38$, $E(-u_c) = 1.083 \times 0.2 \times 0.8848 = 0.192$. For the maximum observed value of $\pi = 0.84$, $E(-u_c) = 1.083 \times 0.2 \times 1.4888 = 0.323$. This would correspond to a range of variation of 0.136 in ℓn (output) between the highest and lowest observed degrees of capacity utilization and this is not an unreasonable range. The logistic approximation is very close, also giving 0.192 for $\pi = 0.38$ and 0.323 instead of 0.328 for $\pi = 0.84$.

Some experiments in which the production function was estimated without any overtime variables but with $CU = (\pi/1-\pi)^{\theta}$ for a range of $\theta = 0.2$ to 0.6 suggested the best fit at $\theta \approx 0.4$ which implies $\sigma = 0.64$ with the corresponding value of z estimated at 0.090. This suggests that 'full capacity' is about 91% of the physical maximum. However, with a standard error of 0·0113 and a Durbin-Watson statistic of 1.42, this is a much less satisfactory equation than R.2.

REFERENCES

Amemiya, T., (1981), "Qualitative response models: a survey", Journal of Economic Literature, 19, 1483-1536.

Baily, M., (1981), "Productivity and the services of capital and labour", Brookings Papers on Economic Activity, no.1, 1-65.

Baily, M., (1982), "The productivity growth slowdown by industry", Brookings Papers on Economic Activity, no.2, 423-453.

Bosworth, D. and P. Dawkins, (1981), Work Patterns: An Economic Analysis, Aldershot: Gower.

Bruno, M., (1984), "Raw materials, profits and the productivity slowdown", Quarterly Journal of Economics, 99, 1-30.

Bruno, M. and J. Sachs, (1982), "Input price shocks and the slowdown in economic growth: the case of U.K. manufacturing", Review of Economic Studies, 49, 679-706.

Central Statistical Office (CSO), (1959), The Index of Industial Production: Method of Compilation, Studies in Official Statistics no. 7, London: HMSO.

Central Statistical Office (CSO), (1970), The Index of Industrial Production and Other Output Measures, Studies in Official Statistics no.17, London: HMSO.

Central Statistical Office (CSO), (1976), The Measurement of Changes in Production, Studies in Official Statistics no.25, London: HMSO.

Central Statistical Office (CSO), (1983), Industry Statistics, Occasional Paper no.19, London: HMSO.

Chatterji, M. and M. Wickens, (1982), "Productivity, factor transfers and economic growth in the U.K.", Economica, 49, 21-38.

Craine, R., (1973), "On the service flow from labour", Review of Economic Studies, 40, 39-46.

Darby, M., (1984), "The U.S. productivity slowdown: a case of statistical myopia", American Economic Review, 73, 301-322.

Deaton, A. and J. Muellbauer, (1980), Economics and Consumer Behaviour, New York: Cambridge University Press.

Denison, E., (1969), "Some major issues in productivity analysis: an examination of estimates by Jorgenson and Griliches", Survey of Current Business, part II, 49, 1-27.

Denison, E., (1974), Accounting for U.S. Economic Growth, 1929 to 1969, Washington, D.C.: The Brookings Institution.

Denison, E., (1979), Accounting for Slower Growth: The U.S. in the 1970's, Washington, D.C.: The Brookings Institution.

Divisia, F., (1952), Exposes d'economique, vol.I, Paris, Dunod.

Engle, R., (1982), "A general approach to Lagrange multiplier model diagnostics", Journal of Econometrics, 20, 83-104.

Feldstein, M., (1967), "Specification of the labour input in the aggregate production function", Review of Economic Studies, 34, 375-86.

Fishwick, F., (1980), The Introduction and Extension of Shiftworking, London: National Economic Development Council.

Gollop, F. and D. Jorgenson, (1980), "U.S. productivity growth by industry, 1947-73", in New Developments in Productivity Measurement and Analysis, NBER Studies in Income and Wealth, Vol.44, (ed.) J. Kendrick and B. Vaccara, Chicago: University of Chicago Press.

Griffin, T., (1976), "The stock of fixed assets in the U.K.: how to make the best use of the statistics", Economic Trends, October, HMSO.

Grubb, D., (1984), "Raw materials and the productivity slowdown: some doubts", Quarterly Journal of Economics, forthcoming.

Hansen, B. D., (1970), "Excess demand, unemployment, vacancies and wages", Quarterly Journal of Economics, 84, 1-23.

Heathfield, D., (1972), "The measurement of capital utilization using electricity consumption data for the U.K,", Journal of the Royal Statistical Society, Series A.

Heathfield, D., (1983), "Productivity in the U.K. engineering industry 1956-1976", mimeo, University of Southamption.

Johnson, N. and S. Kotz, (1972), Distributions in Statistics: Continuous Multivariate Distributions, New York: J. Wiley.

Jorgenson, D. and Z. Griliches, (1967), "The explanation of productivity growth", Review of Economic Studies, 34, 249-283.

Kendrick, J., (1973), Postwar Productivity Trends in the U.S., 1948-1969, New York: NBER.

Leslie, D. and J. Wise, (1980), "The productivity of hours in U.K. manufactuing and production industries", Economic Journal, 90, 74-84.

Leslie, D., (1984), "The productivity of hours in U.S. manufacturing industries", Review of Economics and Statistics, 66, .

Malcolmson, J. and M. Prior, (1979), "The estimation of a vintage model of production for U.K. manufacturing", Review of Economic Studies, 46, 719-736.

Malcolmson, J., (1980), "The measurement of labour cost in empirical models of production and employment", Review of Economics and Statistics, 62, 521-528.

Mendis, L. and J. Muellbauer, (1984), "British manufacturing productivity 1955-1983: measurement problems, oil shocks and Thatcher effects", Centre for Economic Policy Research Discussion Paper.

Mizon, G. and S. Nickell, (1983), "Vintage production models of U.K. manufacturing industry", Scandinavian Journal of Economics, 85, 295-310.

National Board for Prices and Incomes, (1970), Report no.161: Hours of Work, Overtime and Shiftworking, London: HMSO.

Price, R., (1977), "The CBI Industrial Trends Survey - an insight into answering practices ", CBI Review, Summer issue.

Raasche, R. A. and J. A. Tatom, (1981), "Energy price shocks, aggregate supply and monetary policy - the theory and the international evidence", in K. Brunner and A. Melzer (eds.), Supply Shocks, Incentives and National Wealth, Carnegie-Rochester Conference on Public Policy, vol.14, 9-93.

Revankar, N. S. and M.J. Hartley, (1973), "An independance test and conditional unbiased predictions in the context of simultaneous equation systems", International Economic Review, 14, 625-631.

Scott, M.FG., (1976), "Investment and growth", Oxford Economic Papers, 28, 317-63.

Scott, M.FG., (1981), "The contribution of investment to growth", Scottish Journal of Political Economy, 28, 211-116.

Theil, H., (1954), Linear Aggregation of Economic Relations, Amsterdam: North Holland.

Theil, H., (1966), Applied Economic Forecasting, Amsterdam: North Holland.

Triplett, J., (1983), "Concepts of quality in input and output price measures: a resolution of the user-value resource-cost debate", in M. Foss (ed), The U.S. National Income and Product Accounts: Selected Topics, NBER Conference on Research in Income and Wealth, vol.47.