NBER WORKING PAPER SERIES

PROCYCLICAL PRODUCTIVITY: INCREASING RETURNS OR CYCLICAL UTILIZATION?

Susanto Basu

Working Paper 5336

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
November 1995

I would like to thank, without implicating, Robert Barsky, Chris Carroll, John Fernald, Miles Kimball, and Matthew Shapiro for helpful discussions, and the editor, Olivier Blanchard, and two anonymous referees (one of them Robert Hall) for suggestions. Seminar participants at the Federal Reserve Board, the University of Chicago, the University of Michigan, Northwestern University and the NBER provided useful comments. I am grateful to the National Science Foundation for financial support. This paper is part of NBER's research programs in Economic Fluctuations and Monetary Economics. Any opinions expressed are those of the author and not those of the National Bureau of Economic Research.

NBER Working Paper 5336
November 1995

PROCYCLICAL PRODUCTIVITY: INCREASING RETURNS OR CYCLICAL UTILIZATION?

ABSTRACT

It has long been argued that cyclical fluctuations in labor and capital utilization and the presence of overhead labor and capital are important for explaining procyclical productivity. Here I present two simple and direct tests of these hypotheses, and a way of measuring the relative importance of these two explanations. The intuition behind the paper is that materials input is likely to be measured with less cyclical error than labor and capital input, and materials are likely to be used in strict proportion to value added. In that case, materials growth provides a good measure of the unobserved changes in capital and labor input. Using this measure, I find that the true growth of variable labor and capital inputs is, on average, almost twice the measured change in the capital stock or labor hours. More than half of that is caused by the presence of overhead inputs in production; the rest is due to cyclical factor utilization.

Susanto Basu
Department of Economics
University of Michigan
611 Tappan Street
Ann Arbor, MI 48109-1220
and NBER

Why is productivity procyclical? The answer to this question sheds light on the relative merits of different models of business cycles. The literature suggests three main explanations. First, measured fluctuations in productivity might reflect exogenous changes in production technology. Second, productivity (appropriately measured) may be procyclical because of increasing returns to scale: in this case the economy endogenously becomes more efficient by moving to higher levels of activity. Third, if inputs are systematically mismeasured, *measured* productivity may be procyclical even if true productivity does not change. The gap between actual and measured productivity most likely comes from cyclical errors in measuring inputs: unobserved changes in capital utilization or in the intensity of work effort.

This paper represents an effort to evaluate the merits of these three competing hypotheses. Of course they are not mutually exclusive, but from the standpoint of business-cycle theory it is important to know how much of the variation in productivity is due to each of the three sources we mention. Researchers have long been interested in distinguishing between the first, "supply-side," explanation for cyclical productivity and the other two "demand-side" hypotheses (e.g. Hall [1988], Evans [1992]). But it is equally interesting to distinguish between increasing returns and cyclical utilization, since these hypotheses have radically different implications for modeling business cycles.[1] For example, recent papers that incorporate substantial increasing returns and large price-cost markups show that these features introduce a menagerie of possibilities, including multiple equilibria in which sunspots or monetary non-neutralities drive business cycles.[2] Variable-utilization models have no such implications. But for increasing-returns models to generate sunspot equilibria, the degree of returns to scale must be large.[3] One objective of this paper is to obtain an accurate estimate of the degree of returns to scale in order to judge the plausibility of these new business-cycle models.

This paper presents a new method for detecting changes in the utilization of capital and labor that are not directly observed, in a framework that nests increasing returns and technology shocks as alternative possibilities. Its central insight is that firms may extract unmeasured services from their own capital stocks or from workers with whom they have a long-term relationship, but that in order to produce greater output, they need more materials input. Materials use is a convenient indicator of cyclical factor utilization because its input does not have an extra effort or intensity dimension.[4]

By this logic, changes in the input of materials relative to measured capital and labor are an index of unmeasured capital and labor input, so we can estimate the degree to which the procyclicality of productivity is driven by variable utilization. Using this proxy to control for cyclical utilization also allows us to estimate the degree of returns to scale accurately. A major criticism of previous estimates of returns to scale (e.g. Hall [1990]) is that such estimates are biased upward by unobserved factor utilization. Finally, once we account for changes in productivity coming from increasing returns and variable utilization, the residual should represent true changes in production technology. We can investigate the properties of this technology series along dimensions especially relevant for evaluating real-business-cycle models: the variance of technology shocks, and their covariance with output changes.

If cyclical productivity is largely due to unobserved utilization of capital and labor, then one would expect materials input to be more procyclical than measured inputs of the other two factors. This is clearly the case in postwar U. S. manufacturing. Figure I shows annual growth rates of output, materials, and a measure of combined capital and labor input (the growth rates of capital and labor weighted by the relative shares of each input) for aggregate U. S. manufacturing from 1949 to 1985.[5] Clearly materials input is much better synchronized with output than are measured capital and labor inputs. (Ignoring capital and simply studying the pattern of output, materials, and labor input also gives similar conclusions: materials input tracks output much more closely than does labor input.) So the first look at the evidence is consistent with the cyclical utilization hypothesis.

Exploiting the basic insight of the paper, the initial results suggest that cyclical factor utilization is very important for explaining procyclical productivity. Controlling for cyclical utilization, however, there is no evidence of increasing returns to scale in production — indeed, the estimates indicate that returns to scale are strongly diminishing. The implied technology series has a somewhat lower variance and a substantially smaller correlation with output than the standard Solow residual.

I next consider a variety of modifications to the model and a number of sources of error that might be responsible for these conclusions. First I allow for substitution between materials and capital and labor in the production of output, and for the possibility that energy input is not separable from capital input. Neither modification makes a significant difference to the results. Then I discuss whether the results might be driven by cyclical errors in measuring the relative price of materials. Again I argue that this is unlikely to be a significant source of error.

The last modification is to allow for non-homotheticity of the production function. The initial setup assumes that the production function is homogeneous of some degree γ, which is then the degree of returns to scale. However, it is possible that returns to scale are concentrated in the production of value added from capital and labor input. In that case, materials input may grow faster than capital and labor input due to a feature of the production technology, not because of variable factor utilization. I investigate two methods that may allow us to disentangle non-homotheticity from variable utilization, although both are likely to be biased towards finding increasing returns to scale. With this modification, I find that overall returns to scale are about constant, not diminishing as the earlier results suggested. Consequently, variable factor utilization, although it is significant, seems less important than the first set of results suggest. Nevertheless, even the final results suggest that a one percent change in the capital stock or labor hours leads, on average, to a 0.17 percent change in capital and labor utilization. The estimated properties of technology change are unaffected by considering this modified production function.

The results are still a marked contrast to much of the recent conventional wisdom on the degree of returns to scale and markups. I find that returns to scale are about constant, far from the large increasing returns found in earlier empirical work and used to calibrate recent business-cycle models. Constant returns to scale imply that markups of price over marginal cost must also be small: otherwise we would see large pure profits, which we do not observe. Thus, neither increasing returns nor imperfect competition appears to be a major source of procyclical productivity. Instead, cyclical productivity appears to be driven by a combination of variable factor utilization and technology shocks.

The paper is organized in seven sections. Section I presents the method used to measure variable utilization, under the assumption that the production function is homothetic. Section II discusses the data and Section III presents results. Section IV modifies the method of the previous analysis to make it robust to a frequently-cited type of non-separability in the gross-output production function. Section V investigates whether the results are being driven by measurement error in prices. Section VI modifies the method of Section I to allow for non-homotheticity. Section VII concludes.

I. METHOD

I start from a production function for gross output, which includes intermediate goods as one of the factors of production:

$$Y_i = F(Z_i K_i, C_i L_i, M_i, T_i). \tag{1}$$

Here Y is gross output (not value added). K and L are primary inputs of capital and labor, while M is the quantity of materials input. Z and C are the levels of capital and labor utilization, which are chosen by the firm but unobserved by the econometrician. T is an index of technology. I assume initially that the production function is homogeneous of degree γ (and hence homothetic) in capital, labor, and intermediate goods. F is not constrained to have constant returns to scale, so γ is a free parameter. I omit time subscripts for clarity; i indexes productive units.

Hall [1990] shows how one can use first-order conditions from cost-minimization to write the total differential of (1) as the "cost-based Solow residual." Replacing the infinitesimal differentials in Hall's derivation with finite differences, and using lower-case letters to represent logs, we find:

$$\begin{aligned} \Delta y_i &= \gamma\left[\alpha_i^k(\Delta k_i + \Delta z_i + \alpha_i^l(\Delta l_i + \Delta c_i) + (1-\alpha_i^k - \alpha_i^l)\Delta m_i\right] + \Delta t_i \\ &= \gamma \Delta x_i + \Delta u_i + \Delta t_i, \end{aligned} \tag{2}$$

where $\Delta x_i \equiv \left[\alpha_i^k \Delta k_i + \alpha_i^l \Delta l_i + (1-\alpha_i^k - \alpha_i^l)\Delta m_i\right]$, $\Delta u_i \equiv \left[\alpha_i^k \Delta z_i + \alpha_i^l \Delta c_i\right]$, and α_i^l and α_i^k are the shares of labor and capital in total <u>costs</u> (not revenue). Δx is a cost-weighted sum of the growth rates of the *observed* inputs. Δu is a cost-weighted sum of the growth rates of *unobserved* capital and labor utilization. If there are constant returns to scale and perfect competition — so that the cost shares are also revenue shares and $\gamma = 1$ — then (2) is just the defining equation for the Solow residual. Note, however, that unless Δu is identically zero the Solow residual measures both changes in utilization and changes in technology. This is a problematic issue for the large literature on real business cycles, which takes the Solow residual as a measure of technological change alone (e.g. Plosser [1989]).

Hall [1990] interprets the coefficient of a regression of Δy on Δx as an estimate of the degree of returns to scale.[6] Hall estimates the regression by instrumental variables, where the instruments are chosen to be uncorrelated with Δt. However, the resulting estimate of γ is not a consistent measure of returns to scale unless Δu is identically zero or is uncorrelated with the instruments. Neither condition is plausible. Much anecdotal evidence, plus work on shifts by Shapiro [1994], indicates that true inputs of capital and labor are much more cyclical than the observed changes in the capital stock and labor hours. And Hall's aggregate-demand instruments are likely to be uncorrelated with Δt, but not with Δu; plausible models of variable utilization predict that increases in demand will typically increase factor utilization.[7] Thus, estimates of returns to scale, γ, from an equation like (2) are likely to be biased upward.

The intuition behind this paper is that the available data on materials input make it possible to test for the existence of cyclical utilization of capital and labor and to measure their extent. In the process, we can also obtain a consistent estimate of returns to scale. The idea is a simple one: workers putting in longer hours and more effort, or machines being worked extra shifts, need more materials in order to create more output. Materials use is a convenient indicator of cyclical factor utilization because its input does not have an extra effort or time dimension. An hour worked may represent very different amounts of labor input and a machine may be operated at different intensities, but a nail, a sheet of steel, or a piece of lumber always makes the same contribution to output: no amount of coaxing can make one nut fit on two bolts.

In order to derive the relationship between the true (unmeasured) inputs of capital and labor and the measured input of materials, I put some additional structure on the production function. I assume that gross output is produced using a combination of value added and materials, where value added is produced by capital and labor:

(3) $$Y_i = T_i G\big(V(Z_i K_i, C_i L_i), H(M_i)\big).$$

Since F was assumed to be homothetic we can, without loss of generality, assume that V and H have constant returns to scale. Log-linearizing (3) and using the first-order condition for cost-minimization, we find that the growth rate of value-added equals the growth rate of materials, plus a substitution term that depends on the growth rate of the relative price of value added to materials and on the elasticity of substitution between the two inputs, σ:

(4)
$$\Delta v_i = \Delta m_i - \sigma\left(\Delta p_i^v - \Delta p_i^m\right).$$

Δp_i^v and Δp_i^m are the changes in the prices of value added and materials.[8] $\sigma \geq 0$ is the (local) elasticity of substitution between value-added and materials, and is a potentially important parameter. Two reference values for σ are the plausible Leontief case where materials are used in strict proportion to value added ($\sigma = 0$), and the Cobb-Douglas case with unit elasticity of substitution ($\sigma = 1$).

We can, in turn, express the Divisia index of value added, Δv, in terms of changes in observed capital and labor input and changes in unobserved utilization:

(5)
$$\Delta v_i = \frac{\alpha_i^k(\Delta k_i + \Delta z_i) + \alpha_i^l(\Delta l_i + \Delta c_i)}{\alpha_i^k + \alpha_i^l},$$

which implies

(6)
$$\alpha_i^k(\Delta k_i + \Delta z_i) + \alpha_i^l(\Delta l_i + \Delta c_i) = \left(\alpha_i^k + \alpha_i^l\right)\Delta m_i - \sigma\left(\alpha_i^k + \alpha_i^l\right)\left(\Delta p_i^v - \Delta p_i^m\right).$$

Substituting equation (6) into equation (2) gives

(7)
$$\Delta y_i = \gamma\left[\Delta m_i - \sigma\left(\alpha_i^l + \alpha_i^k\right)\left(\Delta p_i^v - \Delta p_i^m\right)\right] + \Delta t_i.$$

Note that we have eliminated both the observed and unobserved inputs of capital and labor: Δu does not appear in (7). In the absence of variable capital and labor utilization, estimates of (2) and (7) should yield the same estimate of γ. If, however, there is variable utilization, then the estimate of returns to scale from (7) should be significantly smaller than that from (2), which is biased upward by changes in unmeasured capital and labor inputs, Δu.

The test is easiest to perform in the Leontief case: if $\sigma = 0$ the growth of gross output is just proportional to the growth of materials input, with the constant of proportionality given by the degree of returns to scale. However, equation (7) shows that if the production function allows materials to be substituted for value added, we must also include changes in the relative price of value added to materials. In principle these prices are observed, and can be included in an estimate of (7).[9]

Suppose, however, that we do not wish simply to test the hypothesis that there is cyclical factor usage, but also to measure its extent. It is intuitively plausible that the average extent of variable utilization can be inferred by comparing the sizes of the estimated γs from equations (2) and (7). This is essentially correct. The estimate of γ from (2) divided by the estimated γ from (7) gives a consistent estimate of (one plus) the elasticity of utilization with respect to share-weighted changes in inputs. Thus, this measure gives the average growth in unobserved inputs of capital and labor as a function of the growth in *total* input use.

However, since we hypothesize that the cyclical variation in utilization is confined to capital and labor, we would prefer to express the change in utilization as a function only of the change in capital and labor input. So my simple parameterization of cyclical factor utilization (essentially cyclical utilization *qua* measurement error) is that the change in utilization of capital and labor is proportional to the growth rate of the observed input:

$$\Delta z = \chi \Delta k \text{ and } \Delta c = \chi \Delta l,$$

where $\chi \geq 0$.[10,11] Under this relationship between true and observed inputs, we can show that the probability limit of γ from estimating (2) (denoted γ^e) exceeds the true γ even when the endogeneity of inputs in response to technology shocks (the "transmission problem") is solved by instrumenting:

(8)
$$\text{plim}(\gamma^e) = \gamma\left[1 + \chi \frac{\alpha^v\left(1 + (1+\chi)(1-\alpha^v)\right)}{\left(\alpha^v + (1+\chi)(1-\alpha^v)\right)^2}\right].$$

$\alpha^v \equiv \alpha^k + \alpha^l$ is the share of value added in total cost and is taken to be a constant. Intuitively, equation (8) shows that the extent of the bias depends on the average degree to which the true input of value added exceeds the measured input, χ, and the use of value added in production, α^v. If $\chi = 0$ or if $\alpha^v = 0$, then there is no bias. As argued above, we obtain a consistent estimate of the true γ by estimating equation (7), and of course the estimate of equation (2) provides the baseline value for the biased γ. Inserting the estimates of γ from (2) and (7) into (8) allows us to infer the average degree of cyclical utilization, as measured by χ.

Finally, the residuals from equation (7) should give us a series of true technology shocks, purged of the effects of imperfect competition, increasing returns, and variable utilization. It will be interesting to compare the properties of this series to those of the standard Solow residual, especially along two dimensions that are critical for assessing real-business-cycle models. The first is the variance of the innovations to technology. Standard real-business-cycle models typically require that this variance be quite large. The second statistic is the correlation between technology shocks and output. Real-business-cycle models typically predict that this correlation is large as well.

II. DATA

I use unpublished data on industry-level inputs and outputs provided by Dale Jorgenson. The data consist of a panel of U. S. manufacturing industries, at approximately the 2-digit S.I.C. level, for the years 1953-1984.[12] In the following paragraphs I highlight a few key features of the data. For a complete description, see Jorgenson, Gollop and Fraumeni [1987].

Output is measured as gross output, and inputs are separated into capital, labor, energy, and materials. In creating a series for labor input, Jorgenson, Gollop and Fraumeni [1987] assume that wages are proportional to marginal products. This allows them, in essence, to calculate quality-adjusted labor input by weighting the hours worked by different types of workers (distinguished by various demographic and occupational characteristics) by their relative wage rates.[13] Hence, labor input can increase either because the number of hours worked increases, or because the "quality" of those hours increases. This quality-adjustment also implicitly adjusts the wage series for the labor force composition effect stressed by Solon, Barsky and Parker [1994]. Similarly, Jorgenson, Gollop and Fraumeni [1987] adjust inputs of capital and intermediate goods for changes in quality.[14]

The Jorgenson data report the quantity of materials purchased per year, not actual materials usage. Since, as Ramey [1989] documents, materials and work-in-process inventories are procyclical, substituting materials purchases for usage can bias downward the γ estimated from equation (7). Thus I have modified the Jorgenson data by adjusting the materials data for changes in inventories.[15]

To estimate required payments to capital, I follow Hall and Jorgenson [1967], Hall [1990], and Caballero and Lyons [1992], and compute a series for the user cost of capital, r. The required payment for any type of capital is then rP_KK, where P_KK is the current-dollar value of the stock of this type of capital. Basu and Fernald [1994b] describe the procedure for constructing the rate of return. Given required payments to capital, calculating the cost shares is straightforward.

The price series for composite capital and labor input is constructed as a Divisia index from the underlying quality-adjusted wage series and the series on required payments to capital, where the weights are the cost shares of capital and labor in value added. The aggregate materials price is similarly constructed from the underlying energy and non-energy materials price series.

Since input use is likely to be correlated with technology shocks, I seek demand-side instruments for input use. To solve the "transmission problem" of endogeneity between productivity shocks and input growth, I use the instruments advocated by Ramey [1989] and Hall [1988], as modified by Caballero and Lyons [1992]: the growth rate of the price of oil deflated by the price of manufacturing durables; the growth rate of the price of oil deflated by the price of manufacturing nondurables; the growth rate of real government defence spending; and the political party of the President.

The change in the price of oil is a valid instrument only if technological progress is Hicks-neutral. However, Jorgenson, Gollop and Fraumeni [1987] present evidence that technological progress is energy-biased, making the oil price instrument the most dubious of the three. Unfortunately, it also has the most explanatory power. Thus, I typically assume that the change to a different energy intensity following an oil price shock requires more than one year to effect, which makes the current oil price change a valid instrument. However, I report the main results, in Tables I-III using both current and lagged oil price changes, and the results barely change depending on which instrument set is used.[16]

III. RESULTS.

First I estimate equation (2) as it stands, using the measured values of all three inputs.[17] I use SUR to estimate the equations for all the industries together, imposing the constraint that the degree of returns to scale, γ, is equal across industries. The resulting estimates of returns to scale are shown in Table I. The estimates are 1.09-1.10, depending on the choice of instrument sets, and indicating statistically significant evidence of increasing returns. The estimate is also quite large economically, though much smaller than those reported by Hall [1990]. Basu and Fernald [1994a] show that Hall's use of value-added rather than gross-output data creates a specification error that biases upward his estimates of γ.[18]

Next I perform the simplest test of the cyclical utilization hypothesis. This corresponds to the case of a homogeneous production function with Leontief technology in the production of gross output, so that output growth is just proportional to materials growth. The results are reported in the first column of Table II. Here the results are very different. The estimate of γ ranges from 0.79 to 0.84, depending on the instrument set used. Clearly these are significantly different, both statistically and economically, from the estimates in Table I. We can decisively reject the null hypothesis that true capital and labor inputs are accurately measured by changes in the capital stock and labor hours. The calculated values of χ, reported in the lines below each set of estimates, show that the true growth of capital and labor input is, on average, about 70 percent greater than the measured growth.

Does allowing for the possibility that materials can be substituted for value added change the conclusion that the unobserved inputs of labor and capital are very large? To see how the estimate of χ changes when we allow for the possibility that firms substitute towards using more materials in a boom, I now allow for substitution between materials and value added.

Since the effect of allowing for substitution between materials and value added depends on the size of σ, the major question of course is the right σ to assume. Rotemberg and Woodford [1992, Appendix] use a version of equation (4) to estimate the elasticity of substitution. They find $\sigma = 0.7$. Bruno [1984] surveys a number of papers and reports a consensus range for σ of 0.3-0.4.[19] Clearly, cyclical changes in the relative price of materials can overturn the initial findings only if σ is relatively large. Therefore, since the bias from an overestimate works against the results of the paper, I use 0.7 as the baseline value of σ.

I estimate (7) for two non-zero values of σ: $\sigma = 0.7$, and the Cobb-Douglas case with unit elasticity of substitution, $\sigma = 1$. I use the average share of value added over the sample period, and the observed change in the ratio of the cost of value added to the price of materials.

The estimates for the non-zero values of σ are in the second and third columns of Table II. There is almost no change in the estimated γ's when we allow for substitution: the estimates range from 0.83 to 0.86, depending on the instrument set and the value of σ. Based on these estimates, the true growth of variable capital and labor inputs is between 68 percent and 55 percent greater than the measured growth of these inputs. For plausible parameter values, the results are barely sensitive to variations in the parameters governing substitution between value added and materials. Even when we allow for the possibility that materials can be substituted for value added, we find that there is a considerable degree of cyclical factor utilization.

As noted in Section I, the residuals from equation (7) should be interpreted as true technology shocks. Table III compares the statistical properties of these technology shocks to those of the standard Solow residual. For the purposes of computing the technology series, I have allowed the degree of returns to scale to differ by sector. The results are for the Leontief case — $\sigma = 0$ — and use only the second instrument set (current oil price change; current and lagged values of the other instruments). The first column of Table III gives the average variance of the usual Solow residual for the 21 sectors: 0.00067. The average variance of the residuals from (7) is 0.00060. Thus, the variance of technology shocks falls by about 10 percent once we control for cyclical utilization. The next three columns report correlations between each of these two measures of technology change and three concepts of output growth: the growth of own-industry output, the growth of aggregate manufacturing gross output, and the growth of manufacturing value added. The correlation with own-output growth falls from 0.44 to 0.34 once we control for cyclical utilization. The average correlations with the two measures of aggregate output change also fall, though not by as large a percentage. Thus we see that the utilization-adjusted technology series is rather less friendly to the assumptions and conclusions of real-business-cycle models than the standard Solow residual.

However, Table III illustrates an interesting fact. All of the correlations with output growth reported there, including the correlations of the standard Solow residual, are much smaller than the figures that are familiar from aggregate data. For example, Burnside, Eichenbaum and Rebelo [1995] report that the correlation between the aggregate Solow residual and aggregate output growth is 0.80. Thus, it appears that productivity is more procyclical at higher levels of aggregation. Basu and Fernald [1994b] show that much of this effect comes from a composition bias: sectors with higher levels of productivity or larger returns to scale account for a larger fraction of output in booms. (Durable-goods industries, which have high productivity and highly procyclical output, are one example.) Thus, once we allow for heterogeneity between sectors (and firms), we see that the explanation of procyclical productivity depends not only on changes within each sector but also on changes in the share of each sector in the aggregate.

IV. PROBLEMS WITH A VALUE-ADDED PRODUCTION FUNCTION

Much of the empirical literature estimating production functions from gross-output data tests and rejects the conditions needed for the existence of a value-added function: see, e.g. Jorgenson, Gollop and Fraumeni [1987]. Of course, once one admits that this estimation is based upon data that contain systematic cyclical errors, it is difficult to know how one should interpret the rejection of the necessary separability conditions. The separability restriction is typically rejected because energy use, which is one component of overall materials consumption, is generally found to be a complement with respect to capital, but a substitute for labor. However, taking the nonseparability of energy into account, it is reasonable to assume that non-energy inputs can be separated from inputs of all the other factors: capital, labor, and energy. While cheap oil may lead to production using more machines and fewer workers, and expensive energy lead firms to economize on machines and use more labor, in either case one needs a certain amount of intermediate input to create output. Under this new assumption, we can write (3) as:

(3') $$Y_i = T_i G\left(V(Z_i K_i, C_i L_i, E_i), H(N_i)\right).$$

where E is the input of energy and N is all non-energy materials input. In this case V no longer has an interpretation as value added, but the essential idea of exploiting the relationship between V and H remains unchanged. Following a sequence of steps exactly analogous to those detailed in Section I, we can show that this new formulation leads to a slightly different estimating equation:

(7') $$\Delta y_i = \gamma\left[\Delta n_i - \sigma\left(\alpha_i^l + \alpha_i^k + \alpha_i^e\right)\left(\Delta p_i^{ve} - \Delta p_i^n\right)\right] + \Delta t_i.$$

Now the price index for V includes energy while the materials price index does not, and the cost share of energy is added to the shares of capital and labor. The growth rate of materials is now defined over non-energy materials only. But this equation is also easy to estimate, and the estimate of γ has just the same interpretation as before.

I present estimates of equation (7') in Table IV. The estimates of γ from this specification are even smaller than those in Table II, implying even higher estimates of unmeasured factor utilization. It is clear that the results are not being driven by this type of non-separability.

V. ERRORS IN MEASURING PRICES

A. Substitution of Inputs

Errors in measuring input price changes lead to problems in assessing firms' desires to substitute inputs. The presence of long-term labor contracts, either implicit or explicit, imply that the shadow wage might well differ from the observed wage. In a boom, where workers work longer hours, we would expect the shadow wage to exceed the observed wage. Also, in a boom the marginal cost of an "efficiency unit" of labor might well be significantly higher than the average cost.[20] Kydland and Prescott [1988] and Solon, Barsky and Parker [1994] document the long-standing conjecture that the real wage is more procyclical than it appears because the marginal worker hired in an upturn has lower human capital than the average employed worker. (However, as noted above, the Jorgenson data controls for much of this composition bias.) These considerations would argue that cyclical changes in Δp_i^v should be greater than measured. On the other hand, Carlton [1983, 1987] has argued that delivery lags and other reductions of service in upturns greatly increase the effective price of manufactured intermediate goods. These factors would argue that Δp_i^m should also be more procyclical than measured.

This issue is important, but equation (7) shows that it is not likely to change the results dramatically. First of all, errors in measuring prices alone create a bias only to the extent that σ is larger than zero. As noted above, the consensus range for σ is 0.3-0.4, so it seems unlikely that reasonable errors in measuring prices could overturn the results. For example, in order to reconcile the differences between the estimates of γ from Tables I and II for $\sigma = 0.3$, one would have to assume that the true *relative* price of value added rises by 0.90 percent more than measured for each percent increase in materials use. Errors of this magnitude are not credible; by comparison, Bils [1987] estimates that *nominal* wage costs increase by only 0.21 percent for each one percent change in production-worker employment, and his estimates are an upper bound for the reasons given in Trejo [1991].

Secondly, the bias works against the results of the paper only if the true price of value added is more procyclical than the true price of materials. Surprisingly, I find that this is true for the observed prices in the data set. Conventional wisdom holds that raw commodities prices are highly procyclical, while the real wage, much the larger input to the price of value added, is acyclical or only mildly procyclical (e.g. Murphy, Shleifer and Vishny [1989]). Both parts of this statement, however, need careful interpretation. While it is true that commodities prices are volatile and procyclical, raw commodities are only a small fraction of the total input of intermediate goods. Since "materials" are correctly classified by use and not by type of good, an industry's materials input includes all of its purchases from other industries (and from itself). For example, Boeing's purchase of a computer to fit into its newest jet is a purchase of an intermediate good, although a computer is typically classified as a high-technology finished good. Following the standards of national income accounting, the Jorgenson data set uses the input-output tables for the U.S. to incorporate such purchases into the definition of materials, along with more conventional materials inputs like raw commodities. Also, the labor cost data that are used for the estimation are adjusted for the fact that labor quality is highly countercyclical, making the cost of an efficiency unit of labor about twice as procyclical as the standard wage series, which is subject to composition bias. These two considerations together lead to the finding that materials prices are countercyclical relative to capital and labor costs.

As a shortcut for capturing these effects, I assume that the percent change in the true price of each input is a multiple of the observed change:[21]

$$\Delta p_i^{j*} = \beta^j \Delta p_i^j, \qquad \text{for } j = v, m$$

where β^m and β^v are parameters.

Substituting the hypothesized relationship between the true and observed prices into equation (7) gives:

(9) $$\Delta y_i = \gamma \Delta m_i - \sigma\beta^v\left(\alpha_i^l + \alpha_i^k\right)\Delta p_i^v + \sigma\beta^m\left(\alpha_i^l + \alpha_i^k\right)\Delta p_i^m + \Delta t_i.$$

Note that if $\beta^v = \beta^m$, changing β is equivalent to changing the elasticity of substitution between materials and value added. However, since we do not directly observe either σ or β, it seems best to simply estimate these unknown parameters (actually, the product of σ and the β's, since they cannot be identified separately from (9)).

The results are in the first line of Table V. They confirm the hypothesis of unmeasured factor inputs and Carlton's conjecture that the effective price of materials is procyclical: when we do not impose a priori values of σ and β, the regression chooses β^m to be large relative to β^v, and hence the degree of returns to scale to be smaller than even in the case when σ was constrained to be zero. As noted previously, the interpretation of a large β^m is that the effective price of materials is more procyclical than observed: that is to say, the effects suggested by Carlton are relatively more important. This is reasonable, because the price of value added computed from the data already contains a correction for the major bias identified in the literature, the cyclical change in the composition of the labor force. Thus, the major omitted factor that the regression estimates is the change in the effective price of materials. So it appears that the conventional wisdom may be right after all — the price of materials may be procyclical relative to the price of capital and labor — but not for the reasons usually given. These results imply an even greater degree of cyclical utilization than previously found, strengthening the findings of the paper. The estimate of γ using materials and energy is 0.75. So allowing for substitution, the estimated degree of variable factor utilization actually rises. The second line of Table V shows that the same is true when we allow for the possibility that the value-added production function is not separable and include energy as one of the inputs to the production of V.

B. Quantity Measurement Problems

A further problem is that errors in measuring materials prices can also imply errors in measuring their quantities, since the real quantities are not independently observed but rather derived from deflating nominal quantities by measured prices.[22]

At this point, we must distinguish between two price concepts. The first is the idea of opportunity cost, which is the price on which firms base their decisions. The second is the transactions price — the dollar price that firms pay for the intermediate goods they receive. The two need not be identical if the opportunity cost has a non-pecuniary component. For example, if there are delivery lags then in upturns the shadow value of an intermediate good can exceed its transactions price if a new good is not immediately available at that price. Other changes that can affect the opportunity cost include reductions in service or training. But these changes, although they can lead to problems in measuring opportunity cost, do not constitute errors in measuring transactions prices. Thus, they do not lead to problems in measuring real materials input.

However, it is of course correct that the observed price indices are not always accurate measures of transactions prices. This can lead to problems in measuring intermediate input growth. But this critique cuts both ways: since real output is also measured as nominal output deflated by a price index, there is a corresponding bias which has the opposite effect. Suppose the true prices of materials and output are both measured with error:

(10a) $$P_{m,t} = e^{\eta_t} P^*_{m,t},$$

(10b) $$P_{y,t} = e^{\upsilon_t} P^*_{y,t}.$$

The observed real quantities of materials and output are constructed by dividing nominal purchases by the observed prices:

(11a) $$M_t = \frac{P^*_{m,t} Y^*_t}{P_{m,t}},$$

(11b) $$Y_t = \frac{P^*_{Y,t} Y^*_t}{P_{Y,t}}.$$

Then the measured growth rate of materials usage differs from the actual usage by a factor $\Delta\eta$, and the measured and actual growth rates of inputs differ by $\Delta\upsilon$. For simplicity, assume that $\sigma = 0$, so the true relation between output and materials usage is:

$$\Delta y_t^* = \gamma \Delta m_t^* + \Delta t.$$

Now suppose we regress observed output on observed materials, using a single instrument, h. The probability limit of the estimated γ is given by:

$$\hat{\gamma} = \gamma\left[1 + \frac{\operatorname{cov}(\Delta\eta, h)}{\operatorname{cov}(\Delta m, h)}\right] - \frac{\operatorname{cov}(\Delta\upsilon, h)}{\operatorname{cov}(\Delta m, h)}. \tag{12}$$

If $\Delta\eta$ and $\Delta\upsilon$ were classical measurement errors they would be uncorrelated with the instrument, h. Thus any bias must come from the cyclicality of measurement error. The usual story is that transactions prices are more procyclical than posted prices, leading to a negative correlation between measurement error and the cycle. A negative correlation between $\Delta\eta$ and h would lead to a downward bias in $\hat{\gamma}$, and thus an overestimate of the degree of utilization. But a negative correlation between $\Delta\upsilon$ and h would have the opposite effect, so we cannot sign the bias. Suppose both measurement errors have the same correlation with the instrument. This is reasonable, because in many cases the two are the same — much of manufacturing gross output is also manufacturing materials input. Then the bias is

$$\hat{\gamma} - \gamma = (\gamma - 1)\frac{\operatorname{cov}(\Delta\eta, h)}{\operatorname{cov}(\Delta m, h)}. \tag{13}$$

Since the bias is proportional to (γ-1), the estimate is unbiased under the null hypothesis of constant returns to scale, even if measurement error is correlated with the cycle.

However, even if we take some value of γ greater than 1 as our null, the best evidence indicates that the bias is ambiguous and is likely to be small. Stigler and Kindahl [1970] undertook the most thorough study extant of the relationship between transactions and posted prices. They examined a large number of transactions prices for intermediate goods over the period 1957-1966 and compared their results to the corresponding BLS price indices.

Perhaps surprisingly, Stigler and Kindahl [1970, Table 6-8] found that measurement error in prices is small at business-cycle frequencies. At a monthly frequency, the correlation of the growth rate of their comprehensive price index with that of the corresponding BLS index is only 0.38. But at a quarterly frequency it rises to 0.58, and semiannually it is 0.73. Presumably the correlation is even higher at an annual frequency.

Interestingly, Stigler and Kindahl [1970, Table 6-8] also found that the correlation of $\Delta\upsilon$ with the cycle has opposite signs in recessions and expansions. In recessions $\Delta\upsilon$ is positive (an average monthly growth rate of 0.076), which is what one would expect if transactions prices are more flexible than posted prices. However, in expansions $\Delta\upsilon$ is negative, implying that posted prices actually rise more than transactions prices (average monthly growth of -0.039). So the sign of the bias is not evident a priori, especially if one recalls that recessions are much sharper than expansions and the number of months of expansion outnumber the recession months by about 4:1 over the sample.

Thus, the available evidence indicates that while the shadow price of materials is significantly more procyclical than the observed price, the transactions price is probably not much more procyclical than the observed price. The first finding strengthens the result of the paper, while the second indicates that the results are not driven by mismeasurement of real materials input.

Although a priori reasoning indicates that the net bias from mismeasurement of transactions prices is likely to be small, it would be desirable to confirm this directly. One method is to repeat the test for specific industries where BLS price indices are known to be good. Stigler and Kindahl [1970] mention Textiles as one such industry. Somewhat to their surprise, they also found that the BLS indices were good measures of transactions prices for steel and nonferrous metal products (semiannual correlations between the growth rates of the Stigler-Kindahl and BLS price indices for these product groups are 0.932 and 0.924, respectively). In the two-digit data the last two industries both fall in the category of Fabricated Metals. So we can see whether these two industries, which we know to have accurate price indices, give substantially the same results as the one for pooled manufacturing industries.[23]

The results are reported in Table VI. The first column gives the estimate of γ from equation (2), the second the estimate from (7) for $\sigma = 0$. The implied χ is reported in the third column. Since the estimation is done for each industry separately, the standard errors are quite large: for example, the point estimate for γ in Textiles is about 0.6, but the 95 percent confidence interval includes 0.93. But in both cases the difference between the two estimates has the right sign. And even though the absolute difference between the two estimates of γ for Textiles seems small, the implied χ is quite large. (The reason is that the Textile industry has a very high materials shares, about 0.7, so even a small difference in the coefficients implies a large amount of cyclical utilization of capital and labor.) However, the calculated χ's are about half of what we found in Table II. So, despite the preceding arguments, one might attribute some of the earlier results to cyclical errors in measuring prices. However, it is equally possible that these two industries are not representative of manufacturing industries overall.

Finally, one might argue that there is a separate but related problem: we observe materials purchases but not materials use. The two might differ when firms have the option of using some of their own capital and labor to create intermediate inputs, purchasing them from outside only when the internal workers are too busy.[24] For example, a firm might ask its secretaries to do xeroxing in-house when there is not much other work, but have them send it to a copy shop when other work presses. Formally, this type of situation implies that the firm has the option of substituting between total capital and labor input and *purchased* materials input. So even if the production function for final output is Leontief in primary inputs and materials, the consolidated production function (including the inputs used to produce in-house intermediate goods) might display considerable substitutability between total capital and labor and purchased materials. However, since these are the data used for the estimates of σ cited in Section III, allowing for substitutability of materials to the extent discussed there should suffice to address this issue.

VI. INCREASING RETURNS IN THE VALUE-ADDED FUNCTION

So far, the analysis has proceeded under the assumptions laid out in Section I, notably that the gross-output production function (F in equation (1)) is homothetic. (Actually, we assumed the stronger condition that F is homogeneous.) Thus, we could assume that both the V and H functions in equation (3) have constant returns; this was important later for measuring value added in (5). Suppose, however, that homotheticity does not hold. Then the V and H functions will not generally have the same degree of returns to scale. In particular, if the returns to scale of V exceeds the returns to scale of H, the test previously proposed will overestimate the degree of cyclical utilization.

I now outline a more general setup. Suppose that H is still homogeneous of degree one, but now let V be homogeneous of degree ρ in K and L.[25] G continues to be homogeneous of degree γ in V and H; note that G is thus not homothetic in K, L and M. Then the degree of returns to scale in the production of gross output, denoted RTS, is given by

(14) $$RTS = \gamma\left[\rho\left(\alpha^{l} + \alpha^{k}\right) + \left(1 - \alpha^{l} - \alpha^{k}\right)\right]$$

Note that RTS > γ if $\rho > 1$. If Δu were identically zero, then equation (2) would estimate RTS. Even with cyclical utilization, equation (7) continues to provide a consistent estimate of γ. However, we can no longer compare the two directly as we did before, since for $\rho > 1$ we now expect γ < RTS, even absent cyclical utilization. Thus we see that comparing the estimates from (2) and (7) is really a test of the joint hypothesis of homotheticity and no cyclical utilization. In the more general framework of this section, showing that γ < RTS is no longer sufficient to prove that cyclical utilization is responsible for Hall's finding of increasing returns.

Note, however, that this problem does not call into question all of the results obtained above. Equation (7) is still true; it is only the interpretation of γ that is different. Thus, the error term in (7) should still be interpreted as a true technology shock, and the properties of technology shocks that we have discovered are still correct even in this more general setup.

We should recognize, though, that there are economically relevant cases where the assumption of homotheticity might not hold. Increasing returns to scale, if they exist, probably come from fixed costs in production. Non-homotheticity, then, can be interpreted to mean that these fixed costs are proportionally larger for capital and labor inputs, so that a higher fraction of the inputs to the production of value added is consumed by "overhead" factors.[26]

Is it reasonable to attribute all of the difference between the estimated RTS from (2) and the estimated γ from (7) to non-homotheticity? Probably not. Under the hypothesis of no cyclical utilization we correctly estimate RTS = 1.09, and γ = 0.84. To reconcile these two estimates for the average observed average materials share of 0.55, we would need ρ = 1.62. If the increasing returns in V come from overhead inputs, we would need the ratio of overhead to total capital and labor to exceed 38 percent. Suppose we think of nonproduction workers as a reasonable proxy for overhead workers, as Ramey [1991] suggests. Ramey documents that the ratio of non-production workers to total employment has never exceeded 30 percent, and suggests 20 percent as an average figure. So it seems unlikely that non-homotheticity accounts for all of the difference between the two estimates. On the other hand, a substantial fraction of the difference may be due to this source.

So we need a way to pin down the homogeneity of V relative to H. One way would be to proceed by assuming that fixed costs are the source of non-homogeneity, and looking at detailed breakdowns of worker occupations to distinguish between "overhead" and "variable" workers. Another way is to use the idea that over the long run, all demand-induced input changes should be accommodated along the extensive margin. A third way is to use proxies for the change in utilization. I use the second and third methods.

Define $\Delta x_i^v = \left(\alpha_i^k + \alpha_i^l\right)^{-1}\left[\alpha_i^k \Delta k + \alpha_i^l \Delta l\right]$, so that $\Delta v_i = \rho\, \Delta x_i^v$. Then, using our new assumption that V is homogeneous of degree ρ in its inputs, we have the following cost-minimizing input demands:

(15a)
$$\Delta m_i = \frac{1}{\gamma}\Delta y_i + \sigma\left(\alpha_i^k + \alpha_i^l\right)\left(\Delta p_i^v - \Delta p_i^m\right) - \frac{1}{\gamma}\Delta t_i .$$

(15b)
$$\Delta x_i^v = \frac{1}{\rho\gamma}\Delta y_i - \Delta u_i - \frac{\sigma\left(1 - \alpha_i^k - \alpha_i^l\right)}{\rho}\left(\Delta p_i^v - \Delta p_i^m\right) - \frac{1}{\rho\gamma}\Delta t_i .$$

Subtracting (15b) from (15a), the growth of materials input relative to capital and labor input is:

(16)
$$\left(\Delta m_i - \Delta x_i^v\right) = \frac{\rho - 1}{\rho\gamma}\Delta y_i + \Delta u_i + \frac{\sigma + \sigma(\rho - 1)\left(\alpha_i^k + \alpha_i^l\right)}{\rho}\left(\Delta p_i^v - \Delta p_i^m\right) - \frac{\rho - 1}{\rho\gamma}\Delta t_i .$$

Equation (16) shows that there are three reasons why Δm might exceed Δx^V when Δy is positive. First, Δu might also be positive, so that total inputs of capital and labor increase more than measured inputs. Second, changes in relative prices might incline producers to use more materials. Third, V might be homogeneous of degree more than one: $\rho > 1$. The analysis in Section I allowed for the first and second reasons but not the third: note that equation (16) reduces to equation (4) if $\rho = 1$.

We can try and estimate ρ from the coefficient of Δy, (since we already have an estimate of γ from equation (7)), but we face the problem that the unobserved Δu and Δy are sure to be positively correlated with one another and with aggregate-demand instruments that are uncorrelated with Δt.

To alleviate this problem, we can try to use our intuition that utilization should not change permanently in response to a permanent change in demand. In any reasonable (convex) model of variable utilization, there must be increasing marginal costs of utilizing inputs more intensively, otherwise inputs would always be utilized at zero or 100 percent.[27] With increasing marginal cost of utilization, firms will not plan to accommodate long-run changes in output by changing utilization permanently, since in the long run it is always cheaper to adjust by adding more capital and labor rather than by changing their intensity of use.[28] Thus, suppose that utilization is a function of changes in output and other variables:

$$\Delta u_i = c_i + A_i(L)\Delta y_i + \xi_i,$$

where ξ is a vector of variables other than output that change utilization. Since utilization does not change permanently in response to a long-run change in y, A(1) = 0. Thus, the test is to regress the excess growth of materials input on a distributed lag of output and on changes in relative prices. Summing up the coefficients on the distributed lag of output will give us an estimate of $\frac{\rho-1}{\rho\gamma}$. Since the estimates of equations (7) and (9) give us estimates of γ, we can back out the parameter of interest, ρ. Note that this estimate of ρ will probably be too high, since the omitted variables ξ are likely to be positively correlated with Δy. For example, in times of high output firms probably find it more difficult to find qualified workers in external labor markets, making it more likely that employers will increase labor input by increasing utilization. Thus, even if we use instruments that are uncorrelated with technology change, the coefficient on Δy is likely to be biased upward. However, this bias favors Hall's hypothesis of large increasing returns and works against the basic method of the paper.

But this test is not quite dispositive as it stands. Suppose that there are increasing returns to V alone. Then it is true that if each firm were operating at a permanently larger scale, the ratio of materials to value added should also change permanently. However, if the estimation is done with industry data, we might fail to detect this outcome. Suppose the following story describes industry evolution. Initially, there is a permanent demand shock. Because V is homogeneous of degree more than 1, materials input grows faster than capital and labor input. But then, as the industry makes supernormal profits, new firms enter. Ultimately, the industry returns to long-run equilibrium with a larger number of firms but the same amount of output per firm. Clearly, the ratio of materials to value added returns to its long-run value, which might be interpreted as evidence that the short-run excess growth of materials is driven by cyclical utilization. To avoid this problem in the exercise proposed above, I examine all of the variables on a per-establishment basis. With this transformation of the data, we can estimate the extent to which the average per-establishment ratio of materials to capital and labor is affected by changes in the quantity of output per establishment.

Thus I estimate equations of the form:

$$(17) \qquad \left(\Delta m_i - \Delta x_i^v\right) = \delta(L)\Delta\tilde{y}_i + \beta_1\left(\Delta p_i^v - \Delta p_i^m\right) + \beta_2\left(\alpha_i^k + \alpha_i^l\right)\left(\Delta p_i^v - \Delta p_i^m\right) + v_i.$$

where $\tilde{y}$ is the log of output per establishment. The regression also includes industry-specific constants and time trends, a post-1973 dummy, and a trend interacted with the post-1973 dummy. These are meant to control for secular changes in productivity that might change the ratio of materials to the other inputs, especially labor. The estimate of $\delta(1)$ from (17) allows us to calculate ρ, given our previous estimates for γ; I use the current value and three lags of change in output per establishment.[29]

The results are in Table VII. The initial growth in output leads to a large increase in materials input relative to capital and labor inputs. However, the coefficient on the first lag of output is significantly negative, suggesting that a sizable fraction of the initial change in the materials-value-added ratio is driven by cyclical utilization. The sum of the lag coefficients is slightly negative, but not enough to erase a significant fraction of the change in the first period. Thus we see that a change in output per establishment does indeed cause a permanent change in the usage of materials relative to other inputs, which is consistent with the hypothesis that there are larger returns to scale in the production of value added. Since we can reject the hypothesis that $\delta(1) = 0$, we can use the estimate of $\delta(1)$ and our Table II estimate of γ of 0.84 to calculate ρ, the degree of returns to scale of V. The implied ρ is 1.38. Inserting this value into equation (14) and letting $\alpha^k + \alpha^l = 0.45$, we find that RTS = 0.99.

As a check on this result, we can try other proxies for Δu. Abbott, Griliches and Hausman [1988] suggest that the number of hours per worked by an average worker is a good index of work effort. Basu and Kimball [1994] confirm this intuition in a dynamic, optimizing model of variable factor utilization. Furthermore, they find that if changes in capital utilization take the form of adding additional shifts, and the major cost of capital utilization is a shift premium paid to workers on late shifts, then the hours per worker variable proxies for both capital and labor utilization. Thus, I replace Δu in (16) with Δh, where h is (the log of) hours per worker. The estimating equation is:

(18) $$\left(\Delta m_i - \Delta x_i^v\right) = \delta \Delta \tilde{y}_i + \zeta dh_i + \beta_1\left(\Delta p_i^v - \Delta p_i^m\right) + \beta_2\left(\alpha_i^k + \alpha_i^l\right)\left(\Delta p_i^v - \Delta p_i^m\right) + v_i.$$

The results are in Table VIII. One of the estimates is rather puzzling: the coefficient on Δh is highly significant but negative. Theory predicts that hours per worker should covary positively with utilization, and Δu enters (16) with a positive sign. However, the main coefficient of interest is δ. We find δ = 0.36, which is extremely close to the sum of the estimated lag coefficients in (17): 0.34. The implied ρ for δ = 0.36 is 1.43, and the implied RTS is 1.01.

Thus, using two different methods to control for variable utilization, we find that returns to scale, as conventionally defined, are almost exactly constant. This result is a sensible one, because it does not counterfactually imply large pure profits. There is a useful relationship between returns to scale, the markup ratio, μ, and the profit rate, π:

$$\mu \equiv \frac{P}{MC} = \frac{P}{AC}\frac{AC}{MC} = \frac{1}{1-\pi} RTS.$$

Since the average markup cannot be smaller than 1, diminishing returns to scale on the order of 0.84 would imply a pure profit rate of at least 16 percent (of gross output). There is no evidence of such large profit rates: the rental cost of capital series implies that the average profit rate in U. S. manufacturing is less than four percent, even without allowing for a return on intangible capital assets. Thus, finding that $RTS \approx 1$ implies a consistent picture of technology and market structure: constant returns in production, small pure profit rates, and hence small markups of price over marginal cost.

Finding ρ equal to about 1.4 is also broadly consistent with other evidence. If all the increasing returns in V come from overhead inputs (fixed costs, not diminishing marginal cost), then the implied ratio of overhead to total inputs is 0.28. This is not far from Ramey's [1991] suggested figure of 0.20.

Ramey considered only overhead labor; it is reasonable that total overhead inputs, including capital, might be a larger fraction of overall inputs.

If all of the increasing returns in V (and thus in the production of gross output as well) come from overhead capital and labor, then we can say something about the slope of the marginal cost curve as well. In this case, V and H are still homogeneous of degree one in the *variable* factors of production. But since we estimate $\gamma < 1$, production is homogeneous of degree less than one in the variable factors: i.e., there is increasing marginal cost. The fact that we find overall constant returns to scale means that the fixed costs and increasing marginal cost just offset one another. In terms of the standard U-shaped cost curve, firms produce at the minimum points of their average cost curves, the point of (locally) constant returns to scale. The implied slope of the marginal cost curve is $1/\gamma$, which equals 1.19 for a γ of 0.84. Thus, we conclude that marginal costs rise about 0.2 percent for each one percent increase in output.

Shea [1993] also concludes that the output elasticity of marginal cost is about 1.2, based on completely different evidence: the response of price to exogenous demand shocks. This figure is also quite consistent with the findings of Burnside, Eichenbaum and Rebelo [1995], who build on the approach of this paper.

VII. CONCLUSION

What should we conclude about the procyclicality of productivity? If one accepts the results of the paper, then it appears that all of the procyclicality of productivity in response to demand shocks stems from cyclical utilization, since the final estimate of returns to scale controlling for cyclical utilization is almost exactly 1.30 Changes in utilization are quite large: comparing the initial Table I estimate of 1.09 to a true returns to scale of 1 implies that utilization increases 0.17 percent for each percent increase in measured capital and labor. But even with this degree of variable utilization, we find that technology shocks are a significant source of variations in productivity. However, controlling for cyclical utilization reduces the variance of technology shocks by 10 percent and their correlation with output by more than 20 percent.

This last result is not favorable for the standard real-business-cycle model. That model anyway predicts that the correlation of output growth with technology change should be larger than the observed correlation of output growth with the Solow residual. If the true correlation of output and technology change is even smaller, the model is that much further from fitting the facts.[30]

The results also point up the lack of evidence for the kind of returns to scale estimates that would be necessary to validate sunspot-driven models of business cycles based on large increasing returns. Unless these models can be based on much smaller returns to scale, they seem to be theoretical curiosa rather than serious candidates for explaining business cycles.[31] In any event, these models rely crucially on the assumption that increasing returns come from diminishing marginal costs. If increasing returns instead come from overhead inputs, which seems more likely, then these models are incorrect in using returns-to-scale estimates to calibrate the slope of the marginal cost curve.

One of the implications of the results is that marginal costs rise quite quickly with output: the elasticity is 1.2. This finding predicts that firms should have strong motives to smooth production by using inventories. The early inventory literature failed to find evidence of production-smoothing, but more recent work by Fair [1989] and Krane and Braun [1991] finds production-smoothing using data on physical quantities.

This paper, therefore, sheds some light on the propagation mechanisms of business cycles. It shows that variable utilization is a plausible propagation mechanism, while imperfect competition, increasing returns to scale, and flat marginal costs are not. But what are the impulses driving business cycles? The results do not support the current generation of sunspot-driven models. Technology shocks seem an unlikely candidate: we found that allowing for cyclical utilization reduces the importance of technology shocks (Basu and Kimball [1994] present more evidence to this effect), and Basu and Fernald [1994b] show that much of what we measure as procyclical technology is in fact an artifact of composition bias. Thus, while the propagation mechanisms of business-cycle models are coming into focus, their ultimate cause remains a mystery.

University of Michigan and NBER

REFERENCES

Abbott, Thomas, Zvi Griliches, and Jerry Hausman, "Short Run Movements in Productivity: Market Power versus Capacity Utilization," Harvard University manuscript, 1988.

Aizcorbe, Ana, "Procyclical Labour Productivity, Increasing Returns to Labour and Labour Hoarding in U. S. Auto Assembly Plant Employment," *Economic Journal*, CII (1992), 860-73.

Bartelsman, Eric J., "Of Empty Boxes: Returns to Scale Revisited," Board of Governors of the Federal Reserve System manuscript, 1993.

Basu, Susanto, and John G. Fernald, "Constant Returns and Small Markups in U.S. Manufacturing." International Finance Discussion Paper 483, Board of Governors of the Federal Reserve System , 1994a.

Basu, Susanto, and John G. Fernald, " Aggregate Productivity and the Productivity of Aggregates," Hoover Institution manuscript, 1994b.

Basu, Susanto, and Miles S. Kimball, "Cyclical Productivity with Unobserved Input Variation," Hoover Institution manuscript, 1994.

Beaudry, Paul, and Michael Devereux, "Monopolistic Competition, Price Setting, and the Effects of Real and Nominal Shocks," Boston University manuscript, 1994.

Benhabib, Jess, and Roger E. A. Farmer, "Indeterminacy and Sector-Specific Externalities," *Journal of Monetary Economics* (forthcoming).

Bils, Mark,"The Cyclical Behavior of Marginal Cost and Price," *American Economic Review,* LXXVII (1987), 838-855.

Bils, Mark and Jang-Ok Cho, "Cyclical Factor Utilization," *Journal of Monetary Economics,* XXXIII (1994), 319-354.

Bruno, Michael, "Raw Materials, Profits, and the Productivity Slowdown," *Quarterly Journal of Economics,* XCIX (1984), 1-30.

Burnside, Craig, and Martin Eichenbaum, "Factor Hoarding and the Propagation of Business Cycle Shocks," NBER Working Paper 4675, 1994.

Burnside, Craig, Martin Eichenbaum, and Sergio Rebelo, "Capital Utilization and Returns to Scale," in *NBER Macroeconomics Annual*, Julio J. Rotemberg and Ben S. Bernanke, eds. (Cambridge, MA: MIT Press, 1995).

Caballero, Ricardo J., and Richard K. Lyons, "External Effects in U.S. Procyclical Productivity," *Journal of Monetary Economics,* XXIX (1992), 209-226.

Carlton, Dennis W., "Equilibrium Fluctuations when Price and Delivery Lag Clear Markets," *Bell Journal of Economics,* XIV (1983), 562-72.

Carlton, Dennis W., "The Theory and the Facts of How Markets Clear: Is Industrial Organization Useful for Understanding Macroeconomics?" in *Handbook of Industrial Organization*, R. Schmalensee and R. Willig, eds. (Amsterdam: North-Holland, 1987).

Evans, Charles L., "Productivity Shocks and Real Business Cycles," *Journal of Monetary Economics,* XXIX (1992), 191-208.

Fair, Ray C., "The Production-Smoothing Model is Alive and Well," *Journal of Monetary Economics,* XXIV (1989), 353-70.

Farmer, Roger E. A., and Jang-Ting Guo, "Real Business Cycles and the Animal Spirits Hypothesis," *Journal of Economic Theory,* LXIII (1994), 42-72.

Gordon, Robert J., "Are Procyclical Productivity Fluctuations a Figment of Measurement Error?" Northwestern University manuscript, 1990.

Hall, Robert E., "The Relation Between Price and Marginal Cost in U.S. Industry," *Journal of Political Economy,* XCVI (1988), 921-947.

Hall, Robert E., "Invariance Properties of Solow's Productivity Residual," in *Growth, Productivity, Employment,* Peter Diamond, ed. (Cambridge, MA: MIT Press, 1990).

Hall, Robert E., "Labor Demand, Labor Supply, and Employment Volatility," in *NBER Macroeconomics Annual,* Olivier J. Blanchard and Stanley Fischer, eds. (Cambridge, MA: MIT Press, 1991).

Hall, Robert E. and Dale W. Jorgenson, "Tax Policy and Investment Behavior," *American Economic Review,* LVII (1967), 391-414.

Jorgenson, Dale W., Frank Gollop, and Barbara Fraumeni, *Productivity and US Economic Growth,* (Cambridge, MA: Harvard University Press, 1987).

Jorgenson, Dale W., and Zvi Griliches, "The Explanation of Productivity Change," *Review of Economic Studies,* XXXIV (1967), 249-283.

Krane, Spencer D. and Steven N. Braun, "Production Smoothing Evidence from Physical-Product Data," *Journal of Political Economy* XCIX (1991), 558-581.

Kydland, Finn E. and Edward C. Prescott, "Cyclical Movements of the Labor Input and Its Real Wage," Federal Reserve Bank of Minneapolis Working Paper 413, 1988.

Murphy, Kevin M., Andrei Shleifer, and Robert W. Vishny, "Building Blocks of Market-Clearing Business Cycle Models," in *NBER Macroeconomics Annual,* Olivier J. Blanchard and Stanley Fischer, eds. (Cambridge, MA: MIT Press, 1989).

Plosser, Charles, "Understanding Real Business Cycles," *Journal of Economics Perspectives,* III (1989), 51-78.

Ramey, Valerie A., "Inventories as Factors of Production and Economic Fluctuations," *American Economic Review,* LXXIX (1989), 338-54.

Ramey, Valerie A., "Comment," in *NBER Macroeconomics Annual,* Olivier J. Blanchard and Stanley Fischer, eds. (Cambridge, MA: MIT Press, 1991).

Rotemberg, Julio J., and Michael Woodford, "Markups and the Business Cycle," in *NBER Macroeconomics Annual,* Olivier J. Blanchard and Stanley Fischer, eds. (Cambridge, MA: MIT Press, 1991).

Rotemberg, Julio J., and Michael Woodford, "The Effects of Energy Price Increases on Economic Activity," MIT manuscript, 1992.

Sbordone, Argia., "Cyclical Productivity in a Model of Labor Hoarding," Federal Reserve Bank of Chicago Working Paper, WP-93-20, 1993.

Schmitt-Gröhé, Stephanie, "Comparing Four Models of Fluctuations Due to Self-Fulfilling Expectations," Board of Governors of the Federal Reserve System manuscript, 1994.

Shapiro, Matthew D., "Cyclical Productivity and the Workweek of Capital," *American Economic Review,* LXXXIII (1993), 229-33.

Shapiro, Matthew D., "Capacity Utilization and the Marginal Premium for Work at Night," University of Michigan manuscript, 1994.

Solon, Gary, Robert Barsky, and Jonathan Parker, "Measuring the Cyclicality of Real Wages: How Important is Composition Bias?" *Quarterly Journal of Economics,* CIX (1994), 1-26.

Trejo, Stephen J., "The Effect of Overtime Pay Regulation on Worker Compensation," *American Economic Review,* LXXXI (1991) 710-40.

1. Recent papers in this literature include Abbott, Griliches and Hausman [1988], Aizcorbe [1992], Basu and Kimball [1994], Bils and Cho [1994], Burnside and Eichenbaum [1994], Burnside, Eichenbaum and Rebelo [1995], Gordon [1990], Hall [1990], Sbordone [1993], and Shapiro [1993].

2. Farmer and Guo [1994] and Beaudry and Devereux [1994].

3. See Schmitt-Gröhé [1994], who compares a number of such models in a common framework.

4. A similar idea is behind the use of electricity consumption data as a proxy for capital utilization by Jorgenson and Griliches [1967]. My procedure also allows me to investigate cyclical fluctuations in labor utilization.

5. All the series have had means removed.

6. Hall uses real value added as his measure of output, and thus uses only primary inputs of capital and labor. Basu and Fernald [1994a] show that using value added creates additional biases in estimating γ. Also, Hall estimates the reciprocal of γ by regressing Δx on Δy; see Bartelsman [1993] for a criticism of this procedure.

7. See Burnside and Eichenbaum [1994] and Basu and Kimball [1994].

8. Hall [1991] also takes time differences in first-order conditions for factor demands, though he focuses on labor demand as a function of output rather than value-added use as a function of materials demand.

9. There are a number of reasons why the true costs of inputs may differ from their observed costs. I discuss some of them and their consequences in Section V.

10. This parameterization is unlikely to be literally true, since the relationship between changes in input quantity and utilization is generally state-dependent. It is, however, a convenient way of presenting the results.

11. Whether the mismeasurement parameter is the same for both inputs depends on the degree of complementarity between capital and labor in production. If the utilization of the two factors can vary independently, the χ estimated from the data will be a weighted average of the parameters for capital and labor, where the weights are the cost shares of each input.

12. The only major difference between the industry groupings in the Jorgenson data and the standard 2-digit S.I.C. classification is that Motor Vehicles (S.I.C. 371) are separated from other transportation equipment (S.I.C. 372-79). (This is the standard practice of the BLS, though not of the BEA.) Thus, there are 21 manufacturing industries in the panel rather than the usual 20.

13. Note that the validity of this procedure is unaffected by markup pricing because markups affect the level of wages, not relative wages.

14. Quality-adjustment makes very little difference to the main results of the paper. Most of the regressions in the paper use materials as the only quantity variable on the right hand side, and there has been almost no change in materials "quality" (i.e. composition). The only noticeable effect is on the conventional estimates of γ (e.g. in Table I), which are about 5 percent larger with non-quality-adjusted data.

15. I am extremely grateful to John Fernald for providing me with the adjusted data.

16. The other tables in the papers use only the current oil price change but current and lagged values of the other instruments.

17. The Jorgenson data set separates intermediate inputs into energy and non-energy materials. Since this distinction is often not germane for my purpose, I typically combine both types of inputs into a single Divisia index.

18. Also, the two concepts of returns to scale are not the same. Conceptually, Hall is estimating the degree of increasing returns in the production of value added, which is larger than the degree of increasing returns in the production of gross output. Given an average materials share of 0.6 over the sample period, gross-output returns to scale of 1.10 translate to value-added returns to scale of 1.30 — still far from the estimates of about 3 reported by Hall [1990].

19. Of course, the data these papers use for their estimation may be flawed because of cyclical measurement error. The presence of cyclical utilization would tend to bias upward the estimate of the elasticity of substitution, since in the data one sees output growth being accompanied by low capital and labor growth but high materials growth, implying a high degree of substitution of materials for capital and

labor. As I have argued here, the better interpretation of that observation is that the increase in materials growth is accompanied by unobserved increases in the growth of variable capital and labor inputs. So the cited elasticities of substitution should be regarded as upper bounds.

20. Bils [1987] argues that the mandated 50 percent overtime premium implies that the marginal wage increases much faster than the average wage. However, Trejo [1991] finds that much of these legally mandated overtime payments are not allocative, since most of them are offset by compensating changes in the base wage.

21. True, unobserved, quantities are indicated by *'s.

22. I thank an anonymous referee and John Fernald for independently pointing out this problem.

23. I have concentrated on the output price indices, since this was the focus of the Stigler-Kindahl studies. But these two industries are likely to also have excellent input price indices. First, the largest source of industry inputs is usually itself, so to that extent the input and output price deflators are identical. Second, these two industries use large quantities of primary materials rather than processed goods as intermediate inputs. Since the prices of these commodities (agricultural products and metals) are largely set in auction markets, these too will probably be measured very accurately.

24. I am indebted to an anonymous referee for pointing out this issue.

25. Only the relative homogeneity of V to H matters, so the normalization of H is innocuous.

26. Of course, if production requires the same average ratio of overhead to total inputs for all factors — including materials — then the production function is again homothetic and the method for detecting variable utilization of capital and labor presented in Section I is exactly correct.

27. See Burnside and Eichenbaum [1994] and Basu and Kimball [1994].

28. Some have argued that the recent upward trend in the workweek of capital is evidence against this argument. Utilization may indeed have a trend, since technological progress may reduce the marginal cost of utilization over time. My argument requires only that these technology changes be substantially independent of changes in demand.

29. We could in principle calculate ρ from the ratio of β_1/β_2, but the standard errors for the β's are very large since the variables they multiply are nearly collinear.

30. Note, however, the caveat about composition bias discussed at the end of Section III.

31. However, modeling variable utilization as a propagation mechanism for technology changes may reverse this conclusion: see Burnside and Eichenbaum [1994].

32. Benhabib and Farmer [forthcoming] have made some progress on this issue by considering a multi-sector model with small sector-specific externalities.

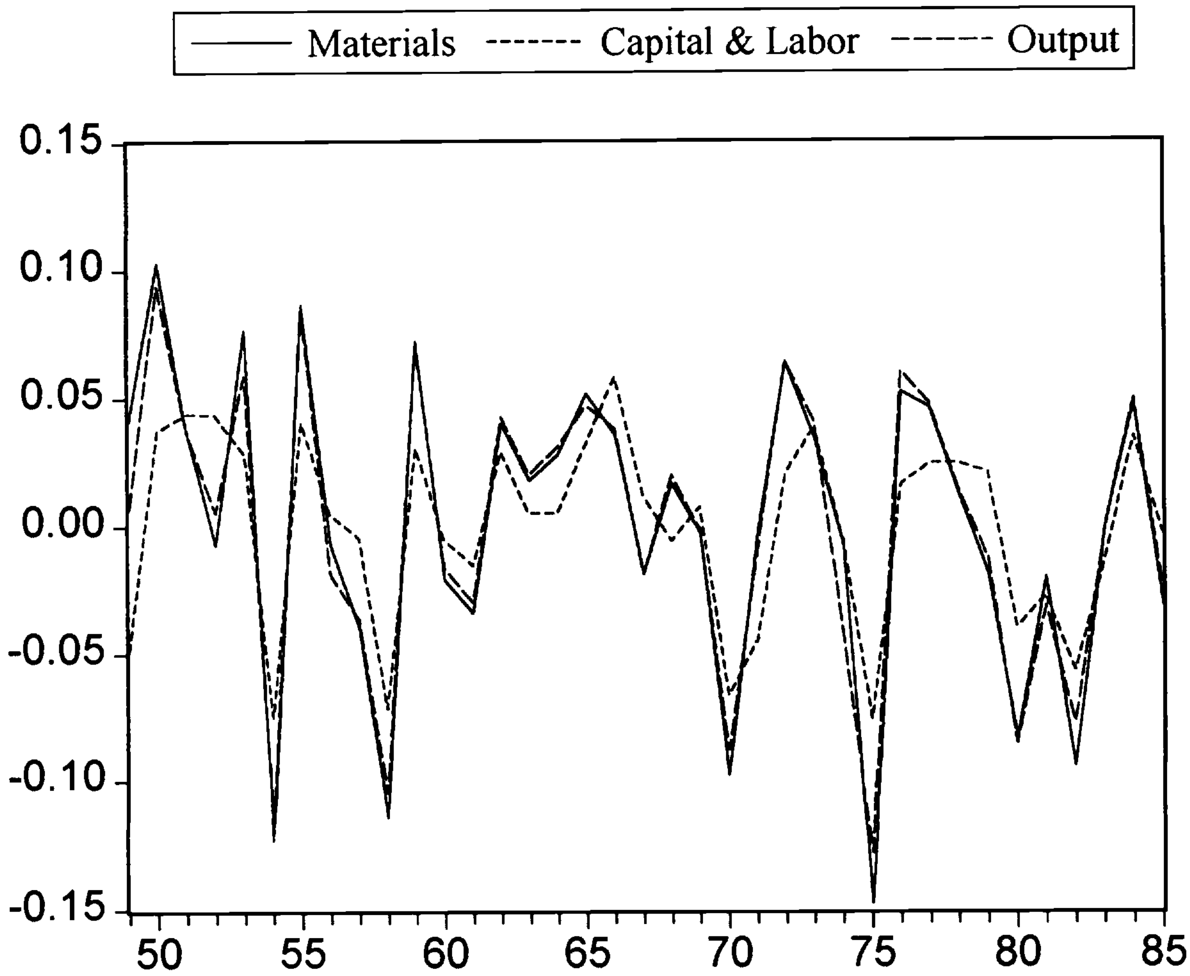

FIGURE I.

Growth Rates of Inputs and Output, U.S. Manufacturing

TABLE I

ESTIMATE OF RETURNS TO SCALE FROM EQUATION (2)

(STANDARD ERRORS IN PARENTHESES)

Instrument Set	Parameter	Estimate
1	γ	1.100 (0.015)
2	γ	1.092 (0.015)
3	γ	1.092 (0.013)

Note: Instrument set 1 contains current oil price, defense spending, and political dummy.

Instrument set 2 contains current oil price, and current and lagged defense spending and political dummy.

Instrument set 3 contains current and lagged oil prices, defense spending, and political dummy.

TABLE II

ESTIMATES OF RETURNS TO SCALE FROM EQUATION (7)

(STANDARD ERRORS IN PARENTHESES)

Instrument Set	Parameter	$\sigma = 0$	$\sigma = 0.7$	$\sigma = 1$
1	γ	0.79 (0.029)	0.83 (0.031)	0.83 (0.027)
	Implied χ	0.80	0.68	0.68
2	γ	0.83 (0.017)	0.83 (0.021)	0.84 (0.022)
	Implied χ	0.68	0.68	0.63
3	γ	0.84 (0.011)	0.86 (0.012)	0.86 (0.012)
	Implied χ	0.63	0.55	0.55

Note: Instrument set 1 contains current oil price, defense spending, and political dummy.

Instrument set 2 contains current oil price, and current and lagged defense spending and political dummy.

Instrument set 3 contains current and lagged oil prices, defense spending, and political dummy.

TABLE III.

PROPERTIES OF SOLOW RESIDUALS AND TECHNOLOGY RESIDUALS

(AVERAGES FOR 21 INDUSTRIES)

	Variance × 10^3	Correlation with Own Output	Correlation with Manufacturing Gross Output	Correlation with Manufacturing Value Added
Solow Residual	0.67	0.44	0.21	0.26
Regression Residual	0.66	0.34	0.17	0.22

TABLE IV

ESTIMATES OF RETURNS TO SCALE FROM EQUATION (7')

(STANDARD ERRORS IN PARENTHESES)

	$\sigma = 0$	$\sigma = 0.7$	$\sigma = 1$
γ	0.80 (0.017)	0.81 (0.020)	0.79 (0.019)
Implied χ	0.87	0.80	0.94

TABLE V

ESTIMATES OF RETURNS TO SCALE AND THE CYCLICALITY OF PRICES.

(STANDARD ERRORS IN PARENTHESES)

	γ	$\gamma\sigma\beta^{v}$	$\gamma\sigma\beta^{m}$	Implied β^{m}/β^{v}	Implied χ
Materials and Energy (Δm)	0.75 (0.015)	0.47 (0.064)	0.80 (0.065)	1.70	2.31
Materials only (Δn)	0.72 (0.015)	0.40 (0.071)	0.97 (0.046)	2.42	2.75

TABLE VI

ESTIMATES FOR INDUSTRIES WITH RELIABLE PRICE DEFLATORS

(STANDARD ERRORS IN PARENTHESES)

	Estimated γ from Equation (2)	Estimated γ from Equation (7)	Implied χ
Textiles	0.61 (0.16)	0.54 (0.13)	0.36
Fabricated Metals	1.06 (0.08)	0.90 (0.09)	0.35

TABLE VII.

ESTIMATES OF NON-HOMOTHETICITY IN PRODUCTION FROM EQUATION (17)

(STANDARD ERRORS IN PARENTHESES)

δ_0	δ_1	δ_2	δ_3	β_1	β_2	$\delta(1)$	p-value for $\delta(1)=0$
0.38 (0.032)	-0.11 (0.027)	0.03 (0.026)	0.02 (0.022)	0.20 (0.074)	0.11 (0.024)	0.34	0.0

TABLE VIII.

ESTIMATES OF NON-HOMOTHETICITY IN PRODUCTION FROM EQUATION (18)

(STANDARD ERRORS IN PARENTHESES)

δ	ζ	β_1	β_2	p-value for $\delta = 0$
0.36 (0.043)	-1.53 (0.131)	0.20 (0.074)	0.11 (0.024)	0.0