Volume Title: NBER Macroeconomics Annual 1995, Volume 10

Volume Author/Editor: Ben S. Bernanke and Julio J. Rotemberg, eds.

Volume Publisher: MIT Press

Volume ISBN: 0-262-02394-6

Volume URL: http://www.nber.org/books/bern95-1

Conference Date: March 10-11, 1995

Publication Date: January 1995

Chapter Title: Capital Utilization and Returns to Scale

Chapter Author: Craig Burnside, Martin Eichenbaum, Sergio Rebelo

Chapter URL: http://www.nber.org/chapters/c11017

Chapter pages in book: (p. 67 - 124)

Craig Burnside, Martin Eichenbaum, and Sergio Rebelo

UNIVERSITY OF PITTSBURGH; NORTHWESTERN UNIVERSITY, NBER, AND THE FEDERAL RESERVE BANK OF CHICAGO; AND UNIVERSITY OF ROCHESTER AND NBER

Capital Utilization and Returns to Scale

1. *Introduction*

This paper studies the implications of procyclical capital utilization rates for inference regarding cyclical movements in labor productivity and the degree of returns to scale. To study cyclical movements in capital utilization we use two measures of capital services: industrial electrical use and data on the workweek of capital. The investigation addresses five questions using different assumptions about the production technology:

1. Is the phenomenon of near or actual short-run increasing returns to labor (SRIRL) an artifact of the failure to accurately measure capital utilization rates?
2. Can we find a significant role for capital services in aggregate and industry-level production technologies?
3. Is there evidence against the hypothesis of constant returns to scale?
4. Can we reject the notion that the residuals in our estimated production functions represent technology shocks?
5. How does correcting for cyclical variations in capital services affect the statistical properties of estimated aggregate technology shocks?

Briefly, the answers are: (1) yes, (2) yes, (3) no, (4) no, and (5) a lot.

Our investigation utilizes aggregate data and two new data sets: a panel on two-digit standard industrial classification code (SIC) industries and a panel on three-digit SIC industries. We argüe that the data are well described by a constant-returns-to-scale production function. The esti-

The authors owe a deep debt of gratitude to Joe Beaulieu. This paper would not have been possible without his help. In addition, we would like to thank Ben Bernanke, Mark Bils, Lawrence Christiano, John Fernald, Joe Mattey, Julio Rotemberg, Matt Shapiro, and Steve Strongin for useful conversations.

mated coefficients on labor and capital services are similar to the shares of labor and capital in national income: 0.64 and 0.36, respectively. The estimated residuals from our estimated production technology have two important properties. First, in most cases, they pass a variant of Hall's (1988) invariance test; they are consistent with a set of orthogonality conditions that candidate measures of technology shocks ought to satisfy. In contrast, the traditionally calculated Solow residual does not pass the Hall test. Second, they are much less volatile and less correlated with aggregate output than the empirical measure of technology shocks used in the real business cycle (RBC) literature.

The observation that average labor productivity is procyclical, which goes back at least as far as Fabricant (1942), is closely related to a well-known puzzle: capital appears to play no role in explaining cyclical movements in output. This puzzle has been stressed by Solow (1964), Lucas (1970), and Bernanke and Parkinson (1991), among others. Exploring different data sets over different sample periods and using different estimation strategies, they arrive at the same conclusion: capital enters estimated production functions either with the wrong sign or not at all.

The typical reaction to this finding is to ignore movements in capital when studying cyclical productivity fluctuations. While disheartening, Perry's (1973) rationale for doing this seems compelling:

If capital is ignored, it is for a simple pragmatic reason: one cannot find an important or statistically significant role for capital in a freely estimated aggregate production function or any equivalent relation that one might use in estimating potential output.

An alternative response is to obtain better measures of capital services. This is the strategy we pursue. And with better measures, we find that there is an important and statistically significant role for capital services. Moreover, estimated returns to scale are roughly constant.

Yet another reaction to the apparent unimportance of capital in estimates of production functions is to stop estimating production functions. In the macro literature, authors like Hall (1988) have studied returns to scale by relating the growth rate of output to a cost-weighted sum of the growth rates of inputs. We implement the Hall-type strategy using our measures of capital services to assess the robustness of our findings. The key result is that, with this approach too, we cannot reject the hypothesis of constant returns to scale.

Using our measures of capital utilization, we argue that neglecting cyclical variations in capital services affects inference about why average labor productivity is procyclical. This is important because the procyclical nature of average labor productivity has played a central role in recent debates about the causes of aggregate economic fluctuations. RBC theorists, such as Kydland and Prescott (1982) and Long and Plosser (1983), emphasize the importance of exogenous shocks to productivity as the main impulse to postwar U.S. business cycles. With shocks to the aggregate production technology, RBC models can account for the observed procyclical nature of labor productivity. Other researchers, sometimes organized under the "new Keynesian" banner, have sought to revive much of the common wisdom associated with the IS-LM paradigm using models grounded on microeconomic foundations.[1] These researchers emphasize the importance of demand shocks as impulses to economic fluctuations. In conjunction with increasing returns to scale, demand shocks too can generate procyclical movements in productivity.[2]

Increasing returns to scale are also an essential ingredient in a recent strand of literature that emphasizes the importance of multiple equilibria for understanding business cycles.[3] In standard RBC models the competitive equilibrium can generally be characterized as the solution to a planning problem, which, being a concave program, has a unique solution. With increasing returns the resource constraints facing the economy no longer define a convex set, so there can be more than one equilibrium path. Under these circumstances, recessions can be the result of pessimistic, self-fulfilling beliefs of agents in the economy. With increasing returns to scale, low output and employment levels will be associated with low levels of labor productivity.

An alternative explanation of procyclical productivity, and the one which is most relevant to this paper, focuses on cyclical movements of capital utilization and labor hoarding. This explanation has recently been explored by, among others, Greenwood, Hercowitz, and Huffman (1988), Kydland and Prescott (1988), Burnside, Eichenbaum, and Rebelo (1993), Finn (1991), Basu and Kimball (1994), Bils and Cho (1994), and Burnside and Eichenbaum (1994).

Given the importance of disentangling the sources of procyclical productivity, analyzing the properties of the Solow residual and estimating the degree of returns to scale have become priority items in the macroeconomics research agenda. Authors like Basu and Kimball (1994) use industry-level annual data to assess the contribution of unobserved in-

1. See Mankiw and Romer (1991) as well as the references therein.
2. Rotemberg and Summers (1990) combine labor hoarding behavior along with nominal price rigidities as a way of rationalizing the cyclical behavior of average labor productivity. Perhaps this defines them as "old" Keynesians.
3. See, for example, Farmer and Guo (1994), Rotemberg and Woodford (1995), and the references therein.

put variation to cyclical movements in total factor productivity. Shapiro (1993b) uses annual data to study the importance of movements in the workweek of capital. A different body of research, originally associated with Hall (1988, 1990), focuses on the returns to scale and externalities. Hall (1988, 1990) claimed to find evidence of large markups and increasing returns to scale. Using similar methods, Caballero and Lyons (1992) and Bartlesman, Caballero, and Lyons (1994) argue that there are large spillover externalities at the industry level.

Bartlesman (1993) suggests that Hall's evidence of large increasing returns to scale can be explained entirely by the presence of small-sample bias in Hall's econometric procedures. A different criticism has been levied by Basu and Fernald (1994a,b), who argue that with imperfect competition, the use of value-added data leads to spurious findings of large increasing returns to scale and external effects. Indeed, they show that when gross output data on two-digit SIC industry-level data are used, evidence of increasing returns and externalities disappears.[4] In addition, at this level of aggregation, findings of external spillover effects are associated with an exceedingly improbable implication: estimated total returns to scale are roughly constant, so spillover effects emerge only at the cost of concluding that there are very large internal *decreasing* returns to scale (see Basu and Fernald, 1994a, and Burnside, 1994). One exception to this characterization is the four-digit SIC industry-level study by Bartlesman, Caballero, and Lyons (1994).

All of the previous studies use variants of Hall's (1988) methodology in conjunction with annual data. Rather than rely solely on annual data, we consider different specifications of technology that allow us to attack the problem with quarterly aggregate and industry-level panel data. As it turns out, there are interesting tradeoffs involved in using different specifications of technology. These involve the generality of the specification being considered, the assumptions about market structure, and the data requirements that are needed to estimate the parameters in question. But overall returns to scale is a dimension across which all of the specifications can be compared. And as it turns out, inference is very robust on this dimension.

In all cases we estimate the parameters of technology using a three-

4. The fact that value-added and gross output data yield different estimates of the degree to scale can be explained even in the presence of perfect competition. In order for value-added output to correctly measure the marginal productivity of primary inputs, one of the following three restrictions has to hold: (1) materials and energy are used in fixed proportions with gross output (Leontief aggregation), (2) the relative price of materials and energy in terms of gross output is constant (Hicks aggregation), and (3) the gross output function has the form $Y = F[(K, L), M, E]$ (weak separability), where K, L, M, and E denote capital, labor, materials, and energy, respectively. See Bruno (1978).

stage least-squares procedure that exploits the fundamental identifying assumption proposed by Hall (1988): shocks to technology ought to be orthogonal to variables that are "known neither to be causes of productivity shifts nor to be caused by productivity shifts." In our view, shocks to monetary policy, say as measured by Christiano, Eichenbaum, and Evans (1994), as well as variables like the relative price of oil qualify to be included in this class.

The remainder of this paper is organized as follows. Our model is presented in Section 2. Econometric procedures and data sources are detailed in Section 3. Empirical results are presented in Section 4. Some important limitations of our analysis are discussed in Section 5. The key problem, as in the related literature (see for example Hall, 1988, 1990, and Basu, 1993) is the potential effect of unobserved overhead capital and labor on the interpretation of our estimated parameters. Concluding comments are contained in Section 6.

2. *Model Specification*

We begin by providing an overview of the three specifications of technology used in our empirical work. In addition we summarize the tradeoffs with each specification. These pertain to the generality of the specification, assumptions about market structure, and the data needed to implement the model empirically.

Let Y_t denote time t gross output. In our first specification we assume that

$$Y_t = \min(M_t, V_t), \tag{1}$$

where M_t denotes time t materials and V_t denotes a function that involves hours worked (L_t), the stock of capital (K_t), and electricity use (E_t). Capital services and E_t are related via a Leontief technology. Our second specification relaxes this assumption and allows for substitution between capital services and E_t. The third specification abandons the assumption that M_t and V_t are related via a fixed-coefficients technology. Here we assume that Y_t is a differentiable function of capital services (S_t), energy (E_t), L_t, and M_t:

$$Y_t = F(S_t, L_t, E_t, M_t). \tag{2}$$

The following table summarizes the tradeoffs involved in using the different specifications:

	Specification 1	*Specification 2*	*Specification 3*
Data frequency	Quarterly, Annual	Quarterly, Annual	Annual
Industry level	2-, 3-digit SIC	2-digit SIC	2-digit SIC
Goods market	No assumptions	No assumptions	No assumptions
Factor markets	No assumptions	Hours, electricity: perfect competition	All factors: perfect competition

The advantages of the first specification are that it allows us to use quarterly two- and three-digit SIC data and makes no assumptions about market structure. The cost is that it imposes a Leontief relationship between M_t and V_t and a Leontief relationship between capital services and E_t. The advantages of the second specification are that it allows us to use quarterly two-digit SIC data and assumes only that labor and electricity markets are perfectly competitive. The cost is that it imposes a Leontief relationship between M_t and V_t. The advantage of the third specification is that it imposes no restrictions on the production technology other than differentiability. The cost is that we can only use annual data and we must assume that all factor markets are perfectly competitive.

We turn to a more detailed discussion of the three technology specifications.

2.1 SPECIFICATION 1: THE SIMPLEST STRUCTURE OF PRODUCTION

In our simplest production specification, Y_t is produced by combining value added (V_t) and materials (M_t) according to the Leontief production function (1). Basu (1993) has argued persuasively that this Leontief form provides a good approximation to the structure of production in manufacturing, since movements in materials track movements in gross output very closely. An additional motivation for working with this specification, emphasized by Bernanke and Parkinson (1991), is that it allows us to work with industry-level gross output data, despite the absence of observations on material inputs.

The value added produced in one hour by one worker is $A_t F(1, K_t/N_t)$. In setting up our benchmark case, we suppose that the function $F(\cdot)$ is homogeneous of degree one, concave, and twice differentiable. The variable N_t is the number of time t workers, and A_t reflects the state of time t technology and other exogenous factors that affect productivity. Since each worker is employed for H_t hours, the total value added produced by the firm in period t is[5]

5. This specification of technology is similar to the one used in Chari, Christiano, and Eichenbaum (1995).

$$V_t = N_t H_t A_t F(1, K_t/N_t) = A_t F(N_t H_t, K_t H_t). \tag{3}$$

So to measure capital services we need to multiply the capital stock by its workweek. In our formulation this coincides with the number of hours that each worker is employed.[6] This correction for capital utilization is similar to the one originally employed by Solow (1957), which involved multiplying the stock of capital by the employment rate.

The key problem involved in using this production structure is the absence of good direct measures of capital services. Certainly none are available at the quarterly frequency. However, following Griliches and Jorgenson (1967), we can measure these services indirectly via electricity consumption. This strategy has also been employed by Costello (1993) in her study of the properties of the Solow residual in an international context.

Suppose that electricity consumption per machine is proportional to its workweek H_t. Then total electricity consumption E_t is given by

$$E_t = \phi H_t K_t. \tag{4}$$

Defining total time t hours as $L_t = N_t H_t$, and using equation (1), we obtain

$$Y_t = A_t F(L_t, E_t/\phi). \tag{5}$$

From an empirical standpoint, this formula has an important advantage: observations on all of its variables are available at the quarterly frequency for two- and three-digit SIC industries.[7] The disadvantage is that it imposes the strong restriction that the elasticity of electricity use with respect to capital use is equal to one. There are a variety of reasons why this may not be true, such as the existence of overhead capital. The generalized technology discussed in Section 2.3 relaxes the unit-elasticity assumption. In Section 5 we discuss how neglecting overhead capital (and labor) can bias our results.

6. Notice that the production function exhibits increasing returns to scale in N_t, H_t, and K_t. It is standard to assume that there are increasing marginal costs associated with increases in H_t, say because the rate of depreciation is an increasing function of H_t. In this case we can optimize with respect to H_t and obtain a reduced-form production function that is concave in N_t and K_t. See Greenwood, Hercowitz, and Huffman (1988).

7. A standard criticism of the use of electricity as a measure of capital utilization is the possible presence of a trend in the electricity–capital ratio. This could reflect a change in the composition of capital away from structures to equipment. We could capture this effect by allowing ϕ to be a deterministic function of time. If the function $F(\cdot)$ were Cobb–Douglas, this would simply change the unconditional growth rate of technology.

2.1.1 Line Speed and Labor Hoarding We now consider the effects of variations in labor effort and in the intensity with which capital is used. Suppose that hourly capital services per worker in equation (3) are $\lambda_t K_t / N_t$, so that the value added produced in one hour by a worker is $A_t F(1, \lambda_t K_t / N_t)$. Here λ_t denotes the intensity with which capital is used, or "line speed." Also suppose that electricity consumption per machine is proportional to the effective workweek of the machine, $\phi \lambda_t H_t$. Then E_t is equal to $\phi \lambda_t H_t K_t$ and equation (5) remains unchanged. So, according to this simple formulation, using electricity consumption allows us to measure capital services in a way that is robust to changes in line speed.

To allow for unobserved changes in labor effort, i.e., "labor hoarding," define the number of efficiency units of labor as $\zeta_t H_t$. Here ζ_t measures effort per hour. Suppose that total electricity use depends on effort, so that $E_t = \phi \zeta_t H_t K_t$. Then equation (5) becomes

$$V_t = A_t F(\zeta_t L_t, E_t / \phi). \tag{6}$$

Notice that E_t still measures total capital services. However, the production function now involves ζ_t, unobserved labor effort. One way to incorporate labor hoarding into the analysis is to specify the costs associated with supplying effort. We could then use the condition that determines the optimal supply of effort to solve for ζ_t as a function of unknown parameters and observable variables. The resulting "reduced form" production function could be used in empirical work. This is the approach pursued by Burnside, Eichenbaum, and Rebelo (1993), Basu and Kimball (1994), and Burnside and Eichenbaum (1994). Here we wish to see how far we can go in explaining the apparent short-run increasing returns to labor by controlling for capital utilization while remaining as eclectic as possible about market structure and the determinants of labor supply. Because of this we abstract from variations in effort in our empirical analysis. This will tend to bias our results *against* the null hypothesis of constant returns to scale.

2.2 SPECIFICATION 2: A SLIGHT GENERALIZATION

The second production specification that we consider is given by

$$Y_t = \min(M_t, V_t^*),$$

where V_t^* is defined as

$$V_t^* = A_t F(L_t, K_t^*). \tag{7}$$

Here K_t^* is given by a CES function of capital and electricity use:[8]

$$K_t^* = [\mu(H_t K_t)^\rho + (1 - \mu)E_t^\rho]^{1/\rho}, \qquad \rho < 1. \tag{8}$$

This type of two-level production function was first proposed by Sato (1967) and has often been used in the applied general equilibrium literature (see for example Ballard, Fullerton, Shoven and Whalley, 1985). Our assumption that the production function is weakly separable between labor and the two other inputs is consistent with Berndt and Wood's (1979) parameter estimates for a translog cost-function fit to a panel of manufacturing industries.

To implement this formulation we need to make use of the optimality condition that determines the firm's demand for electricity. Suppose that the firm acts as a price taker in the market for labor and electricity. Then cost minimization requires that the firm equate the marginal rate of substitution between N_t and E_t to the relative price of the factors, $W_t H_t / P_{Et}$. Here W_t denotes the real wage rate per hour worked at time t:

$$\frac{A_t F_2(L_t, K_t^*)(1 - \mu)(K_t^*/E_t)^{1-\rho}}{F_1(L_t, K_t^*)} = \frac{P_{Et}}{W_t}. \tag{9}$$

Equation (9) holds regardless of whether the firm is a perfect competitor or not in the goods market. Here F_i denotes the partial derivative with respect to the ith argument of F.

In our empirical work we assume that $F(\cdot)$ has a Cobb–Douglas form so that

$$Y_t = A_t\,(L_t)^{\alpha_1}\,(K_t^*)^{\alpha_2}. \tag{10}$$

Consistent with this notation, we do not impose the *a priori* restriction that the production function is constant returns to scale, i.e., we do not assume that $\alpha_1 + \alpha_2 = 1$. Given (9) and (10), gross output can be written as a geometric average of total hours (L_t), energy consumption (E_t), and the price of electricity relative to labor (p_{Et}):

$$Y_t = \left((1 - \mu)\frac{\alpha_2}{\alpha_1} \right)^{\alpha_2/\rho} A_t (L_t)^{\alpha_1 + \alpha_2/\rho} E_t^{\alpha_2 - \alpha_2/\rho} p_{Et}^{-\alpha_2/\rho} \tag{11}$$

8. V_t^* does not correspond to measured value added, because it depends upon E_t. This is immaterial for our empirical work, since we use gross output data rather than value-added data.

Taking first differences and letting lowercase letters denote logarithms, we obtain

$$\Delta y_t = \gamma_0 + \gamma_1 \, \Delta l_t + \gamma_2 \, \Delta e_t + \gamma_3 \, \Delta p_{Et} + \epsilon_t. \tag{12}$$

Here $\gamma_0 + \epsilon_t$ denotes the growth rate of A_t, $\gamma_1 = \alpha_1 + \alpha_2/\rho$, $\gamma_2 = \alpha_2 - \alpha_2/\rho$, and $\gamma_3 = -\alpha_2/\rho$. Our basic production structure coincides with the special case in which the elasticity of substitution between capital and energy is equal to zero ($\rho = -\infty$). Here (12) becomes[9]

$$\Delta y_t = \gamma_0 + \alpha_1 \, \Delta l_t + \alpha_2 \, \Delta e_t + \epsilon_t. \tag{13}$$

We now turn to a brief discussion of the differentiable technology (2).

2.3 SPECIFICATION 3: THE DIFFERENTIABLE TECHNOLOGY

Much of the recent literature that uses annual data to study productivity assumes that output is produced according to (2). Taking a first-order log-linear approximation to this technology yields

$$\Delta y_t = \eta \, \Delta x_t + \epsilon_t, \tag{14}$$

where η denotes overall returns to scale, and Δx_t is a cost-weighted measure of the growth rate of aggregate inputs,

$$\Delta x_t = c_{St} \, \Delta s_t + c_{Lt} \, \Delta l_t + c_{Mt} \, \Delta m_t + c_{Et} \, \Delta E_t.$$

Here lowercase symbols denote logarithms of upper case symbols and c_{jt} denotes the share of factor j in the total cost, at time t.

In sum, our three specifications of technology give rise to three types of relations between factor inputs and output, (12), (13), and (14). But absent further restrictions, these are without empirical content. They hold as identities. For them to have content, identifying assumptions must be imposed on the stochastic process ϵ_t. We turn to this issue in the next section.

3. *Econometric Method and Data*

The fundamental identifying assumption underlying our analysis is that ϵ_t is a stationary technology shock (not necessarily i.i.d.). Suppose that

9. This relation can be derived directly from (3) under the assumption that V_t is Cobb–Douglas in L_t and K_tH_t, where the weights do not necessarily add up to one.

we have observations on a subset of those variables, which, in Hall's (1988) terminology, are known neither to be causes of productivity shifts nor to be caused by productivity shifts. Let z_t denote the time t realization of these variables. By assumption

$$E[z_t \epsilon_t] = 0. \tag{15}$$

We think of (15) as representing a set of necessary conditions that candidate measures of technology shocks must satisfy. Suppose that the dimension of z_t is greater than or equal to the number of parameters in the production technology. Then (15) can be used to estimate the parameters of (12), (13), and (14). We do so via three-stage least squares.

In some cases we present "restricted" estimates, using panels of industry-level data. These estimates are obtained by imposing the linear restriction that the parameters of the production technology, with the exception of γ_0, are the same in all industries. The intercept term for each industry is left unrestricted. When the dimension of z_t exceeds the number of parameters to be estimated, (15) generates overidentifying restrictions that can be tested. We do so using Hansen's (1982) J-test. Parameter restrictions were tested using the Wald statistic discussed in Eichenbaum, Hansen, and Singleton (1988).

3.1 CHOICE OF INSTRUMENTS

We now discuss our choice of instruments, i.e. the observable analogues to the vector z_t. In principle the vector z_t ought to satisfy two criteria. First, the elements of z_t should be "exogenous" in the sense that they are uncorrelated with the growth rate of technology. Second, they should be correlated with economic activity in the industry under consideration, i.e., they ought to be relevant. Finding instruments that satisfy both of these criteria is difficult. Different variables have been used in the literature. Burnside, Eichenbaum, and Rebelo (1993) use quarterly innovations to government consumption. Caballero and Lyons (1992), Basu (1993), Bartlesman, Caballero, and Lyons (1994), Basu and Fernald (1994a,b), and Burnside (1994) employ variants of the instruments used by Hall (1988) and Ramey (1989). These consist of current and/or lagged values of the annual growth rates of oil prices and real military expenditures as well as the political party of the President.

Shea (1993a,b) has criticized the last two of the Hall–Ramey instruments on the grounds that they are not relevant. To make this point, Shea (1993a) regressed the growth rate of industrial production in 20 manufacturing industries on a time trend, seasonal dummies, and current and four lagged values of real military spending, using quarterly, seasonally unadjusted, data over the period 1958–1985. He found that military spending is not statistically relevant for output in *any* of the industries he looked at. Similar results hold for the political party of the President. According to Shea, results based on irrelevant instruments should not be viewed as "better" than ordinary least-squares estimates. This line of reasoning may provide an additional rationale for the maintained assumption in Shapiro (1993b) that, over the sample period, 1978–1988, there were no aggregate technology shocks. Bernanke and Parkinson (1991) make the same assumption using quarterly data over the interwar period to justify the use of ordinary least squares.

In this paper we utilize a different set of instruments. While we do report results for Hall–Ramey-type instruments, we also use as instruments lags of the monetary policy shock measures discussed in Christiano, Eichenbaum, and Evans (1994). These shock measures are particularly attractive in the present context because they are, by construction, orthogonal to a large set of economic aggregates in the monetary authority's reaction function. Specifically, Christiano, Eichenbaum, and Evans (1994) identify monetary policy shocks with the disturbance term in a regression equation of the form

$$S_t = \psi(\Omega_t) + \epsilon_{st}. \tag{16}$$

Here S_t is the policy instrument of the monetary authority, ψ is a linear function, Ω_t is the information set available to the monetary authority, and ϵ_{st} is a serially uncorrelated shock that is orthogonal to the elements of Ω_t. To rationalize interpreting ϵ_{st} as an exogenous policy shock, (16) must be viewed as the monetary authority's rule for setting S_t. In addition, the orthogonality conditions on ϵ_{st} correspond to the assumption that date t policy shocks do not affect the elements of Ω_t. Christiano, Eichenbaum, and Evans (1994) derive two measures of policy shocks. These correspond to different specifications of S_t. In both cases Ω_t is given by

$$\Omega_t = \{Q_t, P_t, \text{PCOM}_t, Q_{t-\tau}, P_{t-\tau}, \text{PCOM}_{t-\tau}, \text{FF}_{t-\tau}, \text{NBR}_{t-\tau}, \text{TR}_{t-\tau} : \tau = 1, \ldots, 4.\}.$$

Here Q_t, P_t, PCOM_t, FF_t, NBR_t, and TR_t denote the time t value of the log of real GDP, the log of the GDP deflator, the log of an index of sensitive commodity prices, the federal funds rate, the log of nonborrowed reserves, and the log of time t total reserves, respectively. The two measures of S_t are the log level of nonborrowed reserves and the federal funds rate. The corresponding policy-shock measures, denoted by ϵ_{NBR_t} and ϵ_{FF_t}, correspond to the residual from the OLS regression of the corresponding measure of S_t on Ω_t, i.e., they are the time t components of S_t that are orthogonal to the elements of Ω_t.

To see why policy shocks are useful instruments in our context, consider the vector v_t consisting of $\epsilon_{\mathrm{NBR}_t}$ and ϵ_{FF_t}. It follows that v_t satisfies $E[v_t|\Omega_t] = 0$. We assume that the time $t - \tau$ technology shock for industry i, $\epsilon_{i\,t-\tau}$, lies in the space spanned by the elements of Ω_t, for all $\tau \geq 0$, so that

$$E[v_t\epsilon_{i\,t-\tau}] = 0, \qquad \tau \geq 0. \tag{17}$$

Among other things, the statement that ϵ_{it} lies in Ω_t embodies the assumption that Christiano, Eichenbaum, and Evans (1994) include enough contemporaneous information in the Fed's reaction function so that what they call a policy shock is not in part a reaction to current technology shocks. Under our assumptions it is also true that

$$E[\epsilon_{it}v_{t-\tau}] = 0, \tag{18}$$

for all $\tau > 0$. The simplest way to see this is to suppose that ϵ_{it} has an (invertible) infinite ordered moving-average representation $\epsilon_{it} = a(L)u_{it}$ where $E[u_{it}|\Omega_{t-1}] = 0$ and $a(L)$ is a square-summable polynomial in the lag operator L. Then

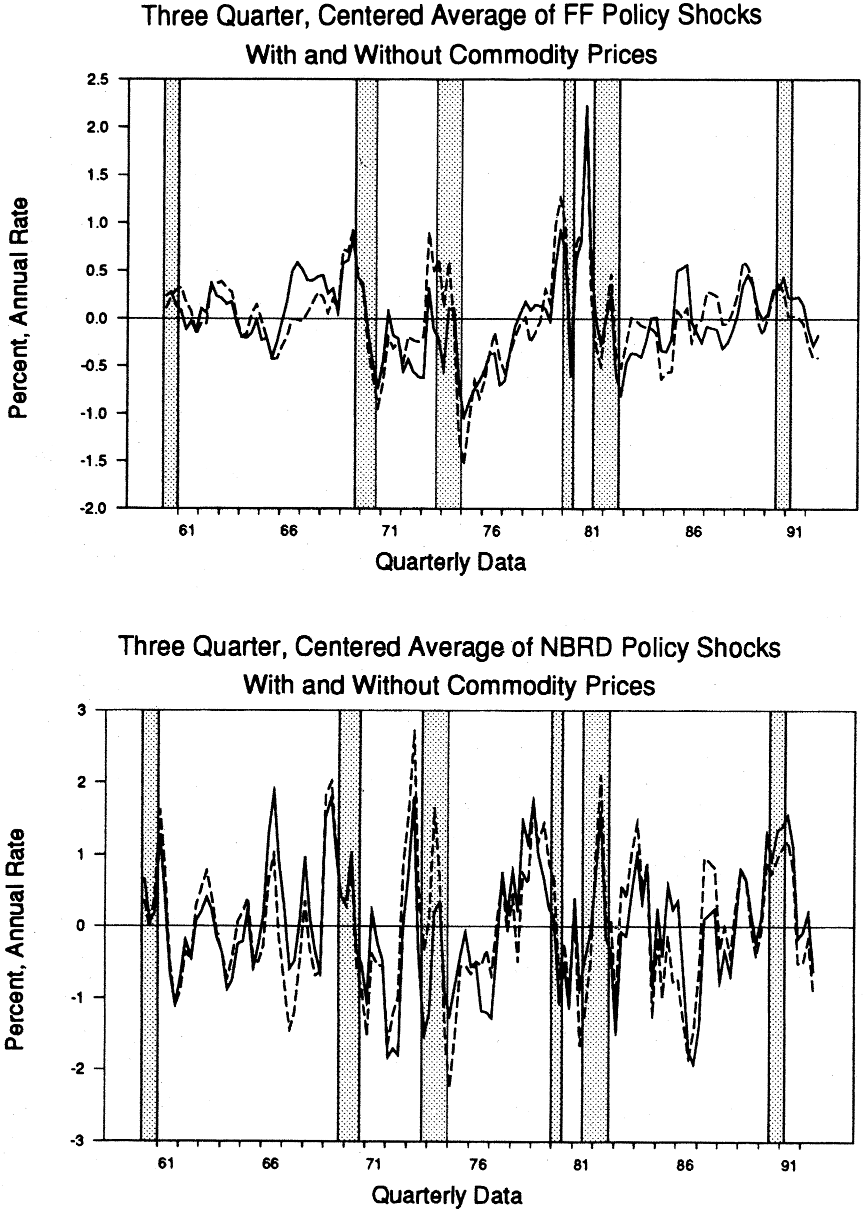
$$E[\epsilon_{it}v_{t-\tau}] = E[(a_0u_{it} + a_1u_{i\,t-1} + \cdots + a_{\tau-1}u_{i\,t-(\tau-1)})v_{t-\tau}] + E[(a_\tau u_{i\,t-\tau} + \cdots)v_{t-\tau}].$$

That the first term on the right-hand side of this expression is zero follows from $E[u_{it}|\Omega_{t-1}] = 0$. That the second term equals zero follows from (17) and the invariability of $a(L)$. Consequently, (18) holds, so that instrument vectors z_t that include current and lagged values of v_t satisfy identifying assumption (15).

An alternative way to rationalize the use of these instruments is to assume that ϵ_{it} is an exogenous stochastic process that has an MA(q) time-series representation. Then it is appropriate to use $v_{t-\tau}$, $\tau > q$, as instruments.

The solid lines in Figure 1, reproduced from Christiano, Eichenbaum, and Evans (1994), depict the estimated time series of $\epsilon_{\mathrm{NBR}_t}$ and ϵ_{FF_t}. Since the policy shock measures are, by construction, serially uncorrelated, they tend to be somewhat noisy. For ease of interpretation we display the centered three-quarter moving average of the shocks. Also, for convenience we include shaded regions, which begin at National Bureau of Economic Research (NBER) business-cycle peaks and end at troughs. The estimated standard deviation of ϵ_{FF_t} is 0.79%, at an annual rate, while

Figure 1 TIME SERIES OF ϵ_{FF_t} AND ϵ_{NBR_t}



For the solid lines, the policy shocks are estimated as the orthogonalized innovations from the six-variable VARs, which include Y, P, PCOM, FF, NBRD, and TR; for the dashed lines, the policy shocks are estimated as the orthogonalized innovations from the five-variable VARs, which include Y, P, FF, NBRD, and TR. In each case, the three-quarter centered averages are computed with equal weights applied to the time $t-1$, t, and $t+1$ orthogonalized innovations.

the standard deviation of $\epsilon_{\mathrm{NBR}_t}$ is 1.61%. The two monetary policy shock measures have a correlation of 0.49.

When we work with quarterly data we consider two specifications of the instrument vector z_t. The benchmark specification of z_t is given by

$$z_{1t} = \{1, \Delta p_{o,t-1-\tau}, \epsilon_{\mathrm{NBR}_{t-2-\tau}}, \epsilon_{\mathrm{FF}_{t-2-\tau}}, \tau = 0, \ldots, 3\}.$$

Here Δp_{ot} denotes the growth rate in the price of oil. We lagged the policy shock measures by two quarters in an attempt to mitigate any spurious correlation between z_{1t} and ϵ_{it} that might arise because of misspecification in the monetary authority's information set. In practice our results were robust to this correction. Our second specification of z_t is given by

$$z_{2t} = \{1, \Delta p_{o,t-\tau}, \Delta g_{t-\tau}, \tau = 0, \ldots, 7\}.$$

Here Δg_t denotes the time t growth rate in military expenditures. We think of z_{2t} as corresponding to the Hall–Ramey instruments. In practice, we measure p_{0t} using the quarterly average of the monthly producer price index of crude petroleum (CITIBASE acronym PW561). We measure g_t as real federal government purchases for national defense (CITIBASE acronym GGFENQ).

When we work with annual data, we choose as our instruments (1) a constant, (2) the current and lagged annual growth rate of the price of oil, and (3) $\epsilon^A_{\mathrm{NBR}_{t-1}}$ and $\epsilon^A_{\mathrm{FF}_{t-1}}$, which are four-by-one vectors containing the quarterly NBR- and FF-based policy shock measures from the year $t-1$. We use shock measures that are lagged by a full year to insure that the instruments do not contain information based on current input or output data.

To investigate the relevance of our instruments, we regress the growth rate of output, the growth rate of hours worked, and the growth rate of electricity consumption on three sets of instruments: (1) z_{1t}, (2) $\{\Delta p_{o,t-1-\tau}, \tau = 0, \ldots, 3\}$, and (3) $\{1, \Delta p_{o,t-\tau}, \Delta g_{t-\tau}, \tau = 0, \ldots, 3\}$. In each case the regression was calculated using data from the aggregate manufacturing sector, the aggregate durable-goods sector, and the aggregate nondurable-goods sector. Table 1 reports the R^2 associated with these regressions. Notice that the R^2 associated with the Hall–Ramey instruments are quite low. They range from a low of 0.03 when the growth rate of electricity consumption in the durables sector is used as the dependent variable to a high of 0.10 when we used the growth rate of output in the manufacturing or durable goods sectors as the dependent variable. Comparable R^2's emerge with $\{\Delta p_{o,t-1-\tau}, \tau = 0, \ldots, 3\}$. In contrast, our benchmark instrument list does much better. Here the

Table 1 R^2 OF INSTRUMENT LISTS WITH OUTPUT AND INPUTS

Sector	*Output*	*Hours*	*Electricity*
Hall Instruments[a]			
Manufacturing	0.10	0.09	0.05
Durables	0.10	0.09	0.03
Nondurables	0.09	0.09	0.08
Growth Rate of Oil Price: Lags 0–3			
Manufacturing	0.07	0.06	0.02
Durables	0.07	0.05	0.00
Nondurables	0.07	0.08	0.03
Benchmark Instruments[b]			
Manufacturing	0.42	0.38	0.34
Durables	0.40	0.39	0.34
Nondurables	0.36	0.29	0.24

[a]Growth rate of oil price, lags 0–3; growth rate of military spending, lags 0–3.
[b]Growth rate of oil price, lags 1–4; ϵ_{FF_t} shock, lags 3–6; $\epsilon_{\mathrm{NBR}_t}$ shock, lags 3–6.

R^2's range from a low of 0.24 when we used the growth rate of electricity consumption in the nondurables sector as the dependent variable to a high of 0.42 when we used the growth rate of output in the manufacturing sector. Evidently, lagged values of $\epsilon_{\mathrm{NBR}_t}$ and ϵ_{FF_t} contain substantial amounts of information regarding the different measures of economic activity that we consider, i.e., they are relevant.

In general, the asymptotic distribution of the technology parameters is affected by the fact that $\epsilon_{\mathrm{NBR}_t}$ and ϵ_{FF_t} are generated regressors. However, this is not the case in our application as long as the growth rate in technology is an exogenous MA(q) process and z_{1t} includes only estimated values of $\epsilon_{\mathrm{NBR}_t}$ and ϵ_{FF_t} that are lagged by at least q periods.[10] To see this write regression equation (16) as

$$Z_t = \beta' X_t + \epsilon_{St}.$$

Denote the estimated values of ϵ_{St} as $\hat{\epsilon}_{St}$. Consider a vector of instruments, $\hat{V}_{t-\tau}$, that includes values of $\hat{\epsilon}_{St}$, lagged at least $\tau > q$ periods. For simplicity's sake, we consider the case in which the number of instruments equals the number of parameters to be estimated. Suppose that we estimate the parameters γ in the relationship

$$W_t = \gamma' D_t + e_t$$

10. We thank Mark Watson for pointing this out to us.

via an instrumental variables procedure that imposes the orthogonality restrictions

$$E[\hat{V}_{t-\tau}e_t] = 0.$$

Then

$$\sqrt{T}(\hat{\gamma} - \gamma)' = \frac{1/\sqrt{T}\Sigma_{t=1}^{T} e_t\hat{V}'_{t-\tau}}{1/T\Sigma_{t=1}^{T} D_t\hat{V}'_{t-\tau}}.$$

Since $\hat{V}_{t-\tau} = V_{t-\tau} + (\hat{\beta} - \beta)'X_{t-\tau}$, it follows that

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{T} e_t\hat{V}'_{t-\tau} = \frac{1}{\sqrt{T}}\sum_{t=1}^{T} e_tV'_{t-\tau} + \sqrt{T}(\hat{\beta} - \beta)'\left(\frac{1}{T}\sum_{t=1}^{T} X_{t-\tau}e_t\right).$$

As long as $\tau > q$, $T^{-1}\Sigma_{t=1}^{T} X_{t-\tau}e_t$ converges in probability to zero. Next note that $T^{-1}\Sigma_{t=1}^{T} D_t\hat{V}'_{t-\tau} = T^{-1}\Sigma_{t=1}^{T}[D_tV'_{t-\tau}] + (\hat{\beta} - \beta)T^{-1}\Sigma_{t=1}^{T}D_tX_{t-\tau}$, so that $T^{-1}\Sigma_{t=1}^{T} D_t\hat{V}'_{t-\tau}$ converges in probability to the same matrix as $T^{-1}\Sigma_{t=1}^{T} D_tV'_{t-\tau}$. It follows that the asymptotic distribution of $\sqrt{T}(\hat{\gamma} - \gamma)'$ is unaffected by the fact that we must estimate $V_{t-\tau}$.

3.2 DATA

Our empirical work utilizes data from a variety of sources. All data referred to in this subsection are seasonally adjusted. We indicate the CITIBASE acronym for each variable in brackets.

3.2.1 Economywide Input and Output Data In some of our empirical work we employ economywide aggregates. Here our measure of output is quarterly real GDP (GNPQ) over the sample period 1972:2–1992:4. Our measure of hours worked is the quarterly average of monthly total employee-hours in nonagricultural establishments (LPMHU). We considered two measures of the quarterly growth rate in the real capital stock. The first is taken from Hall (1994). The second is an updated version (available only through 1988:4) of the measure discussed in Christiano (1988).[11] Our measure of aggregate electricity consumption is a quarterly average of a monthly index of total electrical power usage in the indus-

11. We thank Jonas Fisher for making these data available to us.

trial sector (manufacturing plus mining plus utility industries).[12] When dealing with economywide aggregates, we measure the relative price of electricity using the quarterly average of the producer price index for electric power (PW054) and quarterly compensation per hour in the nonfarm business sector (LBCPU).

3.2.2 Manufacturing-Sector Input and Output Data We measure quarterly labor input at the two-digit SIC level using quarterly averages of monthly production worker hours. For each two-digit industry this measure is constructed as the product of two time series: average weekly hours of production workers (LPHRXX) and production workers on nonagricultural payrolls (LPPXX). Here *XX* refers to the relevant two-digit SIC code. For aggregate manufacturing it is also possible to obtain data on a broader measure of labor input: total hours, of all persons, worked by all employees (LMNM). This broader measure of labor input is not available at the two-digit SIC level. To justify abstracting from nonproduction workers on the basis of the simple model of Section 2, we need to assume that their input is used in fixed proportions with value added. If this Leontief assumption does not hold, the interpretation of our results continues to be valid only if the correlation between nonproduction hours and production hours is one.

Annual labor input measures correspond to the annual averages of the monthly data. All of the data are available over the period 1972:1–1992:4. Corresponding three-digit level data for the sample period 1977:1–1992:4 were obtained from the Board of Governors.

Electricity consumption was measured as kilowatts of electricity used at the two-digit SIC level. These data were obtained from the Board of Governors. The two-digit SIC-level data are available over the period 1972:2–1992:4, while the three-digit SIC-level data are available over the sample period 1977:1–1992:4. Quarterly and annual data correspond to averages of the underlying monthly data.

Obtaining quarterly measures of industry-level output is more difficult than obtaining the corresponding input measures. The Federal Reserve Board uses three sources of data to construct the industrial production index: measures of physical product, kilowatt-hours of electricity, and production worker hours. The weight on each of these underlying sources of information depends on the industry in question. Averaging over all two-digit SIC manufacturing industries, roughly 43%, 31%, and

12. We thank Joe Beaulieu for making these data available to us in machine-readable form. These raw data are published on a monthly basis in *Industrial Production*, Federal Reserve Statistical Release G.12.3.

26% of the output index is based on data on measures of physical product, kilowatt-hours, and production worker hours, respectively. Note that the Board does not use a simple, mechanical rule for inferring output from inputs. Instead it estimates output using time-varying production-factor coefficients. If we conceive of the Board as producing an optimal prediction of output given the information at its disposal, it is reasonable to use the Board data on output.[13] Still, we would be nervous about basing inference entirely on this data set.

Fortunately, there are a number of ways to assess the robustness of our results to the use of alternative data sources. First, we exploit the fact that there are many three-digit SIC industries where the output index produced by the Board is strictly based on physical product. We constructed a database with the subset of these three-digit industries for which we could obtain matching labor input and electricity use over the period 1977:1–1992:4. The net result was a panel of 26 three-digit SIC industries. These are listed in the Appendix, along with the three-digit SIC codes and the corresponding two-digit SIC industries. Second, we repeat our analysis using annual data. At the annual frequency, the Board's measure of output is not based on input data. This is because data from various censuses provide actual production data for most industries. Therefore the problem of inferring output from inputs is almost entirely an issue for within-year variation of industrial output.

4. *Empirical Results*

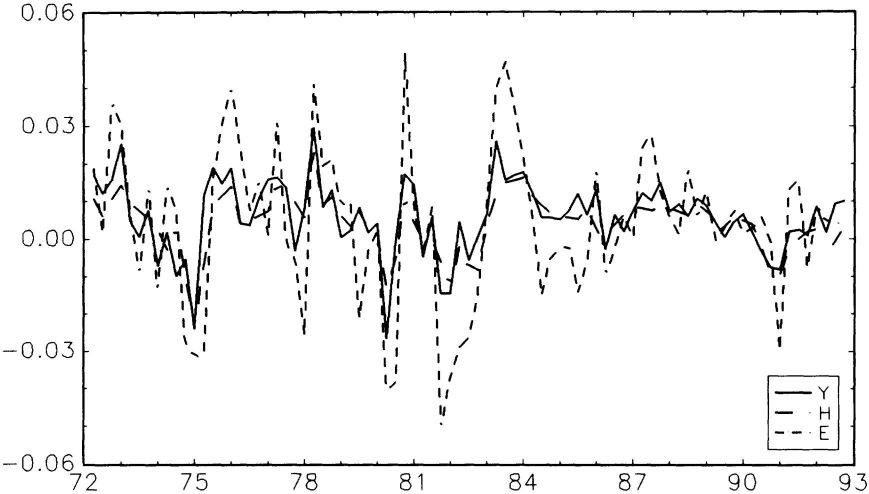
4.1 SOME BASIC FACTS

We begin our analysis with a brief review of some basic facts. Figure 2 displays the quarterly growth rates of real GDP (Δy_t), economywide hours worked (Δl_t), and aggregate industrial electricity consumption (Δe_t). It is clear that Δe_t is highly correlated with both Δy_t and Δl_t, even though at the aggregate level it is difficult to obtain a measure of electricity consumption that matches the output concept. The high correlation between these aggregates is documented in the following table, which presents the unconditional correlations among Δy_t, Δl_t, Δe_t, and the growth rate in our measure of capital, Δk_t. In contrast to Δe_t, Δk_t is basically uncorrelated with Δy_t and Δl_t (as well as Δe_t).[14] This is why

13. See Miron and Zeldes (1989) for a discussion of different models of measurement error in this context.

14. This is also true if we redo the analysis over the sample period 1972:1–1988:4 using the measure of capital discussed in Christiano (1988).

Figure 2 QUARTERLY GROWTH RATES OF ECONOMYWIDE DATA



Y represents real GNP, H represents total hours worked in nonagricultural establishments, and E represents electrical power usage in the industrial sector. All series are plotted as first-differenced logarithms. The data are described in more detail in the text.

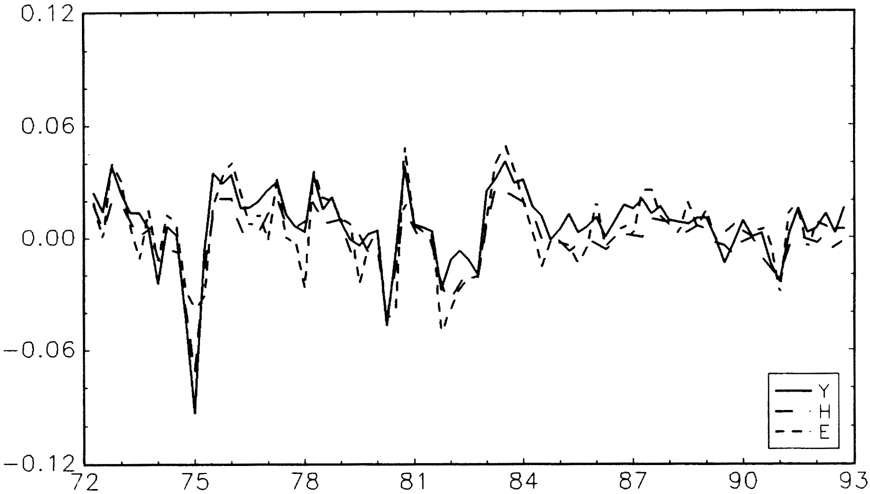
analysts have traditionally found that capital plays no role in explaining cyclical fluctuations in output—existing measures of capital are poor measures of capital services, at least at cyclical frequencies:

Correlations: Economywide

	Δy_t	Δl_t	Δe_t	Δk_t
Δy_t	1.00	.82	.72	.09
Δh_t	.82	1.00	.73	.31
Δl_t	.72	.73	1.00	.07
Δk_t	.09	.31	.07	1.00

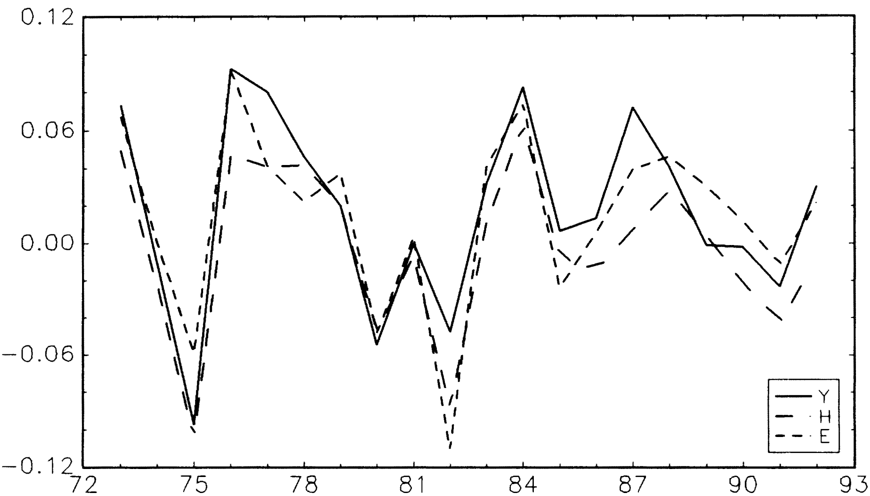
Figures 3 and 4 are the analogues to Figure 2 except that they are based on quarterly and annual manufacturing data. Figures 5 and 6 display, in a graphical manner, the quarterly and annual correlations between (Δy_t and Δl_t), (Δy_t and Δe_t), and (Δl_t and Δe_t) for the individual two-digit SIC industries underlying the aggregate manufacturing data. The following table summarizes the correlations between Δy_t, Δl_t, and Δe_t for the manufacturing sector as a whole:

Figure 3 QUARTERLY GROWTH RATES OF AGGREGATE MANUFACTURING DATA



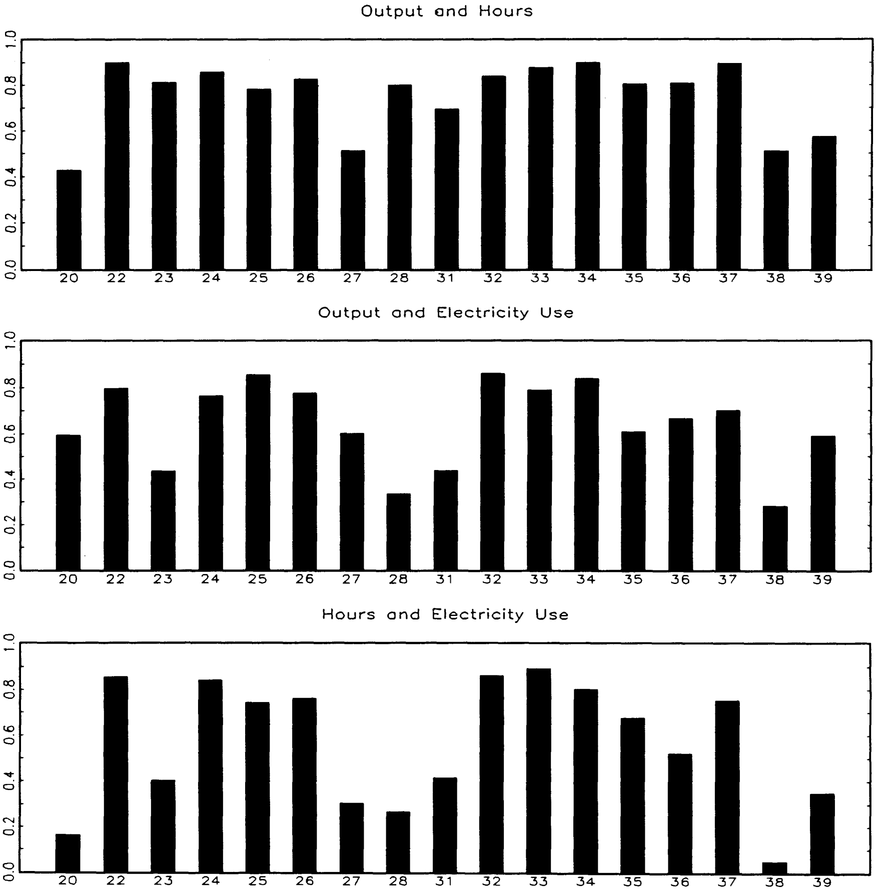
Y represents industrial production in the manufacturing sector, *H* represents total employee hours in the manufacturing sector, and *E* represents electrical power usage in the manufacturing sector. All series are plotted as first-differenced logarithms. The data are described in more detail in the text.

Figure 4 ANNUAL GROWTH RATES OF AGGREGATE MANUFACTURING DATA



Y represents gross output in the manufacturing sector, *H* represents total employee hours in the manufacturing sector, and *E* represents electrical power usage in the manufacturing sector. All series are plotted as first-differenced logarithms. The data are described in more detail in the text.

Figure 5 CORRELATIONS OF QUARTERLY TWO-DIGIT SIC LEVEL DATA

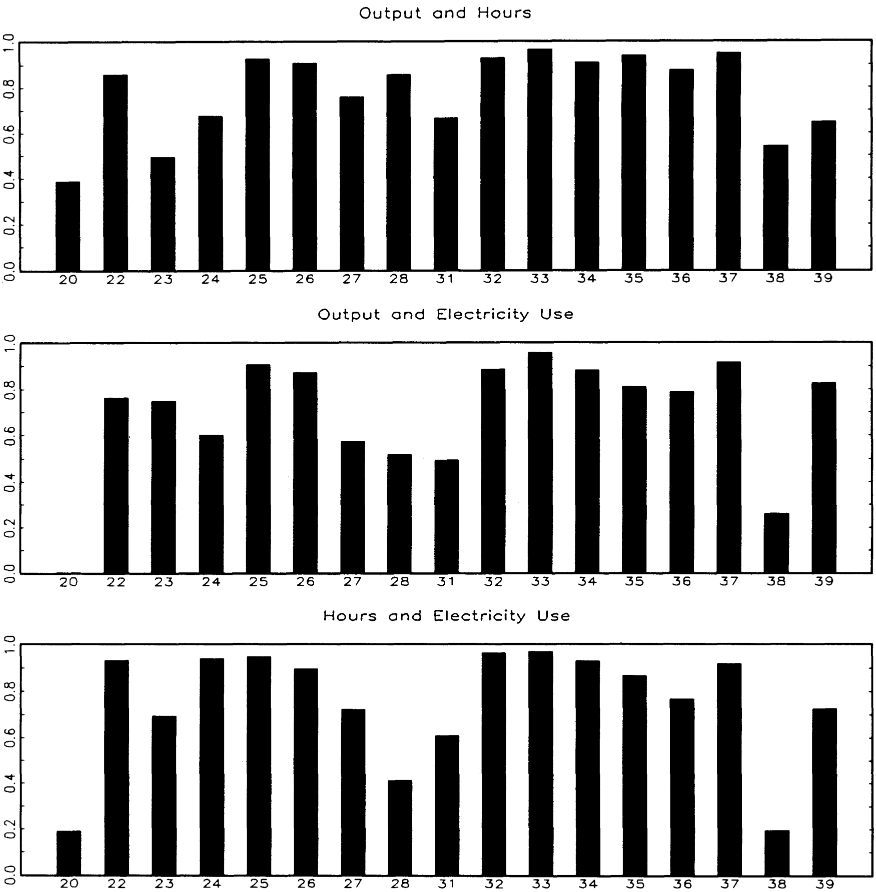


Y represents industrial production, *H* represents production worker hours, and *E* represents electrical power usage. All series are first-differenced logarithms. The *x*-axis labels are the SIC codes.

Correlations: Manufacturing Sector

	Quarterly			*Annual*		
	Δy_t	Δl_t	Δe_t	Δy_t	Δl_t	Δe_t
Δy_t	1	.94	.80	1	.95	.88
Δl_t	.94	1	.81	.95	1	.94
Δe_t	.80	.81	1	.88	.94	1

Figure 6 CORRELATIONS OF ANNUAL TWO-DIGIT SIC LEVEL DATA

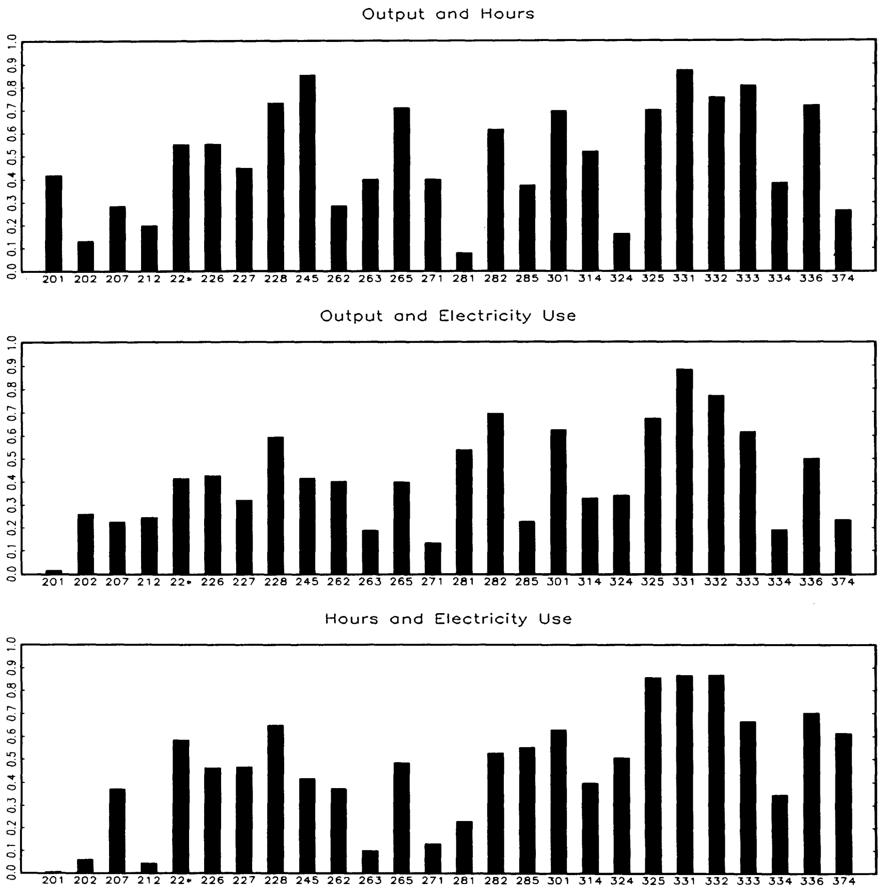


Y represents industrial production, H represents production worker hours, and E represents electrical power usage. All series are first-differenced logarithms. The x-axis labels are the SIC codes. The correlation between Y and E for industry 20 is -0.19.

A number of points are worth making here. First, as in the aggregate data, Δe_t is highly correlated with Δy_t and Δl_t. Indeed, the correlation is even more pronounced in the manufacturing data. This may reflect the fact that our measure of e_t corresponds exactly to the manufacturing sector. Second, the quarterly and annual correlations are very similar. If anything, Δe_t and Δl_t are slightly less correlated with output at the quarterly level. This is very comforting, given possible concerns about the use of input data in the procedure used by the Board to construct some of the quarterly output data. Recall that while these concerns are relevant for the quarterly data, they are not relevant for the annual data. So the basic fact which drives our inferences—namely that Δe_t comoves positively with Δy_t and Δl_t—cannot be dismissed as an artifact of the way the output data are constructed.

A different way to see this is to consider the correlations between (Δy_t and Δl_t), (Δy_t and Δe_t), and (Δl_t and Δe_t) for the individual three-digit SIC-code industries where the Board's measure of output data is not constructed with the aid of any input data. These are displayed in Figure 7. Notice that while there are interesting differences among the industries,

Figure 7 CORRELATIONS OF QUARTERLY THREE-DIGIT SIC LEVEL DATA



Y represents industrial production, *H* represents production worker hours, and *E* represents electrical power usage. All series are first-differenced logarithms. The *x*-axis labels are the SIC codes. Industry 22* is industries 221 and 222 combined. The correlation between *H* and *E* for industry 201 is −0.03.

in the vast majority of cases Δe_t displays a sharp positive correlation with Δy_t and Δl_t.

Next we consider the cyclical properties of a different measure of capital services: the workweek of capital, wk_t. Shapiro (1993b), among others, has suggested that a measure of wk_t might be useful in correcting capital stock data for cyclical variations in capital services. To pursue this point, we obtained the measure of wk_t used by Shapiro (1993b). This consists of an updated version of the series published by Foss (1981). The data are annual, with each observation corresponding to the fourth-quarter workweek of capital. The sample period is 1976:4–1988:4. The following table summarizes the correlations among Δy_t, Δl_t, Δe_t, and Δwk_t:[15]

MEASURED WORKWEEK OF CAPITAL

	Δy_t	Δl_t	Δe_t	Δwk_t
Δy_t	1	.95	.88	.88
Δl_t	.95	1	.90	.89
Δe_t	.88	.90	1	.74
Δwk_t	.88	.89	.74	1

Notice that $\Delta \mathrm{wk}_t$ displays a strong positive correlation with Δy_t, Δl_t, and Δe_t. We take this fact to be supportive of our basic hypothesis that capital utilization rates are procyclical.[16]

We conclude this subsection by briefly documenting the apparent "short-run increasing returns to scale" (SRIRL) puzzle. All the results that we report were obtained using the GMM procedure and the instrument list z_{1t} discussed in Section 3. The following table presents the points estimates of η_1 that result from estimating the relationship $\Delta y_t = \eta_0 + \eta_1 \Delta l_t + \epsilon_t$ using aggregate and manufacturing-sector data (with standard errors in parentheses):

RETURNS TO LABOR

	Economywide		Manufacturing			
	Total Hrs.	Prod. Worker Hrs.	Total Hrs.	Prod. Worker Hrs.	Durables[a]	Nondurables[a]
η_1	1.21	0.96	1.25	0.97	0.92	0.98
	(0.13)	(0.09)	(0.08)	(0.06)	(0.06)	(0.10)

[a]Production worker hours.

15. All growth rates were calculated on a fourth-quarter-to-fourth-quarter basis.
16. We also computed the correlations between the growth rate of total capital services ($k_t \cdot \mathrm{wk}_t$) and (Δy_t, Δl_t, Δe_t). These are similar to the ones between wk_t and (Δy_t, Δl_t, Δe_t).

Notice that when total hours worked are used to construct Δl_t, η_1 is estimated to be significantly greater than one. When production worker hours are used, η_1 is estimated to be approximately one. This is true regardless of whether we work with aggregate data, manufacturing data, durable-goods data, or nondurable-goods data. SRIRL appears to be alive and well, even with our instruments.

The following table presents the point estimates of η_2 that result from estimating the relationship $\Delta y_t = \eta_0 + \eta_2 \Delta e_t + \epsilon_t$:

RETURNS TO ELECTRICITY

		Manufacturing		
	Economywide	*Total*	*Durables*	*Nondurables*
η_2	0.49 (0.07)	1.15 (0.14)	0.83 (0.09)	0.92 (0.19)

Notice that, for the manufacturing sector, measuring factor input by electricity alone or hours alone yields very similar results. Indeed, we even get "short-run increasing returns to electricity." The estimated value of η_2 is positive but smaller for the economywide case. Presumably this reflects the fact that we do not have as good a measure of electricity use for the economywide data.

4.2 CES VERSUS LEONTIEF

The previous subsection documented the basic fact that the growth rate of electricity consumption is highly correlated with the growth rates of hours worked and output. We now consider how this fact affects technology parameter estimates. Table 2 reports the results of estimating the parameters of the technology specification given by (6)–(7), which allows for substitution between capital services and electricity. The first column presents economywide results, while the second column presents results pertaining to the total manufacturing sector. The third column presents results obtained imposing the restriction that the technology parameters are the same in all two-digit SIC industries. The fourth and fifth columns are analogues to the third column that pertain to the durable- and nondurable-goods industries. The row labeled *J* reports the probability value associated with the statistic for testing the overidentifying restrictions of the model. The last two rows report different statistics pertaining to the average "technology shock" ϵ_t. For restricted panel runs, the reported statistic regarding ϵ_t pertains to the average value of the industry-specific statistic. For example, σ_ϵ corre-

Table 2 CES SPECIFICATIONS

		Manufacturing Sector			
			Two-Digit SIC Code Level[a]		
	Economywide	*Aggregate*	*All Industries*	*Durables*	*Nondurables*
α_1	0.74 (0.50)	0.71 (0.34)	0.76 (0.09)	0.63 (0.13)	0.90 (0.15)
α_2	0.24 (0.17)	0.30 (0.30)	0.27 (0.08)	0.41 (0.12)	0.27 (0.12)
σ	0.15 (0.35)	0.03 (0.44)	0.26 (0.17)	0.04 (0.20)	0.30 (0.22)
$\alpha_1 + \alpha_2$	0.98 (0.34)	1.00 (0.08)	1.03 (0.04)	1.04 (0.05)	1.17 (0.08)
J	0.92	0.004	0.34	0.11	0.33
$\sigma_\epsilon/\sigma_{\Delta Y}$	0.60	0.37	0.63	0.60	0.67
$\rho_{\epsilon\Delta Y}$	0.32	0.21	0.54	0.42	0.56

[a]Coefficients restricted across industries, with industry fixed effects.

sponds to the average value of the standard deviation of ϵ_t across the different industries.

The key result is that across all of the cases considered, the estimated value of σ, the elasticity of substitution between the workweek of capital and electricity, is positive but very small. Specifically, it ranges from a low of 0.03 for the aggregate manufacturing sector to a high of 0.30 for the nondurable-goods sector. In no case can we reject the null hypothesis that $\sigma = 0$. This case corresponds to the Leontief specification given by (3).[17]

A different way to assess the Leontief specification is to investigate the empirical relationship between the growth rate of electricity and the growth rate of capital services, as measured by the growth rate of the product of the workweek of capital ($\Delta \mathrm{wk}_t$) and the stock of capital (k_t). The following table reports the results of estimating the relationship

$$\Delta e_t = \beta[\Delta \mathrm{wk}_t + \Delta k_t]$$

17. It is interesting to contrast the restricted point estimates of α_1, α_2, and σ in the manufacturing industries (0.71, 0.30, and 0.03) with the unrestricted point estimates for the underlying industries. One way to summarize the unrestricted estimates is to focus on their median. The median point estimates of α_1, α_2, and σ are 0.60, 0.38, and 0.16. The associated median standard errors are 0.37, 0.35, and 0.55. Evidently, the qualitative nature of inference here is not affected by imposing the (false) restriction that the two-digit industries have the same technology coefficients.

using our three-stage least-squares procedure in conjunction with instrument list z_{1t}:

	Aggregate Manufacturing	*Aggregate Durables*	*Aggregate Nondurables*
β	1.23	1.77	0.53
	(0.55)	(0.77)	(0.35)

Notice that in no case can we reject the null hypothesis $\beta = 1$. In the light of this result and our previous findings regarding σ, through much of what follows we impose the restriction that electricity use is proportional to the workweek of capital. Table 2 contains results generated without imposing that restriction, so the reader can verify that none of the conclusions discussed in the text are affected by the imposition of that restriction.

In the remainder of this section we address five key questions: (1) Does SRIRL vanish once capital services are measured by electricity consumption? (2) Are capital services productive when measured by electricity consumption? (3) Is there evidence against the hypothesis of constant returns to scale? (4) Is there evidence against the overidentifying restrictions of our model? (5) What can we say about the properties of technology shocks? We address these questions at three levels of aggregation: economywide data, two-digit SIC-code level data and three-digit SIC-code level data.

4.3 ECONOMYWIDE DATA

Table 3 reports the results of estimating the model using economywide data. The first column reports results obtained using two different measures of the capital stock. The third column reports results obtained measuring capital services by electricity consumption. A number of results emerge here. First, when we use the capital-stock data, SRIRL appears, i.e., α_1 is estimated to be greater than one. In addition, the estimated value of α_2 is negative and insignificantly different from zero. In sharp contrast, when we measure capital services by electricity use, the SRIRL phenomenon disappears and capital services enter significantly into the production technology. Second, there is no evidence against the hypothesis of constant returns to scale. Finally, according to the statistic J there is no evidence against the model's overidentifying restrictions.

Using electricity consumption as a measure of capital services has important implications for the statistical properties of the technology shocks. As a benchmark, suppose we simply set $\alpha_1 = 0.64$ and $\alpha_2 = 0.36$.

Table 3 ECONOMYWIDE DATA: $\Delta Y_t = \alpha_0 + \alpha_1 \Delta H_t + \alpha_2 \Delta K_t^* + \epsilon_t$

	$\Delta K^* = \Delta K^H$	$\Delta K^* = \Delta K^C$	$\Delta K^* = \Delta e$
α_1	1.23	1.31	0.54
	(0.14)	(0.24)	(0.27)
α_2	−0.32	−0.88	0.30
	(0.85)	(1.81)	(0.11)
$\alpha_1 + \alpha_2$	0.91	0.43	0.84
	(0.80)	(1.61)	(0.19)
J	0.91	0.72	0.41
$\sigma_\epsilon/\sigma_{\Delta Y}$	0.56	0.56	0.60
$\rho_{\epsilon\Delta Y}$	0.38	0.39	0.31

K^H: Hall (1994) measure of capital.
K^C: Christiano (1988) measure of capital: 72:1–88:4.

Using the stock of capital and electricity consumption as measures of capital services, we obtain

	$\Delta S_t = \Delta K_t$	$\Delta S_t = \Delta e_t$
J	.015	.42
$\sigma_\epsilon/\sigma_{\Delta Y}$	.67	.77
$\rho_{\epsilon\,\Delta Y}$	.87	.06

respectively. Notice that with the stock of capital measure, there is substantial evidence against the model's overidentifying restrictions. There is virtually no evidence against these restrictions when capital services are measured using electricity. Perhaps more importantly, with the electricity measure and these parameter values, the technology shocks are virtually uncorrelated with the growth rate of output. Moving to the estimated values of α_1 (0.54) and α_2 (0.30) lowers $\sigma_\epsilon/\sigma_{\Delta y}$ and raises $\rho_{\epsilon\,\Delta y}$ somewhat (see Table 3). But even there, $\rho_{\epsilon\,\Delta y}$ is only equal to .31. This small correlation seems very difficult to reconcile with existing RBC models that are driven primarily by technology shocks. Finally, it is worth emphasizing that our electricity-based technology shocks are much less volatile and substantially less correlated with output than those emerging from the measures of output, hours worked, and stock of capital that are typically used in the RBC literature.[18]

18. For example, suppose we use Christiano's (1988) measure of capital, hours worked, and output. In addition set $\alpha_1 = 0.655$ and $\alpha_2 = 0.345$, the values estimated in Christiano and Eichenbaum (1992). The resulting point estimates of σ_ϵ and $\sigma_\epsilon/\sigma_{\Delta y}$ are 0.0114

4.4 MANUFACTURING-SECTOR DATA

Table 4 reports results based on the two-digit SIC data. Columns labeled "Aggregate" pertain to aggregate manufacturing data, and columns labeled "Restricted" refer to results obtained using the panel on two-digit SIC industries. Results are reported for both quarterly and annual data.

Consider the quarterly results. First, for aggregate manufacturing, the point estimates of α_1 and α_2 are 0.69 and 0.31, respectively. The corresponding restricted panel point estimates are 0.64 and 0.37.[19] These estimates are remarkably close to national-income-based estimates of labor and capital shares obtained using a constant-returns-to-scale Cobb–Douglas production function (see for example Christiano and Eichenbaum, 1992). Second, the standard errors of $\alpha_1 + \alpha_2$ reveal virtually no evidence against the hypothesis of constant returns to scale. Third, the overidentifying restrictions associated with the aggregate model can be rejected at the 1% significance level. However, there is very little evidence against these restrictions for the restricted panel. Fourth, comparing our economy-wide-based estimates of $\sigma_\epsilon/\sigma_{\Delta Y}$ and $\rho_{\epsilon\,\Delta Y}$ (0.60 and 0.31) with those reported in Table 4 (0.37 and 0.21), we see that these fall as we move to the aggregate manufacturing sector. However, we are hesitant to make much of this fact, because our estimates of $\sigma_\epsilon/\sigma_{\Delta Y}$ and $\rho_{\epsilon\,\Delta Y}$ rise to 0.63 and 0.54, respectively, when we work with the restricted panel data. But even these estimates are smaller than those used in the RBC literature.

The key finding with the annual data is that the results are quite similar to those obtained with the quarterly data. There is some difference in the point estimates associated with the restricted panel.[20] This sensitivity is also revealed in the portion of Table 4 reporting annual results for the durable and nondurable goods sector. This point aside, inference seems robust. Specifically, (1) there is no evidence of SRIRL, (2) there is no evidence against the hypothesis of constant returns to scale, (3) there is little evidence against the overidentifying restrictions of the model, and (4) there is overwhelming evidence that capital services, as measured by electricity, are an important factor of production. The fact that inference is robust to the use of annual data is particularly comforting because annual output data are not constructed using information

and 1.05, respectively. The correlation coefficient between ϵ_t and ΔY_t, $\rho_{\epsilon\,\Delta Y}$, is approximately equal to .80.

19. The median unrestricted point estimates of α_1 and α_2 across the two-digit industries are 0.54 and 0.38. The corresponding median standard errors are 0.20 and 0.22.

20. The median point estimates of α_1 and α_2 obtained using the unrestricted annual two-digit SIC data are 0.80 and 0.17, respectively, with corresponding standard errors of 0.13 and 0.15.

Table 4 MANUFACTURING-SECTOR DATA: $\Delta Y_t = \alpha_0 + \alpha_1 \Delta H_t + \alpha_2 \Delta e_t + \epsilon_t$

	Quarterly				*Annual*			
	Aggregate	*2-Digit SIC Code Level*[a]			*Aggregate*	*2-Digit SIC Code Level*[a]		
	Mfg.	*Mfg.*	*Durable*	*Nondur.*	*Mfg.*	*Mfg.*	*Durable*	*Nondur.*
α_1	0.69	0.64	0.61	0.74	0.69	0.43	0.38	0.60
	(0.16)	(0.05)	(0.06)	(0.09)	(0.31)	(0.02)	(0.06)	(0.12)
α_2	0.31	0.37	0.43	0.39	0.21	0.57	0.68	0.30
	(0.17)	(0.05)	(0.07)	(0.08)	(0.35)	(0.02)	(0.06)	(0.12)
$\alpha_1 + \alpha_2$	1.00	1.01	1.04	1. 13	0.90	1.00	1.06	0.90
	(0.07)	(0.04)	(0.05)	(0.08)	(0.09)	(0.01)	(0.03)	(0.07)
J	0.01	0.31	0.12	0.30	0.12	0.20	0.44	0.09
$\sigma_\epsilon/\sigma_{\Delta Y}$	0.37	0.63	0.60	0.67	0.33	0.61	0.55	0.69
$\rho_{\epsilon\Delta Y}$	0.21	0.54	0.42	0.56	0.35	0.58	0.49	0.71

[a]Coefficients restricted across industries, with industry fixed effects.

on factor inputs. As a further check on the robustness of the two-digit SIC results, the Appendix reports results obtained by omitting two-digit SIC industries in which a particularly large proportion of the output index reported by the Board is based on input data.

4.5 THREE-DIGIT SIC SECTOR DATA

Table 5 reports results obtained using our three-digit SIC data set. Recall that this data set consists of three-digit industries for which there are direct measures of physical output. The columns labeled "Restricted" refer to results generated under the restriction that the coefficients α_1 and α_2 are the same in all of the industries we looked at. The columns labeled "Unrestricted" report results generated from the corresponding unrestricted runs. Specifically, we report the median point estimate of α_1 and α_2 as well as the corresponding median standard errors. In addition we report the median point estimates of $\sigma_\epsilon/\sigma_{\Delta y}$ and $\rho_{\epsilon\Delta y}$. The probability value for the statistic J refers to the overidentifying restrictions associated with the entire system of unrestricted runs.

The key features to note here are as follows. First, as above, there is no evidence for either SRIRL or increasing returns to scale. If anything there is some evidence of decreasing returns, but only for the restricted specification where we do not distinguish between durable and nondurable goods. This specification aside, we find very little evidence against the hypothesis of constant returns to scale. Second, as was the case with our

Table 5 THREE-DIGIT SIC CODE LEVEL DATA: $\Delta Y_t = \alpha_0 + \alpha_1 \Delta H_t + \alpha_2 \Delta e_t + \epsilon_t$

	Restricted[a]			*Unrestricted*[b]		
	Mfg.	*Durable*	*Nondur.*	*Mfg.*	*Durable*	*Nondur.*
α_1	0.52	0.73	0.45	0.52	0.64	0.56
	(0.05)	(0.11)	(0.07)	(0.31)	(0.44)	(0.33)
α_2	0.35	0.14	0.35	0.38	0.24	0.21
	(0.04)	(0.08)	(0.06)	(0.25)	(0.36)	(0.26)
$\alpha_1 + \alpha_2$	0.86	0.87	0.81	0.87	0.89	0.92
	(0.05)	(0.08)	(0.09)	(0.28)	(0.23)	(0.32)
J	0.20	0.19	0.26	0.50	0.15	0.79
$\sigma_\epsilon/\sigma_{\Delta Y}$	0.84	0.75	0.88	0.95	0.70	0.97
$\rho_{\epsilon\Delta Y}$	0.84	0.82	0.86	0.78	0.78	0.76

[a]Coefficients restricted across industries, with industry fixed effects.
[b]Median coefficients across industries, with median standard errors reported in parentheses.

Table 6 HALL INSTRUMENTS: $\Delta Y_t = \alpha_0 + \alpha_1 \Delta H_t + \alpha_2 \Delta e_t + \epsilon_t$

	2-Digit SIC Code		3-Digit SIC Code	
	Restricted[a]	*Unrestricted*[b]	*Restricted*[a]	*Unrestricted*[b]
α_1	0.54	0.61	0.56	0.50
	(0.04)	(0.19)	(0.05)	(0.32)
α_2	0.39	0.29	0.23	0.20
	(0.04)	(0.18)	(0.04)	(0.25)
$\alpha_1 + \alpha_2$	0.93	0.91	0.79	0.79
	(0.04)	(0.20)	(0.05)	(0.28)
J	0.41	0.65	0.20	0.81
$\sigma_\epsilon / \sigma_{\Delta Y}$	0.62	0.58	0.83	0.91
$\rho_{\epsilon \Delta Y}$	0.63	0.61	0.87	0.85

[a]Coefficients restricted across industries, with industry fixed effects.
[b]Median coefficients across industries, with median standard errors reported in parentheses.

data sets, we find a substantial role for capital services, as measured by electricity, in producing output. Third, there is virtually no evidence against the overidentifying restrictions imposed by the model. This is true regardless of whether we work with the entire panel or condition on durable- and nondurable-goods industries. Finally, we find that the estimated values of $\sigma_\epsilon / \sigma_{\Delta y}$ and $\rho_{\epsilon \, \Delta y}$ are somewhat larger than those emerging from the manufacturing and economywide data. Still, these estimates are substantially smaller than those used in the RBC literature. We conclude that the main findings obtained with the aggregate and two-digit SIC data are confirmed by the three-digit SIC data.

We now briefly comment on the results of working with the alternative instrument set, z_{2t}. Table 6 reports a subset of the results we obtained with the Hall–Ramey-type instruments. Specifically, we display results for the restricted two-digit and three-digit SIC panels as well as the median estimates from the corresponding unconstrained specifications. The key point to note is the robustness of our results to the change in instruments.

4.6 THE DIFFERENTIABLE TECHNOLOGY

We conclude this section by reporting results obtained from estimating the returns-to-scale parameter η in the production technology given by (2). We estimated η using three measures of the growth rate of capital services, ΔS_t. These measures are the growth rate in the stock of capital, the growth rate of electricity, and the growth rate in the workweek of

Table 7 DIFFERENTIABLE TECHNOLOGY SPECIFICATION

	Measure of ΔS_t		
	ΔK_t	Δe_t	$\Delta[\mathrm{wk}_t \cdot K_t]$
Aggregate Manufacturing			
η	1.10	1.01	0.98
	(0.06)	(0.07)	(0.05)
J	0.05	0.07	0.27
$\sigma_\varepsilon / \sigma_{\Delta y}$	0.24	0.24	0.17
$\rho_{\varepsilon \Delta y}$	0.17	0.26	0.17
Aggregate Durables			
η	1.16	1.06	1.08
	(0.05)	(0.05)	(0.05)
J	0.35	0.20	0.30
$\sigma_\varepsilon / \sigma_{\Delta y}$	0.18	0.16	0.17
$\rho_{\varepsilon \Delta y}$	−0.08	0.11	0.14
Aggregate Nondurables			
η	0.83	0.82	0.86
	(0.13)	(0.14)	(0.12)
J	0.11	0.13	0.22
$\sigma_\varepsilon / \sigma_{\Delta y}$	0.50	0.52	0.42
$\rho_{\varepsilon \Delta y}$	0.67	0.61	0.43

capital times the stock of capital. The corresponding sample periods, which were dictated by data availability, are 1961–1989, 1973–1989, and 1977–1988, respectively. In all cases the data correspond to fourth-quarter-to-fourth-quarter growth rates. The instrument list is given by z_{1t}. Results for aggregate manufacturing, durable goods, and nondurable goods are reported in Table 7.

The key results can be summarized as follows. First, for aggregate manufacturing and durable goods, the estimated value of η is highest when ΔS_t is measured by ΔK_t. Moving to electricity or workweek of capital-based measures of ΔS_t results in smaller η. With these measures we cannot reject the hypothesis of constant returns to scale. If anything, there is some mild evidence of decreasing returns to scale in the nondurable-goods industries. Third, in all cases the estimated shocks to technology and their correlation with the growth rate of output are much smaller than those used in the RBC literature.

On the whole, we conclude that inference about returns to scale is quite robust across the three specifications of technology that we considered. There just is not much evidence in our data sets against the hypothesis of constant returns to scale.

5. *Shortcomings of the Analysis*

In this section we discuss how the presence of capital goods that do not use electricity, overhead labor and capital, and multiple production shifts could affect the interpretation of our results.

5.1 NONPRODUCTION WORKERS AND OVERHEAD COSTS

So far we have stressed mismeasurement of capital services as the main source for the apparent short-run increasing returns to labor. An alternative and perhaps complementary explanation for this phenomenon is the existence of large overhead costs. To see this, suppose that the production function is of the form

$$Y_t = A_t(L_t - \varphi)^{\alpha_1} K_t^{\alpha_2} \tag{19}$$

Here φ represents overhead hours. An infinitesimal increase in hours worked due to a demand shock generates a change in labor productivity equal to:

$$\frac{d(Y_t/L_t)}{dL_t} = \frac{Y_t}{L^2}[(\alpha_1 - 1)L + \varphi]. \tag{20}$$

Suppose $\alpha_1 < 1$. Then, absent overhead costs ($\varphi = 0$), this derivative is negative, suggesting that labor productivity ought to be countercyclical in a model driven primarily by demand shocks. However, for $\varphi > (1 - \alpha_1)L$, this derivative will be positive. This could, in principle, rationalize procyclical productivity even in a model driven by demand shocks. However, a simple back-of-the-envelope calculation suggests that the required overhead costs must be large. If α_1 is roughly equal to 0.65, then $d(Y_t/L_t)/dL_t$ will be positive only if overhead costs represent 35% of L_t.

Even if overhead costs are not this large, the fact that we have neglected them could bias our econometric results. Taking a first-order log-linear approximation to the production function (19) and first differencing yields the following expression for the growth rate of output:

$$\Delta y_t = \Delta a_t + \alpha_1 \frac{L}{L - \varphi} \Delta l_t + \alpha_2 \Delta k_t. \tag{21}$$

As before, lowercase letters denote the logarithms of the corresponding variables. Also, L represents the point around which we linearize (19). The key point is that, as long as $\varphi > 0$, the sum of the coefficients of Δl_t and Δk_t will not equal one even if $\alpha_1 + \alpha_2 = 1$. This is because the coefficient on Δl_t is biased upwards, away from α_1. However, this does not imply that our estimate of local returns to scale, as defined by Rotemberg and Woodford (1995), is biased. These authors define local returns to scale for a production function $F(K, L)$ to be

$$\nu = \frac{KF_1(K, L) + LF_2(K, L)}{F(K, L)} .$$

For the function $F(\cdot\ )$ given by (19), ν is given $\alpha_1\, L/(L - \varphi) + \alpha_2$. So in this case, our estimate of the sum of the coefficients on Δk_t and Δl_t is a consistent estimate of ν.

On *a priori* grounds, it might be reasonable to assume that overhead costs are more important for supervisory labor than for production workers. To the extent that this is true, estimates of the coefficient on Δl_t should be higher when that variable is measured as total hours worked rather than total production worker hours. Because of data constraints, we can only pursue this idea for the aggregate manufacturing sector. When we reestimate (13) with this measure of Δl_t and electricity-based measure of capital services, the point estimates of the coefficients of Δl_t and Δe_t are 0.82 and 0.36. The corresponding standard errors are 0.26 and 0.20, respectively. Recall from Table 4 that the analogous estimates obtained using total production worker hours as the measure of Δl_t are 0.69 and 0.31. The corresponding standard errors are 0.16 and 0.17, respectively. The fact that the point estimate of the coefficient of Δl_t is higher in the case of total worker hours is consistent with the presence of more overhead costs for supervisory workers. However, we cannot reject the hypothesis that the coefficients are actually the same in the two cases. So it is possible that there are important overhead costs associated with labor, and that these might contribute to the procyclicality of labor productivity. But the empirical case that these types of costs are more important for supervisory workers than for production workers is weak.

5.2 ISSUES REGARDING THE STOCK OF CAPITAL

Suppose that there is overhead capital which enters the production in a manner similar to overhead labor. Then, proceeding as above, it is straightforward to show that the coefficient on the change in Δk_t will be biased upwards, away from α_2. As above, this will not induce a bias in our estimate of local returns to scale, ν.

Next we consider the case in which only a subset of the capital stock employs electricity. Specifically, consider our simplest specification of the production technology (3). Suppose that $K_t = K_{1t} + K_{2t}$, so that

$$Y_t = A_t(L_t)^{\alpha_1}(K_{1t}H_t + K_{2t}H_t)^{\alpha_2}, \tag{22}$$

where H_t denotes time t hours of work per worker. Also suppose that electricity use is given by

$$E_t = \phi K_{1t}H_t$$

and that K_{2t} does not require the use of electricity. Taking a log-linear approximation to (22), we obtain

$$\Delta y_t = \Delta a_t + \alpha_1\,\Delta l_t + \alpha_2 \frac{E}{E + \phi K_2 H}\Delta e_t + \alpha_2 \frac{\phi K_2 H}{E + \phi K_2 H}(\Delta k_{2t} + \Delta h_t). \tag{23}$$

In general, the bias depends critically on the correlation between the right-hand-side regressors. As a useful benchmark suppose that $\Delta k_{2t} = \Delta k_{1t} = 0$, so that $\Delta e_t = \Delta h_t$. This is the case in which all variation in capital services corresponds to changes in the workweek of capital. Then (23) can be written as

$$\Delta y_t = \Delta a_t + \alpha_1\,\Delta l_t + \alpha_2\,\Delta e_t,$$

so that there is no bias whatsoever. A similar argument would hold had we written the production function as $Y_t = A_t(L_t)^{\alpha_1}(K_{1t}H_t)^{\alpha_2}(K_{2t}H_t)^{\alpha_3}$. Again the sum of the coefficients of Δl_t and $\Delta(h_t k_{1t}) = \Delta e_t$ would be a biased estimate of total returns to scale. But, to the extent that K_{1t} and K_{2t} do not vary over the cycle, the bias induced by working with the simple Leontief production structure will be small.

5.3 MULTIPLE SHIFTS

In our empirical work we ignored the fact that the capital stock can be utilized more intensely if plants use discrete multiple shifts. The available shift data are scarce, but suggest an interesting puzzle. In U.S. manufacturing the shift premium paid to workers is small. Kostiuk (1990) estimates a premium for the second and third shift of only 5.3%. Despite the small shift premium, most industries whose production process does not require continuous operation make modest use of the second shift and little use of the third. Bils (1992) argues that industries bunch their production in a small number of shifts because of increasing returns to scale. Shapiro (1993a) argues that the marginal premium is much higher than the commonly reported average shift premium. He estimates the marginal premium to be 25%.

To discuss how the presence of multiple shifts could affect our estimates, we extend our basic model to allow for two production shifts. Suppose that output is given by

$$Y_t = A_t H N_{1t}^{\alpha_1} K_t^{\alpha_2} + A_t H N_{2t}^{\alpha_1} K_t^{\alpha_2}. \tag{24}$$

Here N_{it} denotes the number of workers employed in shift i, for $i = 1,2$. To simplify the analysis, we assume that the shift length H is the same for both shifts. Taking a log-linear approximation to (24), we obtain the following expression for the growth rate of Y_t:

$$\Delta y_t = \Delta a_t + \alpha_1 \left[\frac{N_1^{\alpha_1}}{N_1^{\alpha_1} + N_2^{\alpha_2}} \Delta n_{1t} + \frac{N_2^{\alpha_1}}{N_1^{\alpha_1} + N_2^{\alpha_2}} \Delta n_{2t} \right] + \alpha_2 \Delta k_t. \tag{25}$$

The specification used in our empirical work can be written as

$$\Delta y_t = \Delta a_t + \alpha_1 \left[\frac{N_1}{N_1 + N_2} \Delta n_{1t} + \alpha_1 \frac{N_2}{N_1 + N_2} \Delta n_{2t} \right] + \alpha_2 \Delta k_t. \tag{26}$$

Here N_1 and N_2 are the points about which we linearize the production function. To assess the specification error associated with neglecting multiple shifts, we can compare the coefficients used to aggregate Δn_{1t} and Δn_{2t} in (25) and (26). Shapiro (1993a) reports that, for his sample of noncontinuous processor industries, the percentages of workers on the first, second, and third shifts are 68.2%, 20.7%, and 11.1%, respectively. Suppose we aggregate the second and third shifts and assume that α_1 is equal to 0.65. Then the implied coefficients of Δn_{1t} and Δn_{2t} in (25) are 0.681 and 0.319. The corresponding coefficients in (26) are 0.621 and 0.379. We conclude that while there is some bias, it is not of first order magnitude. The basic fact driving this result is the puzzle pointed out by Shapiro (1993a): Why isn't there more shift work?

6. *Conclusion*

This paper has presented evidence that capital utilization rates are sharply procyclical. Our evidence relies on an electricity-based measure of capital services. Standard measures of capital services seriously understate the procyclicality of actual capital services and lead to misleading inference regarding cyclical movements in labor productivity and the degree of returns to scale in the economy. Our results have three important implications for macroeconomists. First, models that depend on large, increasing returns to scale as a source of large propagation effects are inconsistent with the data. Granted, given the sampling uncertainty associated with our parameter estimates, it is possible to maintain that there are small increasing returns to scale. But overall, there is virtually no evidence to suggest that there are important deviations from constant returns to scale in the manufacturing industry. Second, existing RBC models which depend on large, volatile aggregate technology shocks and which predict that the growth rate of output is highly correlated with aggregate technology shocks are empirically implausible. Third, our results strongly support models which emphasize cyclical movements in capital utilization rates as an important determinant of movements in conventional measures of total factor and labor productivity. It seems very difficult to rationalize the properties of electricity use by manufacturing industries in a way that does not involve substantial cyclical movements in capital utilization.

Appendix

In this appendix, we summarize the two- and three-digit SIC codes of the manufacturing industries considered in the paper. In addition we summarize the sensitivity of the results we obtained with the two-digit SIC industries, disregarding industries in which a particularly large percentage of the Board's output measure is based on input data.

In our analysis we used the two-digit SIC industries shown in Table 8, and the three-digit SIC industries shown in Table 9.

Table 8 TWO-DIGIT SIC INDUSTRIES

SIC Code	*Name*
20	Food
21	Textiles
23	Apparel
24	Wood products
25	Furniture
26	Paper
27	Printing–publishing
28	Chemicals
31	Leather
32	Stone, clay, and glass
33	Primary metals
34	Fabricated metals
35	Machinery
36	Electrical machinery
37	Transportation equipment
38	Instruments
39	Miscellaneous

Table 9 THREE-DIGIT SIC INDUSTRIES

SIC Code	*Name*	*Output Units*
201	Meat products	Pounds
202	Dairy products	Pounds or gallons
207	Fats and oils	Pounds
212	Cigars	Units
221,2	Cotton and synthetic fabrics	Bales or linear yards
226	Fabric finishing	Linear yards
227	Carpeting	Square yards
228	Yarns and thread	Pounds
245	Manufactured homes	Units
262	Paper	Tons
263	Paper board	Tons
265	Paperboard containers	Feet
271	Newspapers	Tons
281	Basic chemicals	Tons or cubic feet
282	Synthetic materials	Pounds or tons
285	Paint	Gallons
301	Tires	Units
314	Shoes	Pairs
324	Cement	Barrels
325	Structural clay products	Units
331	Basic steel and mill products	Tons
332	Iron and steel foundries	Tons
333	Primary nonferrous metals	Tons
334	Secondary nonferrous metals	Tons
336	Nonferrous foundries	Pounds
374	Railroad equipment	Units

To assess the robustness of our results we redid the analysis underlying Table 4 excluding two subsets of industries. Excluding SIC industries 23, 25, 34, 35, 38, and 39 leaves us with industries in which at least 30% of the Board's measure of output is based on physical output. If in addition we exclude SIC industries 27, 28, 32, and 36, we are left with a panel of industries in which at least 40% of the Board's measure of output is based on physical output. All of the results in the following table refer to restricted panel estimates based on quarterly data. In Table 10, the row labeled J_2 reports the probability value associated with the statistic for testing the hypothesis of constant returns to scale, $\alpha_1 + \alpha_2 = 1$.

TABLE 10 TWO-DIGIT SIC DATA: $\Delta y_t = \gamma_0 + \alpha_1 \Delta l_t + \alpha_2 \Delta e_t + \epsilon_t$

	All 2-Digit SIC industries			*Exclude SIC 23, 25, 34, 35, 38, 39*			*Also exclude SIC 27, 28, 32, 36*		
	Manuf.	*Dur.*	*Nondur.*	*Manuf.*	*Dur.*	*Nondur.*	*Manuf.*	*Dur.*	*Nondur.*
α_1	.64	.61	.74	.75	.84	.67	.82	.98	.61
	(.05)	(.06)	(.09)	(.06)	(.10)	(.10)	(.09)	(.14)	(.12)
α_2	.37	.43	.39	.32	.24	.47	.27	.13	.51
	(.05)	(.07)	(.08)	(.06)	(.11)	(.08)	(.10)	(.16)	(.12)
J_1	.31	.12	.30	.20	.07	.42	.46	.16	.51
J_2	.75	.38	.10	.18	.15	.09	.09	.10	.22
$\sigma_\epsilon/\sigma_{\Delta Y}$	.59	.57	.63	.60	.52	.69	.59	.50	.63
$\rho_{\epsilon\Delta Y}$	.66	.59	.60	.56	.42	.55	.53	.45	.55

REFERENCES

Ballard, C., D. Fullerton, J. Shoven, and J. Whalley. (1985). *General equilibrium models for tax policy evaluation.* Chicago: University of Chicago Press.

Bartlesman, E. J. (1993). Of empty boxes: Returns to scale revisited. Board of Governors. Manuscript.

———, R. J. Caballero, and R. K. Lyons. (1994). Customer- and supplier-driven externalities. *American Economic Review* 84:1075–1084.

Basu, S. (1993). Procyclical productivity: Overhead inputs or cyclical utilization? University of Michigan. Manuscript.

———, and J. Fernald. (1994a). Are apparent productive spillovers a figment of specification error? Federal Reserve Board. International Finance Discussion Paper 463.

———, and J. Fernald. (1994b). Constant returns to scale and small markups in U.S. manufacturing. Federal Reserve Board. International Finance Discussion Paper 483.

———, and M. Kimball, (1994). Cyclical productivity with unobserved variation. University of Michigan. Manuscript.

Bernanke, B., and M. Parkinson. (1991). Procyclical productivity and competing theories of the business cycle: Some evidence from interwar U.S. manufacturing industries. *Journal of Political Economy* 99(June):439–459.

Berndt, E., and D. Wood. (1979). Engineering and econometric interpretations of energy–capital complementarity. *American Economic Review* 69(3, June):342–354.

Bils, M. (1992). Measuring returns to scale from shift practices in manufacturing. University of Rochester. Manuscript.

———, and J.-O. Cho. (1994). Cyclical factor utilization, *Journal of Monetary Economics* 33:319–354.

Board of Governors of the Federal Reserve System. (1986). *Industrial Production.*

Bruno, M. (1978). Duality, Intermediate Inputs and Value-Added. In *Production economics: A dual approach to theory and applications (Vol. 2),* M. Fuss and D. McFadden, (eds.). Amsterdam: North-Holland, pp. 3–16.

Burnside, C. (1994). What do production regressions tell us about increasing returns to scale and externalities? University of Pittsburgh. Manuscript.

———, and M. Eichenbaum. (1994). Factor hoarding and the propagation of business cycles. NBER Working Paper 4675.

———, ———, and S. Rebelo. (1993). Labor hoarding and the business cycle. *Journal of Political Economy* 101(April):245–273.

Caballero, R. J., and R. K. Lyons. (1992). External effects in U.S. procyclical productivity. *Journal of Monetary Economics* 29:209–226.

Chari, V. V., L. Christiano, and M. Eichenbaum. (1995). Inside money, outside money, and short-term interest rates. *Journal of Money, Credit and Banking,* 27.

Christiano, L. J. (1988). Why does inventory investment fluctuate so much? *Journal of Monetary Economics* 21:247–280.

———, and M. Eichenbaum. (1992). Current real business cycle theories and aggregate labor market fluctuations. *American Economic Review* 82:430–450.

———, ———, and C. L. Evans. (1994). The effects of monetary policy shocks: Some evidence from the flow of funds. NBER Working Paper 4699. *Review of Economics and Statistics,* forthcoming.

Costello, D. (1993). A cross-country, cross-industry comparison of productivity growth. *Journal of Political Economy* 101(April):207–222.
Eichenbaum, M., L. P. Hansen, and K. J. Singleton. (1988). A time series analysis of representative agent models of consumption and leisure under uncertainty. *Quarterly Journal of Economics* CIII:51–78.
Fabricant, S. (1942). *Employment in manufacturing, 1899–1939: An analysis of its relation to the volume of production*. New York: NBER.
Farmer, R., and J.-T. Guo. (1994). The econometrics of indeterminacy: An applied study. In *Carnegie-Rochester Conference on Public Policy*, forthcoming.
Finn, M. (1991). Energy, price shocks, capacity utilization and business cycle fluctuations. Institute for Empirical Macroeconomics. Discussion Paper 50. Minneapolis: Federal Reserve Bank of Minneapolis.
Foss, M. (1981). Long run changes in the workweek of fixed capital. *American Economic Review* 71:58–63.
Greenwood, J., Z. Hercowitz, and G. W. Huffman. (1988). Investment, capacity utilization and the real business cycle. *American Economic Review* 78:402–417.
Griliches, Z., and D. Jorgenson. (1967). The explanation of productivity change. *Review of Economic Studies* 34:249–283.
Hall, G. (1994). Overtime, effort and the propagation of business cycle shocks. Federal Reserve Bank of Chicago. Working Paper No. 94-25.
Hall, R. E. (1988). The relation between price and marginal cost in U.S. industry. *Journal of Political Economy* 96:921–948.
———. (1990). Invariance properties and Solow's productivity residual. In *Growth/Productivity/Employment*, P. A. Diamond (ed.). Cambridge, MA: MIT Press, pp. 71–112.
Hansen, L. P. (1982). Large sample properties of generalized method of moments estimators. *Econometrica* 50:1029–1054.
Jorgenson, D., F. Gollop, and B. Fraumeni. (1987). *Productivity and US Economic Growth*. Cambridge, MA: Harvard University Press.
Kydland, F., and E. C. Prescott. (1982). Time to build and aggregate fluctuations *Econometrica* 50:1345–1370.
———, and ———. (1988). The workweek of capital and its cyclical implications. *Journal of Monetary Economics* 21:343–360.
Kostiuk, P. (1990). Compensating differentials for shift work. *Journal of Political Economy*, 98:1055–1075.
Long, J., and C. Plosser. (1983). Real business cycles. *Journal of Political Economy* 91:39–69.
Lucas, R. E., Jr. (1970). Capacity, overtime, and empirical production functions. *American Economic Review: Papers and Proceedings* 60:23–27.
Mankiw, N. G., and D. Romer. (1991). *New Keynesian Economics*. Cambridge, MA: MIT Press.
Miron, J. A., and S. Zeldes. (1989). Production, sales and the change in inventories: An identity that doesn't add up. *Journal of Monetary Economics* 24:31–52.
Perry, G. L. (1973). Capacity in manufacturing. *Brookings Papers on Economic Activity* 3:701–742.
Ramey, V. (1989). Inventories and factors of production and economic fluctuations. *American Economic Review* 79:338–354.
Rotemberg, J., and L. Summers. (1990). Inflexible prices and procyclical productivity. *Quarterly Journal of Economics* 105:851–874.

———, and M. Woodford. (1995). Dynamic general equilibrium models with imperfectly competitive product markets. In *Frontiers in Business Cycle Research,* Thomas F. Cooley (ed.). Princeton, NJ: Princeton University Press.
Sato, K. (1967). A two level CES production function. *Review of Economic Studies* 34:201–218.
Shapiro, M. (1993a). Capital utilization and the premium for working at night. University of Michigan. Manuscript.
———. (1993b). Cyclical productivity and the workweek of capital. *American Economic Review: Papers and Proceedings* 83:229–233.
Shea, J. (1993a). 'The input selection approach to instrument selection. *Journal of Business and Economic Statistics* 11:145–155.
———. (1993b). Do supply curves slope up? *Quarterly Journal of Economics* 108:1–32.
Solow, R. M. (1957). Technical change and the aggregate production function. *Review of Economics and Statistics* 39:312–320.
———. (1964). Draft of presidential address: On the short-run relation of employment and output. MIT. Manuscript.

Comment

SUSANTO BASU
University of Michigan and NBER

1. *Introduction*

This paper sets itself an ambitious task: it attempts to explain a fundamental stylized fact of modern macroeconomics—the procyclicality of the Solow residual—using an empirical model that nests the three main explanations for this fact: technology shocks, increasing returns to scale, and unobserved input variation. This separation is not really that straightforward, since the first is an impulse and the other two are propagation mechanisms. Even if we conclude at the end of the day that these propagation mechanisms are important, it might still be the case that business cycles are fundamentally technology-driven. But since business-cycle models based on increasing returns or variable factor utilization often invoke other sources of shocks—e.g., government spending or animal spirits—it seems useful to maintain the categorization.

The recent literature has given us some conflicting evidence on these three explanations. As the paper notes, Robert Hall's early work suggested that markups and returns to scale are very large, but more recent papers correct for Hall's use of value-added data and small-sample problems in his econometric procedures. These papers consistently find approximately constant returns and small markups. A number of papers also investigate the role of variable factor utilization in dynamic optimizing models, and generally conclude that variable utilization can explain a substantial fraction of the cyclicality of productivity. So I come to this paper mostly convinced of its two main messages: returns to scale are about constant, and much of the measured cyclicality of productivity is due to variable factor utilization.

Even viewed from this background, the paper makes a number of useful and interesting contributions. Its substantive message is that using electricity consumption as a proxy for capital utilization leads to more sensible results in short-run time-series analysis, and controlling for capacity utilization drastically changes the estimated properties of technology shocks. The subtext of the paper is a methodological message which I strongly endorse: that thinking carefully about the properties of data and measurement, always important, is vital in this line of work.

I shall take issue with BER's argument that their results necessarily indicate that returns to scale are about constant. But since I share their conclusion based on other evidence, this debate is really one about method rather than substance.

Thus, I conclude by asking a substantive question: how should one interpret the paper's conclusion on returns to scale? Does accepting the premise that micro-level returns to scale are approximately constant sound the death knell for increasing-returns models in macroeconomics?

2. *The SRIRL Puzzle*

BER argue persuasively that electricity use helps us resolve the SRIRL puzzle. I prefer to put the "increasing returns" part of that aside for the moment. But a major embarrassment of a long empirical literature is the result that capital services don't matter for short-run production when the capital stock is used as a proxy for capital services. As BER point out, their idea of using electricity as an alternative proxy has old roots. Nevertheless, they provide a production framework that one can use for regression analysis rather than productivity studies. In view of their results, electricity use should become a common proxy for capital utilization. It will be interesting to see how this new method affects results in papers that use more questionable proxies: for example, the Federal Reserve's series on capacity utilization.

However, I would have liked to see a comparison of the electricity-based capital utilization series with others that are implied from estimates of optimizing models of firm behavior (e.g., Burnside and Eichenbaum, 1994). Finding the two series in agreement would greatly strengthen my faith in both.

3. *Increasing Returns to Scale*

Thus, I agree with BER that changes in the capital stock are likely to be bad short-run measures of changes in capital services, and that the change in electricity consumption is a good proxy for the change in true capital services. Does BER's procedure then imply any necessary conclusion about the degree of returns to scale?

The answer is no. For simplicity, I illustrate my point using a small modification of specification (1) in the paper, but the point applies with equal force to all cases in which electricity consumption is used as a proxy for capital services. Suppose we generalize the paper slightly:

$$V_t = A_t(L_t)^{\alpha_1}(K_t^*)^{\alpha_2}, \tag{1}$$

$$K_t^* = \min(\tilde{K}_t^{\beta}, E_t). \tag{2}$$

K_t^* should be interpreted as the input of *variable* capital. The $\tilde{K}_t$ in (2) signifies that this is the true input of capital services, which is not well measured by the capital stock. BER's results on returns to scale turn on the assumption that β is 1: that is, they assume that the K_t^* production function is homothetic, and thus the "output expansion path" between $\Delta\tilde{k}$ and Δe is the 45-degree line. On the other hand, suppose that $\beta > 1$. Then the production function is nonhomothetic and the expansion path no longer has a slope of 1. As the example makes clear, the question of homotheticity is independent of the elasticity of substitution between capital and energy.

Nonhomotheticity has important consequences for BER's estimates of returns to scale. Note that their method still provides consistent estimates of α_1 and α_2. But $\alpha_1 + \alpha_2$ is no longer the degree of returns to scale in this production function. The degree of returns to scale (in the production of value added) is now

$$\text{RTS} = \alpha_1 + \beta\alpha_2. \tag{3}$$

Thus, if $\alpha_2 > 0$ and $\alpha_1 + \alpha_2 = 1$, as BER argue, then there are increasing returns to scale if $\beta > 1$.

Why should we believe that the production of capital services is not homothetic in its inputs? One possibility is overhead capital that does not use much electricity, so that (2) becomes

$$K_t^* = \min(\tilde{K}_t - \bar{K}, E_t). \tag{2'}$$

What is a good candidate for the overhead capital $\bar{K}$? One possibility is structures. Suppose that structures use a negligible amount of energy. Suppose also that they are in large measure a fixed cost of operation. Then one would have exactly the situation I outlined, with

$$\beta = \frac{\tilde{K}}{\tilde{K} - \bar{K}} > 1.$$

In general, any feature of the production technology that makes marginal electricity use higher than average electricity use will lead BER to underestimate the degree of returns to scale. In their discussion of this issue in the paper, BER assume that non-electricity-using capital is used in proportion with electricity-using capital. But this assumption does not hold if the non-electricity-using capital is a fixed cost of operation. For example, a factory building typically must be rented for 24 hours a day, whether the factory works one shift or three.

I have a similar concern regarding overhead labor. As the paper notes, if there is an analogous overhead-labor requirement the production function becomes

$$V_t = A_t(L_t - \bar{L})^{\alpha_1} (K_t^*)^{\alpha_2}, \tag{1'}$$

where K_t^* is still given by (2′). In all but one section, the paper uses production-worker hours as the measure of labor input. Production-worker hours are likely to be a good measure of variable labor input, $L_t - \bar{L}$. But the degree of returns to scale is a function of total labor input, and is given by

$$\text{RTS} = \alpha_1 \frac{L}{L - \bar{L}} + \alpha_2 \frac{\tilde{K}}{\tilde{K} - \bar{K}}. \tag{3'}$$

BER do estimate a regression for aggregate manufacturing that uses total hours instead of production-worker hours. The point estimates in fact suggest there is substantial overhead labor (about 15% of total labor), but BER do not find a significant difference between the two estimates. However, to gain precision they should repeat the test with industry rather than aggregate data. Data on total hours by industry are certainly available at an annual frequency. Using the panel should more than compensate for the shift to a shorter time series.

Thus, I would interpret BER's test for constant returns quite differently. It is actually a test to see whether there are constant returns to the *variable* inputs, i.e., to see whether the marginal-cost curve is flat. But with flat marginal cost, which is the result they cannot reject ($\alpha_1 + \alpha_2 = 1$), any fixed costs would yield globally *increasing* returns.[1] Thus I suggest that BER concentrate more on the point estimates than on the hypothesis tests. Their median point estimates for the unrestricted three-digit data, which I find most compelling, suggest *diminishing* returns to the variable inputs. This raises the possibility that returns to scale are constant if firms operate at the minimum point of their U-shaped average cost curves. But of course, fixed costs and increasing marginal cost do not guarantee constant returns: returns to scale then depend on the average ratio of overhead to total inputs. To settle this issue, BER would have to find some independent method of estimating this ratio.

Basu (1993) addresses this issue and proposes one possible solution. It relies on two principles: that there are increasing marginal costs to changing utilization (otherwise utilization would not vary in any smooth fashion), and that as a consequence firms would prefer to adjust to anticipated, long-run changes in demand along the extensive margin (changing the capital stock) rather than along the intensive margin (changing utilization). From observing the change in the ratio of observed capital use to electricity use in response to a change in demand that was anticipated, and anticipated to be long-lasting, we should be able to estimate the β in the production of capital services, K_t^*. There are substantial problems with this method, the largest being the difficulty in identifying anticipated, long-lasting demand shocks, but it is one approach. I had hoped that these three authors would have suggested other and better methods.

Another source of evidence comes from the realization that we are no longer estimating technology but rather market behavior. With a U-shaped average cost curve the same technology is consistent with a wide range of returns to scale, ranging from constant returns to large increasing returns. Which of these we observe depends on the size of the markup firms charge above marginal cost. If we know the markup μ, we can place an upper bound on the degree of returns to scale using the identity that $\mu(1 - \pi) = \text{RTS}$, where π is the rate of pure profit. Since the profit rate is widely estimated to be small, this upper bound is likely to be a tight one. Thus, we can bring to bear a variety of evidence on industry competitiveness from the industrial organization literature, which will tell us something about the size of the average markup (and

1. Some of the recent literature on sunspot-based models of business cycles (e.g. Farmer and Guo, 1994) has used estimates of returns to scale to calibrate models that really depend on decreasing marginal cost. The two concepts are equivalent only if there are no fixed costs.

thus about the average degree of returns to scale). Knowing these, and with evidence on the slope of the marginal cost curve, we can then estimate the degree of cyclicality of the markup.

As a first step, however, I do a back-of-the-envelope calculation by taking structures and nonproduction labor as proxies for overhead capital and labor. The average ratio of structures to total capital in manufacturing has been about 0.4. Ramey (1991) suggests that the average ratio of nonproduction workers to total employment is about 0.2. Using these figures and the median estimates for the unrestricted three-digit industries in BER's Table 5 to plug into equation (3′), we get average *value-added* RTS of 1.05 for nondurables, 1.20 for durables, and 1.28 for total manufacturing. The first figure agrees closely with results found in the recent literature, but the other two are somewhat higher. Thus, I would say that BER provide mixed evidence in favor of approximately constant returns.

How does this issue affect BER's other conclusions? Not a great deal. It does not affect the validity of using electricity consumption as a proxy for capital services, so long as one is not interested in interpreting the coefficient on electricity use as a measure of the elasticity of output with respect to capital input. Nor does it affect their computation of the statistical properties of technology shocks. If my argument about overhead factors is right, their estimates of the output elasticities of capital and labor are biased down. On the other hand, they then multiply these downward-biased coefficients by input changes that are too large [Δe and $\Delta \ln(L_t - \bar{L})$ rather than Δk^* and Δl]. The two errors just cancel out in expectation, leaving the estimated technology series unaffected.

Thus, BER's finding that technology shocks seem drastically less volatile and less procyclical after controlling for capital utilization remains an interesting and important contribution. One might be tempted to argue that this is bad news for real-business-cycle models. But that need not be true, as Burnside and Eichenbaum (1994) show. Since capital utilization is an additional propagation mechanism, a real-business-cycle model with less-volatile technology shocks might nevertheless account as well for the variance of output, while better matching some of the time-series properties of economic fluctuations.

4. *Aggregation*

Suppose, however, that we accept BER's conclusion: returns to scale are constant at the micro level. Does this imply that macroeconomists should abandon the lessons of increasing-returns models? The answer is, not necessarily.

Let me propose an example loosely grounded in some of the findings of the recent industrial-organization literature. Although three-digit data are quite "micro" enough for macroeconomics, applied production analysis is now frequently done with establishment-level data from the Longitudinal Research Database of the Census of Manufactures. These papers (e.g. Baily, Hulten, and Campbell, 1992) generally find that plant-level returns to scale are about constant. However, they also find that there are substantial differences in productivity *levels* across plants (Caves and Barton, 1990; Baily, Hulten, and Campbell, 1992), even within very narrowly defined industries.

The example is based on these two ideas: constant returns to scale, and productivity-level differences. Suppose the production side of an economy consists of two firms, each producing the same good with a constant-returns-to-scale, Cobb–Douglas production function:

$$Y_1 = A_1 K_1^{\alpha} L_1^{1-\alpha}, \tag{4a}$$
$$Y_2 = A_2 K_2^{\alpha} L_2^{1-\alpha}. \tag{4b}$$

Suppose $A_1 > A_2$. The total inputs of the economy are K and L. Suppose firm 1 gets a share s of the total inputs, so $K_1 = sK$ and $L_1 = sL$. The remainder of the inputs go to firm 2. The economywide output Y is the sum of Y_1 and Y_2. Now suppose we allow the share of inputs going to plant 1 to depend on aggregate output: $s = s(Y)$. In this example, there is formally no aggregate production function. But suppose we hypothesize that one exists, $Y = \Omega F(K, L)$. We now attempt to discover the degree of returns to scale of this hypothetical production function using Hall's (1990) method, which is to regress output growth Δy on cost-weighted input growth Δx. We find

$$\begin{aligned}\Delta y &= \frac{1}{1 - s\varepsilon_s \dfrac{A_1 - A_2}{sA_1 + (1-s)A_2}} [\alpha \Delta k + (1-\alpha)\Delta l] + [s\Delta a_1 + (1-s)\Delta a_2] \qquad (5)\\ &= \gamma \Delta x + \Delta \omega,\end{aligned}$$

where γ is the degree of returns to scale of the "aggregate production function," and ε_s is the elasticity of s with respect to Y. Note that returns to scale exceed 1 if ε_s is positive, i.e., if inputs flow towards high-productivity uses in booms.

What is the "true" degree of returns to scale in this economy? The answer is that it depends on the application. Suppose one wants to estimate the average markup of price over marginal cost in order to calculate the welfare cost of monopoly pricing. Then one would have to use firm-level data, conclude that firms produce with constant returns, and compute the markup from data on profit rates. On the other hand, suppose one is interested in knowing the cyclicality of the average marginal product of labor. Then the appropriate degree of "returns to scale" is γ, since it reflects changes in the marginal product of labor coming from superior allocation of inputs in a boom.

Basu and Fernald (1995) find that a substantial fraction of the cyclicality of aggregate manufacturing productivity can be explained by composition effects of the sort proposed here, that is, one where the "macro returns to scale" exceed the "micro returns to scale." BER's results offer one shred of evidence consistent with a model like the one I propose. Their point estimates for returns to scale in two-digit manufacturing are consistently larger than the corresponding point estimates for the three-digit data. This example would predict such a finding, as composition effects increase the effective degree of returns to scale at higher levels of aggregation.[2]

One would have to do much more to turn this example into a full model. For example, it is clear that in order to have both firms produce in equilibrium at two different productivity levels, we need imperfect competition in either the product or the factor markets. A high price–cost markup by firm 1 would allow both firms to sell at the same price, but large markups combined with constant returns to scale imply large pure profits, which are not observed in the U.S. economy. A more promising alternative may be imperfect competition in the factor market, whereby labor and capital extract most of the rents from high productivity. Baily, Hulten, and Campbell (1992) find some evidence consistent with this hypothesis, which is also advanced by Katz and Summers (1989).

Macroeconomists know that micro-level heterogeneity is a fact of life. This fact is dutifully repeated, and then disregarded in most macroeconomic research. As that research concentrates more on estimating structural parameters from increasingly fine micro data, the danger from disregarding aggregation becomes larger. Parameters estimated from data that are truly at the micro level may bear little resemblance to the parameters of interest to macroeconomists, who try to characterize the behavior of aggregates. I have some reservations about BER's method, but I substantially agree with their conclusion, that the "average" micro-level returns to scale is about constant. But this conclusion is the alpha

2. Of course, this finding is also consistent with the existence of technological spillovers in production, and was interpreted as evidence for such spillovers by Caballero and Lyons (1992).

and not the omega of asking whether increasing-returns models are useful for macroeconomics.

REFERENCES

Baily, M. N., C. Hulten, and D. Campbell. (1992). Productivity dynamics in manufacturing plants. *Brookings Papers on Economic Activity (Microeconomics)* 1992(1):187–267.
Basu, S. (1993). Cyclical productivity: Increasing returns or cyclical utilization? University of Michigan. Mimeo.
———, and J. G. Fernald. (1995). Aggregate productivity and the productivity of aggregates. Hoover Institution. Mimeo.
Burnside, C., and M. Eichenbaum. (1994). Factor hoarding and the propagation of business cycles. NBER. Working Paper 4675.
Caballero, R. J., and R. K. Lyons. (1992). External effects in U.S. procyclical productivity. *Journal of Monetary Economics* 29:209–226.
Caves, R. E., and D. R. Barton. (1990). *Efficiency in U.S. Manufacturing Industries.* Cambridge, MA: MIT Press.
Farmer, R., and J.-T. Guo. (1994). Real business cycles and the animal spirits hypothesis. *Journal of Economic Theory* 63:42–72.
Hall, R. E. (1990). Invariance properties and Solow's productivity residual. In *Growth/Productivity/Employment*, P. A. Diamond (ed.). Cambridge, MA: MIT Press, pp. 71–112.
Katz, L. F., and L. H. Summers. (1989). Industry rents: Evidence and implications. *Brookings Papers on Economic Activity (Microeconomics)* 1989(1):209–290.
Ramey, V. A. (1991). Comment. In *NBER Macroeconomics Annual*, O. J. Blanchard and S. Fischer (eds.).

Comment

ROBERT E. HALL
Stanford University and NBER

The paper by Burnside, Eichenbaum, and Rebelo (BER) differs from earlier contributions to research on cyclical fluctuations in productivity in five ways:

1. BER drop the assumption of cost minimization and the use of factor prices in favor of direct estimation of production-function elasticities.
2. They use data on electricity consumption as a proxy for the flow of capital services.
3. They assume that materials inputs and output are perfect complements, so that data on gross output can be used without measuring materials inputs.
4. They measure the changes in all labor input from the changes in labor input from production workers.

5. They use the lagged unexplained element of monetary policy as an instrument to estimate the production function.

I will start by describing the basic BER method. It starts with the standard log-linear approximation to the production function:

$$\Delta Y = \sum_{F \in \{K,L,E,M\}} \tilde{\alpha}_F \Delta F + \varepsilon.$$

Here ΔY is the change in the log of gross output, $\tilde{\alpha}_F$ is the elasticity of the production function with respect to factor F, ΔF is the change in the log of the amount of that factor in use, and ε is the random growth in productivity.

The first step is to apply the hypothesis that Y, E, and M move in proportion:

$$\Delta Y = \alpha_L \quad \Delta L + \alpha_K \quad \Delta K + \varepsilon,$$
$$\alpha_L \quad = \frac{\tilde{\alpha}_L}{1 - \tilde{\alpha}_E - \tilde{\alpha}_M},$$

and similarly for α_K. Data on gross output can be used in conjunction with data on labor and capital only, given this assumption. Increasing returns can be diagnosed in the usual way after taking this step:

$$\sum_{F \in \{K,L,E,M\}} \tilde{\alpha}_F > 1 \quad \text{iff} \quad \alpha_L + \alpha_K > 1.$$

πBER, in their specification 3, also consider the earlier approach exploiting the proposition that cost minimization implies that the elasticities are in proportion to the cost shares:

$$\alpha_F = \gamma \frac{p_F F}{\Sigma\, p_X X} = \gamma s_F.$$

Thus, the extent of increasing returns can be measured from the one-parameter equation

$$\Delta Y = \gamma(s_L \, \Delta L + s_K \, \Delta K) + \varepsilon.$$

In BER's basic specification, they estimate the elasticities directly. On the one hand, they avoid the assumption of cost minimization and the need to measure factor prices. On the other hand, they need more powerful instruments to estimate more parameters.

The second step is to apply the hypothesis that E and K move in proportion:

$$\Delta Y = \alpha_L \, \Delta L + \alpha_K \, \Delta E + \varepsilon.$$

BER find reasonable values for the two elasticities, and the sum of the two is very close to one. Of the innovations I listed, the one responsible for overturning earlier findings of increasing returns to scale is the use of the electricity proxy.

I find the logic of the electricity proxy much less compelling than do BER. The single most important category of capital is computers. Larger computers are used 24 hours a day—a small staff works overnight running batches of transactions. Computers are not very electricity- or labor-intensive; they are just expensive. BER assume that computer utilization fluctuates along with electrically intensive production operations. The result is an unambiguous downward bias in the sum of the production elasticities. But BER, following Matthew Shapiro's earlier work, confirm the findings by using a direct measure of the workweek of capital. There is a question whether the direct measure may not be superior to the electricity proxy.

My strongest disagreement with BER is about the assumption of perfect complementarity between output and materials inputs. Their own data on electricity usage do not support the hypothesis completely. Any cyclical fluctuations in the extent of vertical integration will invalidate BER's approach. If firms contract out when demand is strong and make their own when it is weak, the perfect complementarity will fail, potentially in an important way. I am not advocating the use of value added, but rather tackling the true production function head-on, using data on gross output and all inputs, including materials. The necessary data are available (at annual frequency) and ought to be used.

BER's assumption that nonproduction and production workers move in exact proportion is plainly refuted by their own data. The assumption clearly biases the results against finding increasing returns, for the same reason I discussed above. By overstating the movements of total labor input, the results understate the elasticity of the production function with respect to labor. Good data on hours of all workers are available at the two-digit level. A high priority should be the development of similar data at finer levels of disaggregation.

Previous research on cyclical productivity has avoided using monetary policy as an instrument, on the grounds that policy may respond to shifts in technology. BER overcome this objection through timing—they remove the contemporaneous relation between major macro variables and their measure of monetary policy. The key, unstated assumption is that monetary policymakers do not have advance knowledge of technology shifts. I do not find this unreasonable. It turns out to yield powerful instruments because the lagged effect of monetary policy on real activity is so potent.

BER deliberately omit all of the developments in research on cyclical productivity based on elaboration of the way labor enters production. The strategy of the paper is to concentrate on capital measurement issues. In brief, research by Susanto Basu and others has shown that the use of weekly hours per worker as a proxy for unmeasured fluctuations in work effort will largely eliminate evidence of increasing returns. The research proceeds in exact parallel to BER's work.

Thus, it is now well established that adding a free variable—electricity consumption, workweek of capital, or workweek of labor—to a cyclical productivity equation will eliminate evidence of increasing returns. These free variables are highly correlated with output. Research now needs to turn to the issue of whether the role of the free variables in the productivity equation is at the level that makes sense as a matter of theory, or whether its role is exaggerated by problems of measurement errors.

Central to this next step is the creation of a complete theory of factor utilization. BER do not inquire into the economics of the workweek of capital, but we already know it is a murky subject. If extra hours of use of capital come at zero cost, it is hard to explain why capital is not used every hour of the week. Shift differentials in labor cost may be part of the story, but depreciation of capital in use may be another. There are also strategic theories of the value of excess capital—they support a subgame perfect equilibrium in which entry is deterred by the credible threat to revert to competition upon entry. Competitive levels of output could not be produced without the extra capacity.

Current research, including this paper, seems in danger of finding implausibly little increasing returns. We know that firms have certain kinds of overhead, including intellectual property and organizational capital. The finding of constant returns to scale, along with its direct counterpart, zero markup of price over marginal cost, leaves no room for any kind of overhead. I suspect that when we solve some more of the measurement problems, we will conclude in favor of mild increasing returns.

Discussion

In response to the comments by Hall and Basu, Eichenbaum agreed that overhead labor and capital were important, at least in principle; Section 5 of the paper provides some discussion of the potential sensitivity of the results to the existence of overhead factors. Hall and Eichenbaum also discussed the implications of the fact that overhead capital (structures, computers) is probably less electricity-intensive than capital whose use varies with production. (The revised version of Section 5 includes an example that investigates this possibility.) Eichenbaum resisted the identification of overhead labor with supervisory workers; he noted that much supervisory labor might vary with production, and that he and his coauthors had been unable to reject statistically that the overhead component was larger for supervisory than production workers. Rebelo interjected that their intention had been to design a consistent theoretical framework which would allow them to estimate the parameters of interest, given the available data; if better or more disaggregated data had been available, it might have been feasible to use a model including a more detailed treatment of overhead labor, overhead capital, and different types of capital.

Hall took Rebelo's comment as justification for a greater emphasis on annual (as opposed to quarterly) data, since more detail is available at the annual frequency. For example, as Basu noted, there are annual, industry-level data on supervisory and production worker-hours; there are also industry data on materials inputs. Eichenbaum defended the emphasis on quarterly data on the grounds that having more observations improved the precision of the estimates. In addition, he noted that the annual capital stock data are constructed using strong maintained assumptions about market structure, user costs, etc.; he felt it was worth exploring alternatives to using these data. Hall responded that going from annual to quarterly data did not increase the real quantity of information by very much, and entailed some sacrifices. Ben Bernanke pointed out that quarterly data have the advantage of better capturing business-cycle phenomena, which are largely what the paper is about. Robert Gordon suggested that the authors should use data at both frequencies to check the bias in their estimates.

Julio Rotemberg made the point that there are no data at either frequency that identify a key parameter, the marginal elasticity of electricity use with respect to variations in capital utilization. Thus there is no real alternative to making strong modeling assumptions.

John Shea focused on the paper's results which indicate that the slope of the short-run marginal cost curve is positive. He noted that this paper agrees with two previous studies—Basu's paper, which uses materials as a proxy for variable input, and Shea's recent article in the *Quarterly Journal of Economics*—which find the elasticity of marginal cost with respect to output to be about 0.2 in the short run. Hall remarked that the long-run supply curve might still slope down. Shea agreed, but pointed out that for many macroeconomic issues, for example the propagation of demand shocks, what matters most is the slope of the short-run marginal cost curve.

Simon Gilchrist noted that because the share of computers in producers' equipment had increased dramatically in recent years, there might be strong trends in the biases of the estimated coefficients, and these trends might be correlated across industries.

Ben Bernanke asked about the implications of the paper's results for real business cycle (RBC) models, as contrasted, for example, with monetary models of the business cycle. Eichenbaum said that he believed that it is very hard to isolate large exogenous shocks coming from technology. This is not fatal for the RBC approach, but it suggests the need to look for strong propagation mechanisms that can amplify relatively small technological impulses into large economic fluctuations. On the issue of the relative explanatory power of technology shocks and money, Eichenbaum cited vector autoregression studies that attribute about 30–35% of the forecast variance of output to monetary policy innovations; this result, even if taken at face value, leaves considerable room for technological and other sources of cyclical fluctuations.

Basu raised the issue of whether omitting data on materials inputs mattered for the paper's results. He noted that even if the production function was Leontief in value added and materials, the omission of materials would matter if materials inputs did not enter with an exponent of one, that is, if the production function were nonhomothetic. He claimed that an instrumented regression of the change in output on the change in materials inputs typically yields a coefficient of about 0.8, rather than 1.0 (as would be the case if the production function were homothetic). Eichenbaum said that he and his coauthors had estimated the same regression and found a similar coefficient, in the vicinity of 0.9. Given measurement errors, he suggested, it is difficult to conclude that these point estimates are significantly different from one.

Gordon asked whether there was a trend in the share of materials. Basu said that there had been a lot of materials deepening over time, with the share of materials rising from 0.5 in the 1960s to 0.6 in the 1980s. Eichenbaum did not consider this trend to be a problem for their framework.