

Productivity Growth in the 1990s: Technology, Utilization, or Adjustment?

Susanto Basu
University of Michigan
and National Bureau of Economic Research

John G. Fernald
Federal Reserve Bank of Chicago

Matthew D. Shapiro
University of Michigan
and National Bureau of Economic Research

November 2000
Revised June 9, 2001
Corrected August 9, 2001

The authors are grateful to Robert Hall, Andrew Abel, Dale Jorgenson, Robert Solow, and seminar participants at Columbia, the Federal Reserve Bank of Chicago, the Federal Reserve Board, the International Monetary Fund, Michigan, Northwestern, and NYU for their comments and suggestions. Basu acknowledges support of the Sloan Foundation. Shapiro acknowledges support of the National Science Foundation (SRB-9617437).

ABSTRACT

Productivity Growth in the 1990s: Technology, Utilization, or Adjustment?

Measured productivity growth increased substantially during the second half of the 1990s. This paper examines whether this increase owes to an increase in the rate of technological change or whether it can be explained by non-technological factors relating to factor utilization, factor accumulation, or returns to scale. It finds that the recent increase in productivity growth does appear to arise from an increase in technological change. Cyclical utilization raised measured productivity growth relative to technology growth in the first part of the expansion, but lowered it subsequently. Factor adjustment leads to a steady-state understatement of technology growth by measured productivity growth. The understatement was greater in the second half of the expansion than the first. Changes in the distribution of inputs across industries with different returns to scale lead to a modest understatement in the growth in technology. Although the increase technological change is most pronounced in durable manufacturing, technological change also increased outside of manufacturing.

Susanto Basu
Department of Economics
University of Michigan
Ann Arbor MI 48109-1220
and NBER
tel. 734 764-5359

John G. Fernald
Federal Reserve Bank
of Chicago
Chicago IL 60690-0834
tel. 312 322-2116

Matthew D. Shapiro
Department of Economics
University of Michigan
Ann Arbor MI 48109-1220
and NBER
tel. 734 764-5419

1. Introduction

Measured productivity growth increased dramatically during the second half of the 1990s. Does this increase in productivity growth herald a new industrial revolution based on computers and information technology? Is this increase just a bit of temporary good luck? Or is it merely mismeasurement arising from the increase in effort, factor utilization, or factor accumulation that accompanies a booming economy?

The answers to these questions cannot be definitive until more time passes.¹ In particular, it is very hard to address the question of whether the current good performance of productivity is temporary. The boom in the stock market provides some ancillary evidence that might bear on this question, yet it is subject to differing interpretations.² Our approach, however, is to limit our attention to the internal evidence on output and inputs and their cyclical relationship. These relationships will allow us to extract an estimate of technology from productivity and therefore shed light on what has happened in the recent past, but these estimates will to an extent leave open the question of the future of growth in technology.

A major contribution of this paper is to analyze two potentially offsetting cyclical factors in measured productivity: factor utilization and adjustment costs. Attention to factor utilization has a long history in productivity measurement. The basic idea is that unaccounted-for changes

¹ Recent papers examining these issues include Baily and Lawrence (2001), Gordon (2000), Jorgenson and Stiroh (2000), Nordhaus (2000), Oliner and Sichel (2000), Stiroh (2001), and Whelan (2000a).

² For example, both Robert Hall (2001a) and Robert Shiller (2000) attribute the boom to information technology, but Hall presumes that the stock market is reacting to underlying fundamentals relating to information technology while Shiller believes that popular perceptions about information technology have given impetus to a speculative bubble.

in utilization and effort will raise measured productivity without having any effect on true technology. Solow (1957) made a correction for utilization of capital in his seminal paper. In the productivity literature that followed, such adjustments were routine (either explicitly or by averaging over the business cycle). Though early real business cycle literature missed the point about cyclical productivity, there has been a resurgence of attention to this issue.³

Adjustment costs similarly require that measured productivity be adjusted to yield an estimate of technology. Broadly speaking, adjustment costs reduce output to the extent that productive resources are diverted from production to adjustment when firms undertake capital accumulation or hiring. Hence, when adjustment is increasing, output growth will be temporarily damped, yielding an underestimate of technological change.

Adjustment costs have received less attention than utilization, at least in the recent literature in macroeconomics. Yet, they have a role in productivity measurement that is closely linked to that of utilization. First, if increases in factor utilization and increases in factor adjustment are positively correlated, then the utilization and adjustment have opposite effects on measured productivity. Second, costs of adjustment presumably drive cyclical variation in utilization. If quasi-fixed factors were costless to adjust (i.e., not really quasi-fixed), then there would be no need to pay for costly variation in their utilization.⁴ Hence, the recent literature that

³ Greenwood, Hercowitz, and Huffman (1988) is an early real business cycle model that does incorporate variable utilization. See Shapiro (1986b, 1993, 1996), Basu (1996), Basu and Kimball (1997), and Burnside, Eichenbaum, and Rebelo (1995) for the importance of variable utilization in cyclical productivity.

⁴ Adjustment and utilization need not move together. First, the timing may be different, as we find during the 1990s. That is, since utilization is confined to a bounded range, it may return to its long-run level at some point during an expansion, whereas factor accumulation continues; see Sims (1974). Second, adjustment and utilization could even move in opposite directions. For example, if capital depreciates in use, a high shadow cost of current capital relative to future capital can decrease utilization and increase adjustment.

emphasizes variable utilization implicitly or explicitly assumes some quasi-fixity or fixed cost. We show how measurement of technology is affected by this inherent interaction when the quasi-fixity is motivated by adjustment costs.⁵

This paper also accounts for a third effect that creates a wedge between measured productivity and true technology. Basu and Fernald (1997) show how shifts in the composition of input growth across industries with differing returns to scale can create such a wedge. Though this effect is not important cyclically in the 1990s, it does have a role in the appropriate measure of technology.

Factor adjustment can have both cyclical and steady-state effects on measured productivity. The steady-state effect comes from the drag of adjustment costs on growth from steady-state capital accumulation. This drag increases when investment rises, as it did in the 1990s. Kiley (2001) also suggests that these effects might be substantial.

Factor utilization and adjustment play a potentially important role in understanding the acceleration in productivity in the 1990s. The 1990s began with a shallow recession. Though the time between the peak in second quarter of 1990 and the trough in first quarter of 1991 was not particularly long, the speed of the recovery was unusually slow.⁶ Once growth accelerated, there was a substantial cyclical contribution of utilization to measured productivity. This cyclical bounce in measured productivity, of course, simply offset the cyclical decline experienced going into the recession. This cyclical effect is quite standard, though it is important

⁵ Time to build would be an alternative explanation of the quasi-fixity. We do not account for it here apart from a one-period lag that it takes investment to enter the capital stock. The costs of reallocation of existing capital emphasized by Ramey and Shapiro (1998) are another channel for delayed adjustment. We lack information on reallocation for recent years, so we will not investigate this channel.

to keep track of it in assessing the performance of the 1990s. We find that utilization contributed about 1/2 percentage point per year to growth in the measured Solow residual in the 1992-1994 period as the economy recovered from recession. Since then, utilization has bounced around from year to year, but on balance, has contributed *negatively* to growth in the Solow residual, and thus does not explain the increase in growth in the second half of the 1990s.

The 1990s are distinct, however, in the changes in factor accumulation, particularly that of capital.⁷ Figure 1 shows the ratio of nominal nonresidential fixed investment to GDP over the post-war period. The 1990s experienced a boom in business investment of unprecedented size and duration. The increase in investment in the 1970s was as sharp, but not as prolonged. In the two other long expansions since the World War II – the 1960s and the 1980s – investment peaked well before the end of the expansions.

Information technology equipment – computers plus telecommunications equipment – has been a major part of the story. Its share in total business fixed equipment investment increased dramatically in the 1990s. Figure 2 shows how during the 1990s, the share of information technology investment in GDP rose from 3 percent to almost 6 percent.⁸ Much of this information processing equipment has been purchased by the non-manufacturing sector.

⁶ See Economic Report of the President (1994, pp. 56-7, 98-101). The growth in output and labor were both unusually slow.

⁷ In contrast with capital, the growth in labor in the 1990s, though substantial, is not atypical. There was a substantial increase in labor supply from every margin – population, participation, and the decline in the unemployment rate. Yet, notwithstanding the drop in the unemployment rate, the growth in labor input during the 1990s is quite similar to that in the 1980s boom. See Blank and Shapiro (2001) for a discussion of the role of labor supply in the recent expansion.

⁸ The change in real terms is even more dramatic owing to the very rapid price decline for information technology. Direct comparison of real quantities is, however, problematic owing to the chain-weighting of the national accounts. See Whelan (2000b).

Our analysis will show how adjustment to this new level of investment has had an important role in the timing of productivity change in the 1990s.

Note, moreover, that this increase in investment is not a typical cyclical phenomenon. Late in the booms of the 1960s and 1980s, the ratio of investment to GDP declined as strong output growth outpaced investment. In contrast, only as of early 2001 (a decade into the expansion) is there any sign of a peak of the investment rate in the current expansion, though Figure 2 shows a hint of a peak in non-information processing investment in 1998.

A key issue for our analysis is the cyclical pattern of factor utilization. As discussed below, our proxy for factor utilization is average weekly hours. The principle underlying our utilization correction is that firms will push on each utilization margin – hours of workers, hours of capital, and effort – so as to equalize the cost of using each margin of adjustment. The theory implies that an observed margin of adjustment, e.g., average weekly hours, can be used as a proxy for other margins of adjustment, such as effort or number of shifts, which are more difficult to measure. Figure 3 shows this series for manufacturing (available since 1947) and the private sector (available since 1964). Average weekly hours increase in the beginning of expansions, but reach their peak well before their end. (See also Sims, 1974.) Since quasi-fixed factors adjust as an expansion proceeds, firms can cut back on costly utilization margins, e.g., increasing hours via overtime. In the 1990s, there are two cycles in average weekly hours. They increased following the trough of the recession, but decreased in the middle of the decade, only to increase again when growth accelerated in the second half of the decade. As this second expansion matured, they again decreased.

The extraordinary performance of investment in the 1990s raises the possibility that adjustment costs play a nontrivial role in measuring technology. In the next section of the paper,

we outline our framework for measuring technology. Our method corrects productivity measurement for adjustment costs, factor utilization, and biases from scale economies and reallocations of inputs across industries. The following sections discuss the data, estimates, and results.

2. Theoretical Framework for Measurement

We seek to identify the technology component of the aggregate Solow residual in the 1990s. Only this component of productivity growth will increase production and income in the long run. In particular, although one must account for short-run deviations from constant returns and perfect competition in estimating a time series for aggregate technical change, growth in productivity due to increasing returns is not likely to be permanent. We also allow for two other biases in measuring technical change: Unobserved changes in the utilization of capital and labor and costs of adjusting the capital stock and the labor force.

Our method begins by estimating technical change at a disaggregated level, allowing for non-constant returns to scale, variations in utilization, and adjustment costs. Our approach modifies Hall (1990), which in turn generalizes Solow's classic (1957) paper. We then define aggregate technology change as an appropriately-weighted sum of the resulting sectoral residuals. Section 2.1 presents the basic framework, while sections 2.2 and 2.3 discuss how to correct the sectoral residuals for changes in utilization and costs of adjustment, respectively. Section 2.4 presents our method of aggregating sectoral technical change into an economy-wide

index, and section 2.5 discusses the relationship between productivity growth and technical change in the long run.⁹

2.1 Sectoral Technical Change

We assume that the representative firm in each sector has a production function for gross output of the form:

$$Y_i = F^i(S_i K_i, E_i H_i N_i, M_i, Z_i) (1 - \Phi^i(I_i/K_i)) (1 - \Psi^i(R_i/N_i)). \quad (1)$$

The firm produces gross output, Y_i , using its capital stock K_i , employees N_i , and intermediate inputs (materials and purchased services) M_i . We assume that the capital stock and number of employees are quasi-fixed, so their levels cannot be changed costlessly; the functions Φ and Ψ give the costs of adjustment, where I and R are, respectively, the flows of investment and hiring. At this level of generality, Φ and Ψ may be convex or non-convex.

Firms may, however, vary the intensity with which they use these quasi-fixed inputs: H_i is hours worked per employee; E_i is the effort of each worker; and S_i is the capital utilization rate (i.e., capital's workweek, the number of shifts). Total labor input, L_i , is the product $E_i H_i N_i$. The firm's production function F^i is (locally) homogeneous of arbitrary degree g in total inputs. If g exceeds one, then the firm has increasing returns to scale, reflecting overhead costs, decreasing marginal cost, or both. Z_i indexes technology.

Following Solow (1957), we implicitly define technical change as the fraction of output growth that cannot be attributed to the growth rate of inputs, via a first-order (log-linear)

⁹ Basu and Fernald (2001) provide detailed derivations and discussion of utilization, reallocation, and aggregation corrections.

approximation to the production function. For any variable J , we define dj as its logarithmic growth rate, approximated by $\ln(J_t / J_{t-1})$. The log-linearization shows:

$$dy_i = \left[\left(\frac{F_1 K}{F} \right)_i^* (dk_i + ds_i) + \left(\frac{F_2 (EHN)}{F} \right)_i^* (de_i + dn_i + dh_i) + \left(\frac{F_3 M}{F} \right)_i^* dm_i \right] - \mathbf{f}_i (di - dk)_i - \mathbf{y}_i (dr - dn)_i + dz_i, \quad (2)$$

where $\mathbf{f}_i = \left(\frac{\Phi' I}{1 - \Phi K} \right)_i^*$, and $\mathbf{y}_i = \left(\frac{\Psi' R}{1 - \Psi N} \right)_i^*$.

A star (*) indicates the steady-state value of a variable. We wish to express the steady-state output elasticities $F_J J / F$ as functions of observed variables, possibly up to a single unknown parameter. Note that the log-linear approximation requires that we treat the output elasticities as constants. Below we relate the output elasticities to observed factor shares and the markup or the degree of returns. In practice, we make the output elasticities constant by taking the sample averages of the annual shares and assuming that returns to scale do not vary over time.

A higher-order approximation to the production function (e.g., the second-order Törnqvist) would estimate time-varying elasticities by using the high-frequency variation in observed shares, but these would be meaningful only if we were sure that the high-frequency fluctuations in observed factor prices correspond to changes in allocative prices for the marginal factor. For reasons discussed by Hall (1980) and Carlton (1983), we doubt that the observed

cyclical changes in wages and materials prices are allocative.¹⁰ We therefore avoid these higher-order approximations, although in theory they are of course more accurate.

The part of equation (2) that may be least familiar is the role of adjustment costs in what is essentially a productivity growth relationship. To understand this issue better, note that we can write the drag on output growth coming from adjustment costs for, e.g., capital as:

$$\frac{\Phi'}{1-\Phi} d\left(\frac{I}{K}\right).$$

Note that the drag on output *growth* is proportional to the *change* in the investment-capital ratio. If investment is a constant fraction of the capital stock (as in steady state), then adjustment costs reduce the *level* of output, but do not affect its growth rate. When interpreting the time series on corrections to productivity growth coming from adjustment costs in Section 4, one needs to remember that this correction is most important when investment is *rising* (or falling), and not when investment is high but stable.

One of the seminal insights of Solow (1957) was that with constant returns, perfect competition, and (implicitly) no adjustment costs, the output elasticities can be observed directly – they are just the payments to each factor divided by total revenue. Hall (1990) dropped the assumption of constant returns to scale and perfect competition in the product market (but assumed price-taking in factor markets), and showed that cost-minimization implies:

$$\frac{F_J J}{Y} = \mathbf{m} \frac{P_J J}{PY} \equiv \mathbf{m}s_J \quad \text{for } J = K, L, M, \quad (3)$$

¹⁰ Since the zero-profit assumption is reasonable only in steady state, we would also have to estimate a time-varying rental cost of capital, allowing for the costs of adjustment we discuss in section 2.3. For a review of the complementary flexible functional form approach with time-varying output elasticities and adjustment costs, see Nadiri and Prucha (2001).

where \mathbf{m} is the markup of price over marginal cost and P_j is the (purchase or rental) price of each factor.¹¹ Note that the shares are defined to be the total *cost* of each factor divided by total revenue; thus, they correspond exactly to Solow's shares if and only if the firm makes zero profit. Otherwise, payments to factors that receive the profits (usually assumed to be capital, but possibly labor as well) will exceed the cost of hiring those factors.

We generalize the derivations of Solow and Hall by explicitly modeling the quasi-fixity of capital, K , and the number of workers, N , which changes our expressions for the output elasticities.¹² We use a dynamic cost-minimization framework to solve for the necessary Euler equations for capital and labor. See Appendix A for details. We find that the steady-state elasticities are:

$$\left(\frac{F_1 K}{F}\right)_i^* = \mathbf{m}_i^* \left(\frac{(r + \mathbf{d})q}{PY}\right)_i^* - \mathbf{f}_i = \mathbf{m}_i^* s_{Ki}^* - \mathbf{f}_i \quad (4)$$

and

$$\left(\frac{F_2(EHN)}{F}\right)_i^* = \mathbf{m}_i^* \left(\frac{WL}{PY}\right)_i^* - \mathbf{y}_i + r^* \left(\frac{\mathbf{y}N}{R}\right)_i^* = \mathbf{m}_i^* s_{Li}^* - \mathbf{y}_i + r^* \left(\frac{\mathbf{y}N}{R}\right)_i^*, \quad (5)$$

where q^* is the steady-state shadow value of capital (which may not equal the purchase price of investment goods), and r^* is the steady-state real interest rate.

Some intuition for the capital elasticity comes from noting that the markup times capital's share (the first term on the right-hand-side of equation (4)) equals the elasticity of market output,

¹¹ These derivations require no assumption about the form of imperfect competition (e.g., static monopoly or monopolistic competition). For the purpose of cost-minimization the markup is treated parametrically, although of course the firm solves for the optimal markup as part of its more complex dynamic profit-maximization problem.

¹² As explained below, we assume that S , E , H and M may be varied without cost of adjustment (although the firm may face upward-sloping factor supply curves or increasing user costs).

Y , with respect to capital. But in our setup, the elasticity of Y with respect to K does not equal the elasticity of F with respect to K , since having more capital also reduces the adjustment cost of investment. Thus, ms_K strictly exceeds the elasticity of F with respect to K , by an amount equal to f_i , the elasticity of adjustment costs with respect to investment.

The intuition for the labor elasticity is similar. Note that the elasticity with respect to N has a similar form to the elasticity with respect to K , but with the subtraction of one additional term. That term represents the annuity value of the adjustment cost of hiring an additional worker. The reason that this term appears in the expression for labor but not for capital is that capital implicitly pays its own adjustment cost, but the firm must pay the adjustment cost for labor. Since hiring an extra worker now leads to costs as well as benefits, one cannot *a priori* sign the difference between ms_L and $F_L L/F$. Depending on parameter values, the difference may be positive or negative.

In order to construct the shares s_{ji}^* , one generally needs to calculate the rental cost of capital. We avoid this difficulty by assuming that firms make zero economic profits in the steady state, so we take total cost as approximately equal to total revenue and follow Solow (1957) in estimating capital's share as a residual. Rotemberg and Woodford (1995) discuss a variety of evidence supporting the zero-profit assumption.

Substituting in our expressions for the output elasticities, we can write equation (2) in terms of the markup of price over marginal cost, which comes directly from the optimization. But we could equally easily, and without additional assumptions, express (2) in terms of returns to scale, which is perhaps more intuitive when discussing a technological relationship. Returns to scale g and the markup m are linked by the following equation:

$$m(1 - s_{pi}) = g, \tag{6}$$

where s_{p_i} is the share of pure profits in gross output and the markup equals one if the firm is perfectly competitive. Thus, as in Chamberlinian monopolistic competition, firms can charge markups even if entry eliminates pure profits, but only if average cost exceeds marginal cost. Since we are assuming that s_{p_i} is zero in steady state, we can interpret the coefficient multiplying the shares as the degree of returns to scale, \mathbf{g} .

We describe in Section 2.3 how we use additional information to calibrate the adjustment elasticities \mathbf{f} and \mathbf{y} . Since terms in these parameters effectively become observable, we put them on the left-hand side of our estimating equation. Thus, that equation is:

$$\begin{aligned}
& dy_i + \mathbf{f}_i di_i + \mathbf{y}_i (dr_i + dh_i) - \left(\frac{r\mathbf{y}N}{R} \right)_i^* (dn_i + dh_i) \\
&= \mathbf{g}_i^* \left[s_{K_i}^* dk_i + s_{L_i}^* (dn_i + dh_i) + s_{M_i}^* dm_i \right] \\
&+ \mathbf{g}_i^* \left[(s_{K_i}^* - \mathbf{f}_i) ds_i + (s_{L_i}^* - \mathbf{y}_i + (r\mathbf{y}N/R)_i^*) de_i \right] + dz_i \\
&\equiv \mathbf{g}_i^* dx_i + \mathbf{g}_i^* du_i + dz_i
\end{aligned} \tag{7}$$

Writing the equation in this form makes it somewhat easier to interpret why terms like $\mathbf{f} di$ should appear symmetrically with the growth rate of market output dy . The reason is that the firm is actually producing two types of output – goods and services that it sells, and the internally-used services of installing its new investment goods and training its new hires. Total output correctly computed should add these implicit services to the explicit market output (see also Abel, 1999).

Note that in principle, it does not matter whether the firm produces these installation services in-house – using its own capital or labor – or out-sources them – hiring as an intermediate input a systems integration firm, say, to install a new computer system. In either case, the firm has a joint product of observed market output and unobserved installation services. Outsourcing the installation rather than doing the task in-house means the firm *purchases* those

installation services, so it affects the composition of inputs dx_i on the right-hand-side of (7), as the firm substitutes purchased intermediate services for its own capital and labor.

The next subsection presents our method for measuring du_i . Lack of an observable counterpart to this variable is the only barrier to estimating equation (7). Our primary interest is to measure true technical change, dz_i , which we cannot observe directly.

2.2 *Controlling for Variable Factor Utilization*

Utilization growth, du_i , is a weighted average of capital utilization, S_i , and labor effort, E_i . The challenge in estimating firm or sectoral technology change using equation (2) is to relate du_i to observable variables. We do so following Basu and Kimball (1997), who use the basic insight that a cost-minimizing firm operates on all margins simultaneously, so the firm's first-order conditions imply a relationship between observed and unobserved variables. Thus, increases in observed inputs can, in principle, proxy for unobserved changes in utilization. Shapiro (1986b) shows that the excess output elasticity of worker hours relative to the number of workers disappears when one explicitly accounts for changes in the number of shifts, which provides direct support for the usefulness of hours per worker as a proxy for capital utilization.

Basu and Kimball (1997) show that under plausible conditions there is a simple, observable proxy for variations in utilization. In particular, if the sole cost of changing the workweek of capital (the number of shifts) is that workers need to be compensated for working at night (a shift premium), then changes in hours per worker proxy appropriately for unobserved changes in *both* effort and capital utilization.¹³ We are thus able to control for variable utilization

¹³ Basu and Kimball (1997) generalize this model to allow the rate of depreciation to vary depending on capital utilization, as in a variety of papers. This modification introduces two new terms into the estimating equation at the end of this section, but Basu and Kimball cannot reject the hypothesis that these terms are insignificant. In any case, we have noted in previous work

without assuming, unrealistically, that one can observe either the firm's internal shadow prices of capital, labor and output, or the true quantities of capital and labor input at high frequencies. The Basu-Kimball model uses only the cost-minimization problem and the assumption that firms are price-takers in factor markets; it does not require any assumptions about the firm's pricing and output behavior in the goods market.

Details of the model are given in Appendix A. Here we present the solution, and discuss the intuition. The appendix shows that

$$du_i = \left[\frac{\mathbf{h}_i}{\mathbf{n}_i} (s_{Ki}^* - \mathbf{f}_i) + \mathbf{z}_i (s_{Li}^* - \mathbf{y}_i + (r\mathbf{y}N/R)_i^*) \right] dh_i, \quad (8)$$

where du_i is the share-weighted change in capital and labor utilization, as defined in equation (7) and, as before, dh_i is the growth rate of hours per worker.

The parameter \mathbf{z}_i is the elasticity of effort with respect to hours per worker, reflecting the fact that observed hours per worker should be an excellent proxy for labor effort. From the firm's perspective, labor input is symmetric – it can come from either greater work intensity or a longer workday. In the model in the appendix we assume, loosely speaking, that the worker's disutility of labor is convex along both margins. For example, workers' marginal disutility for working an extra hour per day increases as the workday becomes longer, thus requiring a firm to pay an increasingly large "overtime premium" as it lengthens its workday. Under these assumptions, it always pays a firm to push along both margins simultaneously when it needs more labor input from a fixed labor force – that is, to increase both hours per worker and effort.

(Basu, Fernald and Kimball, 1999) that including these terms barely affects estimates of technical change.

Equation (8) also says that the change in hours per worker should be a proxy for changes in the unmeasured workweek of capital, with $\mathbf{h}_i/\mathbf{n}_i$ the elasticity of shifts with respect to hours per worker. The parameter \mathbf{h}_i indicates the rate at which the elasticity of labor costs with respect to hours increases and \mathbf{n}_i indicates the rate at which the elasticity of labor costs with respect to capital utilization increases. The reason that hours per worker proxies for capital utilization as well as labor effort is that shift premia create a link between capital hours and labor compensation. The shift premium is most worth paying when the marginal hourly cost of labor is high relative to its average cost, which is the time when hours per worker are also high. The faster the rate of growth of the overtime premium relative to the rate of growth of the shift premium, the more firms will tend to use capital utilization as the dominant margin of adjustment.

Putting everything together, we have an estimating equation that controls for variable utilization:

$$\begin{aligned}
& dy_i + \mathbf{f}_i di_i + \mathbf{y}_i (dr_i + dh_i) - \left(\frac{r\mathbf{y}N}{R} \right)_i^* (dn_i + dh_i) \\
&= \mathbf{g}_i^* dx_i + \mathbf{g}_i^* \left[\frac{\mathbf{h}_i}{\mathbf{n}_i} (s_{Ki}^* - \mathbf{f}_i) + \mathbf{z}_i (s_{Li}^* - \mathbf{y}_i + (r\mathbf{y}N/R)_i^*) \right] dh_i + dz_i \\
&\equiv \mathbf{g}_i^* dx_i + \mathbf{x}_i dh_i + dz_i
\end{aligned} \tag{9}$$

This specification controls for both labor utilization and capital utilization, as well as costs of adjustment, non-constant returns to scale, and imperfect competition. Since we treat the shares and other parameters as constant, we can simply estimate a single coefficient on dh . This procedure does not allow us to estimate the structural parameters governing changes in utilization, but that is not our objective. We do estimate the markup \mathbf{g} that, as we show below, is needed to aggregate our estimates of sectoral technical change.

2.3 *Calibrating Adjustment Costs*

Equation (7) can, in principle, be estimated directly to control for costs of adjustment (by putting the terms in \mathbf{f} and \mathbf{y} on the right-hand side). As we discussed above, these costs can camouflage the extent to which technical progress has been faster in a period like the 1990s, which has also seen large increases in the rates of investment and hiring. However, our attempts to estimate equation (7) directly almost never produced estimates of adjustment costs that are significant statistically. We do not take this evidence as dispositive, however, since we are disregarding the main sources of information usually used to estimate adjustment costs. These are typically estimates of the Euler equation for capital or a regression of investment on Tobin's q .

In order to incorporate the information from these studies of adjustment costs, we calibrate the \mathbf{f} and \mathbf{y} terms instead of estimating them. We base this calibration on the marginal adjustment cost function, that is, the derivative of output with respect to investment. The marginal adjustment cost function is a well-known object in the context of the capital accumulation literature. In the Euler equation literature, it is directly estimated. In the q -theoretic literature, the q investment function is derived by inverting the condition that marginal adjustment cost equals q . In this paper, we calibrate our parameters based on estimates of an Euler equation. The Euler equation attempts to quantify directly the parameter we require, that is, the output loss from factor adjustment. Moreover, the Euler equation estimates do not suffer from the incredibly high estimates of adjustment costs typically found in regressions of q on the investment rate. Hall (2001b) follows the same procedure to calibrate adjustment costs in his recent work on intangible capital.

Recall that

$$\mathbf{f} = \left(\frac{\Phi'}{1 - \Phi} \frac{I}{K} \right)$$

is the elasticity of the adjustment cost with respect to the investment rate. Marginal adjustment cost, the partial derivative of output with respect to investment, is related to \mathbf{f} as

$$\frac{\partial Y}{\partial I} = -\frac{\Phi'}{1-\Phi} \frac{Y}{K} = -\mathbf{f} \frac{Y}{I}. \quad (10)$$

Shapiro (1986a) estimates marginal adjustment cost from the Euler equation for capital accumulation. His parameterization of marginal adjustment cost is as follows

$$\frac{\partial V}{\partial I} = -g_{kk} I \cdot V \cdot (1-t_k) \quad (11)$$

where V denotes real value added (as opposed to real gross output Y). Using the steady-state approximation that real value added is real gross output times one minus the materials share,

$V = (1-s_M)Y$, combining equations (10) and (11) allows us to calibrate \mathbf{f}_i for each industry i as

$$\mathbf{f}_i = \frac{g_{kk}(1-t_k)I^2}{1-s_{Mi}} \quad (12)$$

where $g_{kk} = 0.0015$ is Shapiro's estimate of the the adjustment-cost parameter, s_{Mi} is the industry-specific average materials share from our data, and t_k and I are set to their average values of 0.49 and 4.83 in Shapiro's data. On a value-added basis, (i.e., abstracting from the materials share adjustment), the estimate of \mathbf{f} is therefore equal to 0.018 quarterly rate. Though this estimate implies relatively rapid adjustment to steady state, adjustment costs are not trivial. They account for 0.7 percent of output or about 9 percent of the cost of investment.

Our data are annual. To convert the coefficient to an annual rate, we multiply it by $1 + I + I^2 + I^3$, where I is the root of the Euler equation. Shapiro's estimates imply a root of 0.75, so the annualization factor is 2.7; this factor is smaller than mechanical multiplication by 4 owing to within-year adjustment of the capital stock. Therefore, the estimate of \mathbf{f} is 0.05 at an annual rate on a value-added basis.

Recall that f is calculated at the steady-state I/K ratio. Thus, the finding that f is non-zero implies that marginal q cannot equal one, even in steady state.

The case of labor adjustment costs, however, is rather different. First, Shapiro's estimates of labor adjustment costs are mixed. He finds that there are significant labor adjustment costs for non-production workers, but not for production workers or their hours. The data sets available to us do not separately report the employment of production and non-production workers. Second, Shapiro's estimates pertain to gross flows of employment, as must any sensible model of adjustment costs – it cannot be the case that a firm that hires 1,000 employees and fires 1,000 other employees suffers the same adjustment cost as a firm that does zero hiring. However, we are again constrained by data limitations – there is no consistent time series of gross flows of employment by industry for the entire economy. For these reasons, we make the assumption that the marginal adjustment cost for labor *is* zero in the vicinity of the steady state. We freely admit that this assumption is strong and in some respects unsatisfactory, and hope that future researchers armed with more extensive data sets will revisit this issue. In terms of the calibration, therefore, we henceforth assume that $\mathbf{y} = 0$. Note that out-of-steady state adjustment costs can nonetheless rationalize the variability of hours per work required for our utilization adjustment.

2.4 *Aggregate Technical Change*

So far we have sketched a framework for measuring firm-level (in practice, industry-level) technology shocks dz_i , the residual in equation (2). The natural question, therefore, is how the sectoral shocks in the production of gross output are related to aggregate technical change in the production of value added, in an economy with increasing returns to scale and imperfect competition. This section gives a condensed version of the analysis of aggregate productivity

growth with imperfect competition found in Basu and Fernald (1997), and uses that analysis to motivate our definition of aggregate technical change.

Define value added as a Divisia index (that is, in growth rates in continuous time; the chain-linked U.S. national income accounts approximate this index in discrete time):

$$dv_i = \frac{dy_i - s_{Mi}dm_i}{1 - s_{Mi}} = dy_i - \left[\frac{s_{Mi}}{1 - s_{Mi}} \right] (dm_i - dy_i). \quad (13)$$

From the production side of the national accounts identity, aggregate output is a Divisia index of firm-level value-added. In growth rates:

$$dv \equiv \sum_{i=1}^N w_i dv_i, \quad (14)$$

where w_i is the firm's share of nominal value added:

$$w_i \equiv \frac{P_i^V V_i}{P^V V}. \quad (15)$$

The aggregate Solow residual is then defined as:

$$dp \equiv dv - dx^V, \quad (16)$$

where dx^V is the growth rate of aggregate capital and labor inputs, multiplied by the shares of each factor in national income. Note that since we are now considering the shares of capital and labor in value added, the shares used to compute dx^V sum to 1. (Henceforth we use a superscript V to denote the value-added version of variables previously defined in a gross-output context.)

We now need to consider the value-added-augmenting form of increasing returns and technical change. Since value added is only a fraction of gross output, increasing returns in the production of gross output translates to larger returns increasing returns in the production of value added, and a gross-output-augmenting technology shock is a larger value-added-

augmenting technology shock. Rotemberg and Woodford (1995) show that the precise formulas are:

$$\mathbf{g}_i^V = \mathbf{g}_i \frac{1 - s_{Mi}}{1 - \mathbf{g}_i s_{Mi}}, \text{ and}$$

$$dz_i^V = \frac{dz_i}{1 - \mathbf{g}_i s_{Mi}}.$$

We also define P_L as the economy-wide average hourly wage (total payments to labor divided by total labor hours) and P_K as the average payment to a unit of capital (total payments to capital divided by the aggregate capital stock).

Basu and Fernald (1997) show that taking a simple weighted sum of equation (2) over the entire economy and substituting definitions (14)-(17) gives the following decomposition for aggregate productivity:

$$\begin{aligned} dp &= R + du + da + dz^V \\ &= (\bar{\mathbf{g}}^V - 1)dx^V + R_g + R_M + \bar{\mathbf{g}}^V R_K + \bar{\mathbf{g}}^V R_L + du + da + dz^V \end{aligned} \tag{17}$$

where

$$\begin{aligned}
\bar{\mathbf{g}}^V &= \sum_{i=1}^N w_i \mathbf{g}_i^V, \\
R_g &= \sum_{i=1}^N w_i (\mathbf{g}_i^V - \bar{\mathbf{g}}^V) dx_i^V, \\
R_M &= \sum_{i=1}^N w_i (\mathbf{g}_i^V - 1) \left[\frac{s_{Mi}}{1 - s_{Mi}} \right] (dm_i - dy_i), \\
R_K &= \sum_{i=1}^N w_i \frac{s_{Ki}}{1 - s_{Mi}} \left[\frac{F_{Ki} - P_K}{F_{Ki}} \right] dk_i, \\
R_L &= \sum_{i=1}^N w_i \frac{s_{Li}}{1 - s_{Mi}} \left[\frac{F_{Li} - P_L}{F_{Li}} \right] (dn_i + dh_i), \\
du &= \sum_{i=1}^N w_i \frac{\mathbf{g}_i du_i}{1 - \mathbf{g}_i s_{Mi}}, \\
da &= \sum_{i=1}^N w_i \frac{1}{1 - \mathbf{g}_i s_{Mi}} (\mathbf{f}_i (di - dk)_i + \mathbf{y}_i (dr - dn)_i), \text{ and} \\
dz^V &= \sum_{i=1}^N w_i dz_i^V.
\end{aligned}$$

The key point to note about equation (17) is that it is qualitatively different from the firm-level equation (2). With a simple rearrangement (subtracting dx_i from both sides) equation (2) relates measured productivity growth in gross output to a scale effect, utilization and adjustment corrections, and technical change. Analogous terms are present in the expression for aggregate productivity growth, but the R_M , R_g and R_K and R_L terms are new. The R_M term corrects for mismeasurement in value added: Value added is computed using s_{Mi} as a measure of the output elasticity of materials, but with increasing returns the output elasticity exceeds material's share. The R_g term shows that the effect of increasing returns on productivity growth is not confined to the average scale effect, the first term in equation (17), because if returns to scale vary across sectors, productivity growth receives an extra boost if the sectors that have above-average increasing returns also experience above-average input growth.

The R_K and R_L terms reflect possible heterogeneity in the marginal products of identical factors. We begin with the labor heterogeneity term, R_L . It is easier to explain the intuition behind this term if we assume that there is a single type of worker, although our empirical

procedure recognizes and corrects for heterogeneity in worker attributes. The Solow residual calculation assumes that the marginal product of all workers is identical, and equal to the economy-wide average real wage. Now suppose that in some sectors workers are unionized, or firms pay efficiency wages. Either would raise the industry wage rate above the competitive level. Since firms would still equate their marginal revenue products of labor to the wage, marginal products would be higher than average in the sectors with labor market imperfections (holding everything else constant). Even if all labor markets are competitive, the fact that there are adjustment costs of capital and labor would also prevent the marginal products of labor from equalizing at all points in time. In such situations the distribution of input growth would again matter: Productivity growth will receive an extra boost if labor input growth is higher in sectors where labor has above-average marginal product. The intuition for the R_K term is identical, although we believe that the heterogeneity in capital's marginal product is due solely to adjustment costs.

Equation (17) highlights an unfortunate implication of imperfect competition and heterogeneity: It shows that there is no unambiguous definition of aggregate technical change in an imperfectly competitive economy. Economists are used to defining economy-wide technical change either from a top-down perspective, as the change in value added for given aggregate inputs of capital and labor, or from a bottom-up perspective, as an appropriately-weighted average of technical change at every firm. Under perfect competition these two approaches coincide, but this is no longer the case with imperfect competition. The bottom-up perspective says that aggregate technical change is just dz^V . But, as equation (17) shows, the ability of the economy to produce final goods and services for given inputs of capital and labor generally fluctuates if there are changes in markups or in the distribution of inputs across heterogeneous

firms, even if aggregate inputs are held fixed and technology at every firm remains unchanged.

Thus, the top-down perspective would add all the R terms to dz^V .

However, in the next section we argue that only dz^V is responsible for *long-run* changes in productivity, which are our primary interest in this paper. We therefore define our measure of aggregate technical change as equal to dz^V only.¹⁴

Given this definition, equation (17) shows why we implement our corrections with disaggregated data. Measuring dz^V properly requires us to estimate the dz_i at a sectoral level, and then aggregate using the observed shares, w_i , and the estimated returns to scale parameters g .

2.5 *Productivity Growth and Returns to Scale in the Short and Long Run*

The focus of our paper is on long-run productivity. Productivity growth is the key to rising living standards; Basu and Fernald (1997) show that productivity growth – measured almost exactly as in Solow (1957) – is the right measure of welfare change for a representative consumer. The intuition for this result is that the consumer values extra output at the marginal utility of wealth, but dislikes supplying more labor or capital by an amount equal to the marginal utility of wealth multiplied by the real wage or the real rental rate. Thus, the relative prices used to compute the Solow residual measure the consumer’s marginal rate of substitution between output and inputs even when – due to imperfect competition or increasing returns – they fail to

¹⁴ We can show that dz^V is the unambiguous measure of technical change (in the sense that the top-down and bottom-up definitions coincide) in certain models with imperfectly competitive product markets (e.g., Rotemberg and Woodford’s (1995) dynamic general-equilibrium model). However, this result is true only under certain restrictive conditions, which are satisfied in the Rotemberg-Woodford model: factor markets must be competitive, and all firms must have identical separable gross-output production functions, charge prices that are the same markup over marginal cost, and always use intermediate inputs in fixed proportions to gross output.

measure the social marginal rate of transformation between these goods.¹⁵ Importantly, to the extent that productivity growth and technical change differ, it is productivity, not technology, that is the key to welfare. For example, the consumer does not care whether a boost in productivity is due to superior technology or greater exploitation of scale economies – the extra output still confers the same utility, and supplying the extra inputs still gives the same disutility.¹⁶

Thus, given that we wish to predict long-run productivity levels, our emphasis on measuring technical change *per se* may seem misplaced. As we showed above, short-run productivity (correctly measured, hence net of utilization) can change for a variety of reasons unrelated to technical change.

We have in mind, however, a model where the non-technological effects in equation (17) do not operate in the long run, so that so that over long horizons productivity is driven solely by technology. In such a model, the distribution of inputs over firms of the same type (e.g., with the same size markup) does not change in the long run, and neither does output per firm. Whenever a shock to the economy increases demand, since the number of firms is roughly fixed in the short run, the increased production mandates a higher output per firm. But if firms produce with fixed costs of operation, the increased output will lead to increases in economic profit. In the model we have in mind, the increased profit would spur entry, and drive per-firm output and profits down to their steady-state levels. This is essentially the canonical model of Chamberlinian

¹⁵ This discussion has ignored taxes on wages and capital rents, which drive a large wedge between the after-tax marginal products relevant for consumer optimization and the before-tax payments used to measure factor shares in the standard Solow residual. If the Solow residual is to measure welfare, it must be calculated using factor prices as perceived by the consumer, i.e., after-tax prices.

¹⁶ Productivity growth must be measured net of utilization. For example, both the firm and the worker perceive extra effort as additional labor input, which requires additional compensation.

imperfect competition, embedded in a growth setting. As we argue next, this kind of model makes the distinction between short- and long-run returns to scale advocated by Solow:

I would be inclined to try very hard to formulate some sort of increasing returns to scale in the short run (when output is below capacity), combined with a closer approximation to constant returns to scale in the long run (Solow, 1990, p. 225).

It is worth elaborating on this point, since it is conceptually important for our approach. Take, for example, the issue of whether we should expect increasing returns at the firm level to drive a wedge between productivity and technology in the long run. By definition, returns to scale equals the ratio of average cost to marginal cost, and increasing returns implies that average cost falls with scale. Abstracting from adjustment costs for simplicity, assume that firms of type i have the production function

$$Y_i = F^i(S_i K_i, E_i H_i N_i, M_i, Z_i) = G^i(S_i K_i, E_i H_i N_i, M_i, Z_i) - C_i, \quad (18)$$

where C_i is a fixed (per-period) cost and G is homogeneous of constant degree \mathbf{r}_i . Then, for

this type of firm, returns to scale at any point in time, \mathbf{g}_i , equals $\mathbf{r}_i \frac{Y_i + C_i}{Y_i}$.

In order for increasing returns to contribute to long-run productivity growth, firms must always expand their scale of operations, thereby reducing unit costs forever. First, this is impossible if (as we think likely) increasing returns are due solely to fixed costs (i.e., $\mathbf{r} = 1$). As these costs are spread over more and more units of output, returns to scale must fall asymptotically to \mathbf{r} . Second, even if $\mathbf{r} > 1$, as long as $C > 0$ firms will make ever-larger profits

The division of inputs into observed factors and unobserved utilization is motivated solely by issues of measurement.

as their scale of operations expand indefinitely (assuming no change in their degree of market power). But presumably these rising profits would trigger entry by other firms, thus limiting the expansion in the size of existing firms. As new firms enter, scale economies would be reduced because per-firm output would fall. If output returns exactly to its pre-shock level at each type of firm i , then firm-level returns to scale in that group will return to the pre-shock g_i .

In such a world, however, long-run *aggregate* returns to scale would always be 1, regardless of the returns to scale at any firm. (Intuitively, assuming zero economic profits in the long run, the long-run industry supply curve is horizontal at the steady-state level of average cost, even if average cost exceeds marginal cost at every firm.) Thus, aggregate returns to scale would tend back to their steady-state value of 1 after a shock as the number of firms adjusts. Furthermore, assuming homothetic demand, the distribution of inputs across the different categories of output would also be unchanged, thus implying no long-run reallocation effects.

This of course raises the interesting question of the length of time that this process of adjustment might take. If entry is relatively slow, the economy might reap extra productivity benefits by exploiting firm-level increasing returns for a non-negligible period of time. On the other hand, if new firms also produce new products, correctly-measured productivity growth might suffer because the steady-state number of varieties is not available during the transition period. Our decision not to model explicitly the process of adjustment of aggregate returns to scale can be justified if we assume that these two effects roughly offset one another.

Rotemberg and Woodford (1995) show that one can formalize a simpler version of the model we have sketched verbally. (In their model firms are not heterogeneous.) Thus, we take as a working approximation the hypothesis that in the long run productivity and technology coincide. However, when measuring technical change, it is important to allow for short-run

deviations from constant returns. (For example, in the Rotemberg-Woodford model the number of firms is essentially fixed in the short run, so with firm-level increasing returns, short-run productivity growth exceeds technical change when output is rising.) Thus, we allow for deviations from constant returns and perfect competition in the short run when *estimating* the extent of technical change from data over short periods of time, as in the recent productivity boom that is our primary interest. But in *projecting* the long-run impact of that boom, we believe that only the technical change is permanent.

3. Data and Estimates

In selecting the data for this paper, we face a trade-off between the span of the data and its timeliness. There are two well-established sources of data for U.S. multifactor productivity at the aggregate and sectoral level – the BLS’s multifactor productivity data, and the data assembled by Dale Jorgenson and his associates.¹⁷ Both datasets control carefully for changes in the composition of the capital stock and the labor force. Nevertheless, for our purposes, both datasets suffer disadvantages. The Jorgenson industry dataset covers the entire U.S. economy at a disaggregated level, but unfortunately, ends in 1996 – too early to address the questions in this paper. The BLS dataset now runs through 1999,¹⁸ but its detailed industry coverage excludes non-manufacturing – the vast majority of the economy.

Given these shortcomings, we construct our own dataset on industry inputs and outputs.

We obtain data on industry gross output and intermediate inputs from the Bureau of Economic

¹⁷ See Jorgenson, Gollop, and Fraumeni (1987) for a detailed discussion of the construction of these data. Jorgenson and Stiroh (2000) provide a recent use and discussion of these data, which we will henceforth refer to as the “Jorgenson-Stiroh” dataset.

¹⁸ The updated data through 1999 were released in April 2001, after this paper was completed.

Analysis's Gross Product Originating dataset. On a consistent basis, the BEA data are available only from 1987 through 1999. We combine these data with estimates of labor hours from the BLS and estimates of industry capital stocks from the BEA. Appendix B describes our dataset (which we will refer to as the "BEA dataset") in greater detail. Relative to the BLS and Jorgenson-Stiroh datasets, the major shortcomings of our dataset are the short time period of coverage (1987-1999) and the lack of controls for changes in the composition of capital and labor.

Since our sample period with the BEA dataset is too short for reliable estimation, we use the Jorgenson-Stiroh dataset for all estimation. Those data are available from 1959 to 1996. (We begin all estimation in 1965, however, because that is the first year for which we have consistent data on hours growth for all industries, as discussed below.) Note that it would be inappropriate to splice the BEA data onto the Jorgenson-Stiroh data. The BEA's data are based on the recent rebenchmark of the National Income and Product Accounts while the Jorgenson-Stiroh are not. In addition, they have conceptually different measures of capital and labor. Finally, we do not want to splice the data at the same point where there might have been a shift in underlying growth rates. Hence our strategy with respect to the data is as follows:

- Estimate parameters using the long span of the Jorgenson-Stiroh data.
- Apply these estimated parameters to the BEA data.

Though we do not report results based on BLS's data, we have referred to them to check our results.

As discussed in Section 2, the equation we use to extract an estimate of technological change from output growth, input growth, and adjustment terms is as follows:

$$dy_i + \mathbf{f}_i di_i = \mathbf{g}_i^* dx_i + \mathbf{x}_i dh_i + dz_i$$

where i index industry, time subscript is suppressed, and an industry-specific constant is suppressed. The residual, dz_i , plus the industry-specific constant, is our estimate of the period-by-period growth in technology for industry i . We estimate the returns to scale/markup parameter, g_i^* , and the utilization parameter, x_i . We calibrate the adjustment cost parameter f_i using the procedure discussed in Section 2.3. For data on output growth, dy_i , and share-weighted input growth, dx_i , we use Jorgenson and Stiroh's data. For the capital adjustment term, di_i , we use industry data from the BEA's Tangible Wealth dataset.

For the dh_i term in the utilization adjustment, we used the growth rate of average weekly hours of production workers (where the log level is detrended with a Hodrick-Prescott filter), since that is the series that is available for the full sample period. (See Appendix B for further details.) Hours per worker has low frequency movements not related to cyclical utilization. For the private sector as a whole, there is a downward trend in hours per worker until about 1990, when the trend flattened. See Figure 3. In manufacturing, the pattern is reversed. Hours per worker are flat until 1990 and then may show an upward trend. Since our utilization measure is the cyclical change in hours per worker, getting the trend right is important. Given the change in the trend at the end of the sample, there is substantial uncertainty about the correct trend.

To estimate the equation, we allow for heterogeneity in the coefficients where there is sufficient variation in the data and signal in the instruments to provide precise estimates. In other cases, we pool the estimates across industries. For our main results, we allow the returns to scale parameter g_i^* , and utilization parameter, x_i , to vary across the broad sectors, specifically, durable manufacturing, non-durable manufacturing, and non-manufacturing. Within these

sectors, we constrain the coefficients to be equal.¹⁹ The calibrated adjustment cost parameters, f_i , reflect industry heterogeneity in materials shares as in equation (12). The estimated equation also includes industry-level constants to allow for differences in growth rates across industries, i.e., the estimated equations have industry-level fixed effects. These constants are added back in to our estimate of growth in technology.

The returns-to-scale and utilization parameters are estimated by instrumental variables owing to the simultaneous determination of inputs and technology. Valid instruments need to be correlated with the various transformations of inputs on the right-hand side of the equation above and uncorrelated with technology. We use oil prices (current and one lag), defense spending shocks [current and one lag; from Ramey and Shapiro (1998)], and monetary policy shocks [one lag; from Christiano, Eichenbaum and Evans (1996)] as instruments.²⁰ These should be correlated with demand for inputs, but not with technology. The instruments appear to be adequate in terms of first-stage fit.

Table 1 gives the estimates of the returns-to-scale and utilization parameters.

The returns to scale (also the markup) parameter g_i^* is precisely estimated. Outside of non-durable manufacturing, where it is significantly below one, the estimated parameter is very close to one. These estimates confirm recent finding that widespread increasing returns to scale are

¹⁹ In earlier work, Basu and Fernald allow for more heterogeneity across industries. In the preliminary work for this paper we found that allowing for this heterogeneity does not have much effect on the aggregate estimate of technology although it does make a difference for the industry-by-industry estimates of scale versus utilization effects.

²⁰ We thank Charles Evans for providing us with updated quarterly estimates of money shocks from 1960 through 1999. We sum the quarterly series to obtain an annual series. In principle, we could use the four quarterly shocks as instruments, but the first-stage F-statistic falls sharply.

generally absent in the data. The estimates of the utilization-correction parameter α_i are less precise, though taken together, they are statistically significant.

4. Corrections to Productivity and the Growth of Technology: Results

4.1. Aggregate Results

The purpose of the estimates presented in the previous section is to yield an estimate of the growth in technology. The aggregate estimate of technology is, rearranging equation (17),

$$dz^V = dp - du - da - R$$

where dp is the Solow residual, du is the utilization correction, da is the adjustment-cost correction, and R is the reallocation/returns-to-scale correction. The corrections du , da , and R are based on the estimated and calibrated parameters. To produce the sectoral and private-industry aggregate, the industry-level estimates are aggregated according to the formulas given in Section 2.4.

Figures 4 through 7 report the year-by-year estimates of technology and its components. Table 2 gives the estimates of growth in technology averaged over various periods. Table 3 decomposes these estimates of the growth in technology into growth in output, inputs, and the corrections to the Solow residual.

The top left panel of Figure 4 shows the growth rates of output and input and the Solow residual. The bottom left panel cumulates these growth rates to yield log levels normalized to be 100 in 1989, the peak of the expansion of the 1980s. We use this same format with growth rates in the top panel and log levels in the bottom panel in all of Figures 4 through 7. Figure 4 shows the decline in output and input going into the recession and sustained recovery following the trough in early 1991. The Solow residual, except for a blip in 1993, has a similar pattern. Its

growth increases in the second half of the 1990s as the gap between output and input growth widens. The question for this paper is, of course, how much of this increase represents acceleration in technology and how much is due to cyclical factors?

The middle panels of Figure 4 give the corrections to the Solow residual that allow us to give an answer to this question. The top middle panel shows the three corrections to the Solow residual. These corrections – for utilization, adjustment, and reallocation – are *subtracted* from the Solow residual to provide an estimate of the period-by-period growth rate in technology. Again, the second panel of the figure cumulates these percent changes. The utilization correction is, as expected, highly cyclical. In the peak year of 1989, utilization added over a percentage point to output and hence would lead the Solow residual to overstate technological change. During the recession years of 1990 and 1991, utilization is a subtraction from output. It rebounds sharply during the recovery year of 1992 and bounces back and forth during the remainder of the decade. This behavior of utilization is what one would expect for a mature recovery. As capacity expands, firms can economize on utilization. There is a second peak in utilization when output accelerated in the second half of the 1990s.

The other corrections are much less volatile. In the recession year of 1990 and in 1991, the first year of the weak recover, the adjustment-cost correction is neutral to weakly positive, that is, lower than steady-state investment was a net boost to measured output growth. With the onset of the investment boom in 1993, it was a drag on growth and therefore causes the Solow residual to understate the growth in technology. From 1992 forward, the adjustment-cost correction is negative and increasing, so it gives an increasing boost to our estimate of growth in technology relative to the growth in the Solow residual.

The final correction is what we term reallocation. This term is driven by movements of inputs across industries when the production function is not constant returns to scale. The reallocation correction uniformly adds to output (i.e., leads to an understatement of growth in technology were this term to be neglected). Again, this correction is being driven almost exclusively by the estimated diseconomies of scale in non-durable manufacturing. There is a slight acceleration of the correction in 1996, but since this correction is otherwise so smooth it simply affects the baseline and not the timing of the growth in technology.

In the top middle panel of Figure 4, the high volatility of utilization makes it appear to swamp the other two corrections. The persistence of the other corrections, however, makes this impression incorrect. The bottom panel shows that the cumulative effect of all the adjustments is significant. Despite the rebound in utilization from 1992, the net effect of utilization summed through 1999 is about neutral for the level of productivity. Hence, though the cyclical factors raised utilization early in the period, our estimates imply that the acceleration in productivity – especially late in the decade – is not due to cyclical factors.

The cumulative effect of the adjustment-cost correction over the decade is significant. Note that this adjustment will be nonzero in the long run because the estimates underlying our calibration find costs of adjustment even in steady state. An implication of this analysis is that traditional estimates of the growth in technology based on the Solow residual understate the growth in technology because they fail to add the growth in unmeasured output (internal installation services) to the residual. As shown in Table 3, we estimate that this correction to the Solow residual is 0.3 percent per year over the entire expansion from 1987 to 1999. It is even higher in the second half of the 1990s when investment continued to boom as the expansion progressed.

Though less volatile than the utilization correction, the adjustment-cost correction does affect the timing of the estimates of growth in technology. In the 1990 to 1992 period, low investment provided a boost to output, but high investment has been a steady subtraction since 1992. Hence, adjustment affects the timing of the acceleration in technology. It is easiest to see this in the right hand panels of Figure 4, which plots the Solow residual and the estimate of technological change. Again, the top panel shows growth rates while the bottom panel cumulates them. The factor adjustment correction adds to output in 1990 through 1992 and subtracts thereafter, so making this correction tends to accelerate estimated technology beginning in 1992. Hence, the main result of accounting for adjustment is the finding that the acceleration in productivity began earlier in the 1990s than is implied by other estimates.

Table 2 shows the annual growth rates of the estimates of technology over various periods. Taking into account the corrections for utilization, adjustment, and reallocation, there is a marked acceleration in technology at the end of 1990s. During the period 1995-1999, we estimate that technology grew at an annual rate of 3.1 percent. This rate is over a percentage point faster than the 2.0 percent rate over the 1987 to 1998 period as a whole. It also represents an acceleration compared with the longer, Jorgenson-Stiroh sample shown in the Appendix Table 1. Bearing in mind that one needs to add on the order of one percentage point to the Jorgenson-Stiroh rates to make them comparable to the BEA data owing to the quality adjustments, the last few years' growth has been substantially faster than the overall period. Moreover, the last few years roughly equal the strong performance of technological change during the 1960 to 1973 period.

Let's summarize our main findings concerning the aggregate:

1. Cyclical utilization accounts for some of the behavior of measured productivity during the early part of the 1990s expansion, but does not explain the acceleration in productivity at the end of the decade.
2. Adjustment costs from the onset of the boom initially obscured the acceleration in technology. In steady state, they add to correctly-measured technology relative to the Solow residual.
3. Reallocation affects the baseline growth in technology, but has not contributed substantially to the increase in its growth rate.
4. There is evidence of a substantial increase in the pace of technological change in the later half of the 1990s.

We conclude this section by noting that the main findings appear quite robust to datasets and parameter estimates. In terms of data, all datasets that we know of show an acceleration in productivity; different datasets (such as the BLS multifactor dataset) may show different magnitudes for the acceleration in productivity (and – after applying our corrections for utilization, adjustment, and reallocation – technology) but those differences do not affect the qualitative conclusion that an acceleration took place. In terms of parameters, consider, for example, the coefficients on hours growth. Reducing the coefficient on hours would reduce the potential impact of utilization – but that would simply strengthen the conclusion that utilization cannot explain the acceleration in productivity. Raising the coefficient on hours would have its largest effect in the 1992-94 period, and hence, would once again strengthen the conclusion that utilization does not explain the acceleration in productivity.

4.2. *Sectoral Results*

The remaining lines of Table 2 report the results for the durable manufacturing, non-durable manufacturing, and non-manufacturing. To summarize the results,

1. Durable manufacturing has the fastest rate of growth in technology and the largest acceleration. In the last half of the 1990s, technology in durable manufacturing increased over 6 percent per year.
2. Non-durable manufacturing shows very slow growth in technology recently.
3. Non-manufacturing also shows acceleration in technology relative to its recent performance.

Figures 5, 6, and 7 and Table 3 show the detail underlying these calculations. In durable manufacturing, there has been a substantial acceleration of output and inputs. This acceleration, of course, corresponds to the boom in investment discussed in the introduction. The utilization adjustment in durable manufacturing is very volatile, but has had little cumulative effect over the 1990s. The point estimate of returns to scale in durables is slightly greater than one (Table 1) and input growth has been strong, so the reallocation/scale effects are positive over the period. The adjustment-cost correction also plays a small role, with the pattern being similar to the aggregate. Its cumulative magnitude durable manufacturing is the same as in the aggregate private sector (see Table 3A and 3B), but the growth in technology is so large in aggregate manufacturing that the Solow residual and technology in Figure 5 move strongly together.

The data on non-durable manufacturing during the 1995-1999 period are somewhat surprising. The BEA's data show an acceleration in intermediate inputs. It is enough so that non-durable manufacturing value added actually falls despite healthy increases in gross output (see Figure 6 and Table 3C). This feature of the data warrants further investigation. It does

imply negative growth in the Solow residual. Figure 6 shows that the reallocation/scale correction is substantial. Given that the point estimates of returns to scale are less than one in non-durable manufacturing (Table 1), the growth in input leads to some negative reallocation correction. Even more importantly, the rising materials-output ratio in the BEA data leads to a substantially negative materials-reallocation term. This correction increases our estimate of technology relative to the Solow residual (see right panels of Figure 6 and Table 3C). Hence, technology is growing modestly. In contrast, the Solow residual is negative.

Non-manufacturing, like durable manufacturing, shows strong growth in output, input and the Solow residual throughout the 1990s in the left panels of Figure 7. The corrections on non-manufacturing are not as volatile as in manufacturing, mainly owing to the much lower volatility of utilization (middle panels of Figure 7). The adjustment correction has an important impact in the early part of the decade with the same pattern as in the aggregate data.

Non-manufacturing contributes a substantial fraction of the acceleration in technology in the second half of the 1990s. Growth in private sector technology increased 2.9 percentage points, from 1.2 percent to 3.1 percent. Growth in non-manufacturing technology increased by 1.6 percentage points, from 0.9 percent to 2.7 percent. Because non-manufacturing constitutes most of GDP, the non-manufacturing acceleration in technology accounts for most of the aggregate acceleration, despite the greater acceleration in durable manufacturing. Moreover, the acceleration in durable manufacturing was offset somewhat by the deceleration in non-durable manufacturing. In particular, non-manufacturing makes up about 80 percent of private output, so it accounted for 1.3 percentage points of the 2.9 percent point acceleration in private technology. Hence, while non-manufacturing contributes less than its share to the acceleration in technology, its contribution is nonetheless substantial.

The finding that non-manufacturing is so important in the increase in technology is important. It shows that the improvement in the pace of technological progress was quite widespread in the 1990s. In particular, it was not merely concentrated in the production of computers and other information technology goods, as claimed by Gordon (2000). Our results, and those of Baily and Lawrence (2001), Nordhaus (2000), Oliner and Sichel (2000), and Stiroh (2001), suggest by contrast that the acceleration in technology was widespread.

5. Discussion

Our results show a marked increase in the rate of technological change since 1995. These results complement Oliner and Sichel (2000), who also find an increase in the rate of technological change. Oliner and Sichel's work carefully focuses on the role of information technology investment. They do not, however, make the corrections to the Solow residual necessary to transform it into a reliable estimate of technological change. High utilization might account for their results. On the other hand, the high rate of investment that underlies the technology boom could be an offsetting factor pulling down the Solow residual relative to technology.

Our finding of a technology acceleration in durable manufacturing seems relatively uncontroversial – even New Economy skeptics such as Gordon (2000) agree that in computers, telecommunications, and other areas of durable manufacturing, technology has been improving at a rapid rate. Outside of durable manufacturing, however, our results contrast sharply with those of Gordon, who argues that the acceleration in labor productivity largely reflects cyclical factors. Gordon uses the quarterly labor-productivity data for the private business and durable manufacturing sectors to “back out” a series on labor productivity outside of durable manufacturing, which he then attempts to decompose econometrically into trend versus cycle.

Because of data limitations, Gordon's mapping is not necessarily precise – output in durable manufacturing is a gross output measure, whereas private business output is a measure of final sales (or value added). Changes in the materials-output ratio will, for example, change the mapping. Our industry data and the form of our utilization correction are also important differences.

Our results provide added support for the view that recent increase in the pace of productivity does correspond to an increase in technology: The corrections for utilization and adjustment have subtle effects on the timing of growth in the 1990s, but do not explain the notable acceleration in the latter half of the 1990s.

Is the increase in the growth of technology sustainable? Our results do not provide evidence one way or the other. Nothing in our technique of calculating technical change can detect momentum. Our framework does, however, provide one forecast that is favorable for the level of technology. Suppose the investment boom experienced in the 1990s is reversed. Our framework suggests, all other things equal, such a reversal will provide a once-and-for-all boost to the estimated level of productive capacity. During the boom, resources are being diverted to adjustment. If this pace of adjustment slows, those resources will be available permanently for production. While the effect on the growth rate is temporary, the effect on the level of output is permanent.

Appendix A: Derivation of the Utilization Proxy²¹

Following Basu and Kimball (1997), we model the firm as facing adjustment costs in both investment and hiring, so that both the amount of capital (number of machines and buildings), K , and employment (number of workers), N , are quasi-fixed. We believe that quasi-fixity is necessary for a meaningful model of variable factor utilization. Higher utilization must be more costly to the firm; otherwise factors would always be fully utilized. If there were no cost to increasing the rate of investment or hiring, firms would always keep utilization at its minimum level and vary inputs using only the extensive margin, hiring and firing workers and capital costlessly. Only if it is costly to adjust along the extensive margin is it sensible to adjust along the intensive margin, and pay the costs of higher utilization.

We assume that the number of hours per week for each worker, H , can vary freely, with no adjustment cost. In addition, both capital and labor have freely variable utilization rates. For both capital and labor, the benefit of higher utilization is its multiplication of effective inputs. We assume the major cost of increasing capital utilization, S , is that firms may have to pay a shift premium (a higher base wage) to compensate employees for working at night, or at other undesirable times.²² Variations in the workday of capital (i.e., the number of shifts) are perhaps the most plausible reason for variations in utilization. The variable-shifts model has had considerable empirical success in manufacturing data, where, for a short period of time, one can observe the number of shifts directly. We take S to be a continuous variable for simplicity. The cost of higher labor utilization, E , is a higher disutility on the part of workers that must be compensated with a higher wage. We allow for the possibility that high-frequency fluctuations

²¹ The material in this appendix is taken from Basu, Fernald, and Kimball (1999).

in this wage might be unobserved, as could be the case if wage payments are governed by an implicit contract in a long-term relationship.

We consider the following dynamic problem, in which an industry's representative firm minimizes the present value of expected costs:

$$\text{Min}_{S,E,H,M,I,R} E_t \sum_{s=t}^{\infty} \left[\prod_{j=t}^s (1+r_j)^{-1} \right] [WNG(H,E)V(S) + P_M M] \quad (\text{A1})$$

subject to

$$\bar{Y} = F(SK, EHN, M, Z)(1 - \Phi(I/K))(1 - \Psi(R/N)) \quad (\text{A2})$$

$$K_{t+1} = I_t + (1 - \mathbf{d})K_t \quad (\text{A3})$$

$$N_{t+1} = N_t + R_t \quad (\text{A4})$$

I is gross investment, and R is hiring net of separations. $WG(H, E)V(S)$ is total compensation per worker (which may take the form of an implicit contract, and hence not be observed period-by-period). W is the base wage; the function G specifies how the hourly wage depends on effort, E , and the length of the workday, H ; and $V(S)$ is the shift premium. P_M is the price of materials. \mathbf{d} is the rate of depreciation. We omit time subscripts where this practice does not cause confusion.

We assume that Ψ and Φ are convex, and make the appropriate technical assumptions on G in the spirit of convexity and normality.²³

²² Basu and Kimball (1997) extend this model to allow utilization to affect the rate at which capital depreciates.

²³ The conditions on G are easiest to state in terms of a function \mathbf{Q} defined by $\ln G(H,E) = \mathbf{Q}(\ln H, \ln E)$. Convex \mathbf{Q} guarantees a global optimum; assuming $\mathbf{Q}_{11} > \mathbf{Q}_{12}$ and

There are six intra-temporal first-order conditions and two Euler equations, for the state variables K and N . To conserve space, we analyze only the optimization conditions that affect our derivation. Let I be the multiplier on constraint (A2); I has the interpretation of marginal cost. Numerical subscripts denote derivatives of the production function F with respect to its first, second and third arguments, and literal subscripts denote derivatives of the labor cost function G . The conditions that we require are those for optimization with respect to choices of S , H , and E . These are:

$$S: \quad I F_1 K (1 - \Phi)(1 - \Psi) = w N G(H, E) V'(S) \quad (A5)$$

$$H: \quad I F_2 E N (1 - \Phi)(1 - \Psi) = w N G_H(H, E) V(S) \quad (A6)$$

$$E: \quad I F_2 H N (1 - \Phi)(1 - \Psi) = w N G_E(H, E) V(S) \quad (A7)$$

Note that the firm's uncertainty about future variables does not affect our derivations, which rely only on (some of) the *intra*-temporal equations for optimization. Uncertainty affects the evolution of the state variables (as the Euler equations would show) but not the minimization of variable cost at a point in time, *conditional* on the levels of the state variables. Our utilization proxy depends only on this static variable-cost minimization.

Equations (A6) and (A7) can be combined into an equation implicitly relating E and H :

$$\frac{H G_H(H, E)}{G(H, E)} = \frac{E G_E(H, E)}{G(H, E)}. \quad (A9)$$

The elasticity of labor costs with respect to H and E must be equal, because on the benefit side the elasticities of effective labor input with respect to H and E are equal. Given the assumptions on G , (A9) implies a unique, upward-sloping E - H expansion path, so that we can write

$Q_{22} > Q_{12}$ ensures that optimal H and E move together.

$$E = E(H), \quad E'(H) > 0. \quad (\text{A10})$$

Equation (A10) expresses unobservable intensity of labor utilization E as a function of the observed number of hours per worker H . We define $\mathbf{z} \equiv H^* E'(H^*) / E(H^*)$ as the elasticity of effort with respect to hours. Log-linearizing, we can write the growth rate of effective labor input as:

$$d \ln(EHN) = dn + dh + de = dn + (1 + \mathbf{z}) dh. \quad (\text{A11})$$

To find a proxy for capital utilization, we combine (A5) and (A6). Rearranging, we find:

$$\frac{F_1 SK/F}{F_2 EHN/F} = \left[\frac{G(H, E)}{HG_H(H, E)} \right] \left[\frac{SV'(S)}{V(S)} \right] \quad (\text{A12})$$

The left-hand side is a ratio of output elasticities, which (as in Hall, 1990) one can show are proportional to factor cost shares. We denote these cost shares by \mathbf{a}_K and \mathbf{a}_L . Define $g(H)$ as the elasticity of cost with respect to hours, and $v(S)$ as the ratio of the marginal shift premium to the average shift premium:

$$g(H) = \frac{HG_H(H, E(H))}{G(H, E(H))} \quad (\text{A13})$$

$$v(S) = \frac{SV'(S)}{V(S)}. \quad (\text{A14})$$

With these definitions, we can write equation (A12) as:

$$v(S) = \frac{\mathbf{a}_K}{\mathbf{a}_L} g(H). \quad (\text{A15})$$

The labor cost elasticity with respect to hours given by the function $g(H)$ is positive and increasing by the assumptions we have made on $G(H, E)$. The labor cost elasticity with respect to capital utilization given by the function $v(S)$ is positive as long as there is a positive shift

premium. We assume that the shift premium increases rapidly enough with S to make the elasticity increasing in S . We also assume that $\mathbf{a}_K/\mathbf{a}_L$ is constant, which requires that F be a generalized Cobb-Douglas in K and L .²⁴

Under this assumption, the log-linearization of (A15) is simply

$$ds = \frac{\mathbf{h}}{\mathbf{n}} dh. \quad (\text{A16})$$

where \mathbf{h} is the elasticity of g with respect to H and \mathbf{n} is the elasticity of v with respect to S . The parameter \mathbf{h} indicates the rate at which the elasticity of labor costs with respect to hours increases. The parameter \mathbf{n} indicates the rate at which the elasticity of labor costs with respect to capital utilization increases.

²⁴ Thus, we assume $Y = Z\Gamma((SK)^{a_K} (EHN)^{a_L}, M)$, where Γ is a monotonically increasing function. Basu and Kimball (1997) argue that estimating the general case where the ratio of the elasticities is a function of all four input quantities would demand too much of the data and the instruments.

Appendix B: Data Sources

Overview

We use three datasets on industry output and inputs.

The first is industry data from Dale Jorgenson and Kevin Stiroh, which runs from 1959 through 1996.²⁵ Unfortunately, much of the recent interest in productivity developments focuses on events after 1996, so we augment the Jorgenson-Stiroh dataset with other data.

The second dataset (which we use mainly for comparison) is from the Bureau of Labor Statistics, which provides measures of input, output, and TFP from 1948 to 1999 for the private business economy, for aggregate manufacturing, and for two-digit industries within manufacturing. Unfortunately, there is no industry coverage outside of manufacturing. This lack of coverage is a substantial shortcoming, since a major question is how pervasive the productivity acceleration is outside of durable manufacturing.²⁶

The third dataset, which we constructed, attempts to overcome some of the shortcomings of these first two datasets, and forms the core of the results we present. This dataset starts in 1987 and runs through 1999, and covers the entire non-farm, non-mining business economy at (roughly) the same level of disaggregation as the Jorgenson-Stiroh dataset. Our dataset has the advantages of being recent and comprehensive. It does not, however, incorporate adjustments for capital and labor composition. Our dataset is also relatively short (because of the availability of consistent measures of gross output and intermediate inputs from the Bureau of Economic Analysis.)

²⁵ These data are available (as of November 6, 2000) at Dale Jorgenson's web site, <http://www.economics.harvard.edu/faculty/jorgenson/data.html>.

Our basic strategy is to use the Jorgenson-Stiroh data to estimate all parameters. We then apply these estimated parameters to our dataset, to construct a parallel set of estimates of productivity and technology growth from 1987 to 1998. Appendix Table 1 provides results analogous to Table 2 for these data.

Details

The Jorgenson-Stiroh dataset is disaggregated at roughly the two-digit level within manufacturing and the one-digit level outside manufacturing.²⁷ For each of the 29 non-farm, non-mining private industries in our sample, the data include real measures of gross output Y , capital input K , labor input L , and intermediate inputs M . For each variable, Jorgenson and Stiroh also provide a corresponding price; by construction, nominal factor payments are equal to the nominal value of output (excluding taxes).

Jorgenson tries to ensure that the data are as consistent as possible with production theory. Notably, inputs are adjusted for composition, or “quality.” Some workers, for example, are paid more than others, which under standard neoclassical assumptions should be associated with differences in marginal products. Similarly, computers depreciate quickly, and rapidly lose their value (largely because technological progress has led to rapid declines in prices), while structures depreciate slowly and have price increases; thus, computers must have a higher

²⁶ Data through 1999 were released only in April 2001, too late to be used extensively in this paper.

²⁷ Jorgenson differentiates between motor vehicles (SIC 371) and other transportation equipment (SIC 372-379). Otherwise, the manufacturing industries match the two-digit SIC classifications. Jorgenson combines wholesale and retail trade into a single “trade” sector. In our dataset for the more recent period, we combine electric and gas utilities into a single ‘utilities’ industry, because of data limitations. For this reason, we compute Tornquist indices of output and inputs for Jorgenson’s two utilities industries before estimating parameters.

expected service flow than structures to ensure that same expected return. The quality corrections in the dataset attempt to take account of these differences in marginal products.

Our theory of utilization suggests that utilization should be proportional to hours per worker. We use data from the BLS establishment survey to measure hours per worker of production and non-supervisory workers (the vast majority of the workforce). In most cases, the data from the establishment survey provides a reasonably consistent time series by industry since 1947 for most two-digit manufacturing industries, and since 1964 for most other industries. For electronic and electrical equipment (SIC 36) and instruments (SIC 38), the hour-per-worker series begins only in 1988. That is, of course, fine for our main focus on the 1990s; but for other purposes, we would like to be able to create a series that goes back farther. We assume that before 1988, hours per worker grows at the same rate as in total durable manufacturing. For services, the data begin in 1989; we assume hours per worker is flat before that date (given that they are relatively flat after that date). We downloaded all data from Haver Analytics at a monthly frequency, seasonally adjusted (except for finance-insurance-and-real-estate, where we used the NSA data, because the SA data are available only since 1995).

To estimate the trend, we applied the Hodrick-Prescott filter to the log of the level of average weekly hours using monthly data from 1964 to 1999. We converted to an annual frequency by averaging the monthly log-levels; and then took first differences to calculate dh_i . We estimated a different trend for each industry. We set the Hodrick-Prescott coefficient to 10,000,000. Using lower values put too much of the high-frequency variation into the trend so that, by construction, utilization could not vary much. The utilization measure, dh_i , is the change in the deviation of the log of average weekly hours from their industry-specific Hodrick-Prescott trend.

Using other data sources, such as the Current Population Survey (downloaded from the BLS web site) or data from Dale Jorgenson on hours per employee (non-production as well as production), does not appear to substantively affect our results. For example, one alternative series we looked at was the CPS series for hours per worker of full-time (more than 35 hours/week) employees. Conceptually, hours per worker of full-time workers seems preferable, but unfortunately the series only begins in 1979. Because of the greater time period of coverage, and because results were largely unaffected, we relied primarily on the establishment-survey measure.

For our constructed industry dataset from 1987 to 1999, we use a variety of data sources.

Gross output, intermediate inputs, and nominal factor shares (which, again, sum to one by construction) are from the Bureau of Economic Analysis, Gross Product by Industry dataset.

These data are available through 1999, and are available at

<http://www.bea.doc.gov/bea/dn2/gpo4799.exe> (Downloaded January 15, 2001). Several issues arise with these data:

- A consistent time series is only available from 1987 onwards; we have not attempted to bridge the pre- and post-1987 data (which requires dealing with some differences in industry definitions) to create a longer time series.
- Electric and gas utilities are combined in the BEA data, so we created a corresponding Divisia aggregate for Jorgenson's data.
- Wholesale and retail trade are distinguished in the BEA data (and in most of the other data which follow), which we aggregate to get a composite trade sector to match Jorgenson's data.

- We remove owner-occupied housing from the finance, insurance, and real estate industry, since our measures correspond to private business output.

Industry investment and net capital stock data are available from the BEA fixed-asset dataset. These data are discussed in the September 2000 *Survey of Current Business*, and are available from the BEA web site at <http://www.bea.doc.gov/bea/dn2/facd.htm> (downloaded March 2001). We take the capital stock at the *beginning* of the year as our measure of capital input during the year. Note the following:

- Conceptually, these data do not make the composition corrections made by Jorgenson or in the BLS multifactor productivity dataset. That is, they are capital *stock* measures (a wealth measure), not capital *input* measures (a productivity measure).²⁸
- Comparing the growth rate between 1995-99 to the rate between 1990 and 1995, the capital stock accelerated by 1.3 percent per year (from 1.5 percent in the first half of the 1990s to 2.8 percent in the second half of the 1990s). By contrast, in the BLS MFP dataset, capital input accelerates by 2.6 percent per year, from 2.7 percent in the first half to 5.3 percent in the second half. Hence, omitted capital-composition effects can account for perhaps 0.4 percentage points of the technology acceleration we find $((2.6-1.3)*.33)$.
- As with the output data, we remove owner-occupied housing from the stock of capital in the finance-insurance-and-real-estate industry as well as from the economy-wide aggregate. (In 1999, owner-occupied housing accounted for \$7.3 billion of the \$19.8 billion in tangible private capital).

²⁸ Note: The BEA methodology is discussed in Department of Commerce (1999). The methodology section from that reference is available at <http://www.bea.doc.gov/bea/ARTICLES/NATIONAL/NIPAREL/Meth/wlth2594.pdf> (downloaded November 6, 2000).

Annual industry data on total employment, and on hours per worker of production and non-supervisory workers, are available from the BLS. We estimate industry hours as the product of these two series, and take that as our measure of industry labor input. We do not control separately for the hours per worker of non-production and supervisory workers. (This probably has a relatively small effect, since the majority of workers are, in fact, production or non-supervisory, and the hours of non-production workers are likely to be correlated with those of production workers.) Note the following:

- We do not make any adjustments for labor composition at this point. Preliminary estimates by Dan Aaronson and Dan Sullivan of the Federal Reserve Bank of Chicago suggest that labor quality grew 0.6 percent per year from 1988 to 1995, but only 0.3 percent per year from 1995 to 2000.²⁹
- According to BLS estimates in the April 2001, labor quality improved contributed 0.6 percent from 1990-95, and by 0.4 percent 1995-99.³⁰
- Hence, by omitting a correction for labor quality we probably *understate* the acceleration in technology in the second half of the 1990s.

In all cases where we need to aggregate data across industries (for example, when we create indices for durable manufacturing), we use Törnqvist indices, which provide a close approximation to the chain-linked indices used in the national accounts (See Whelan (2000b) for a discussion).

²⁹ Forthcoming in the Federal Reserve Bank of Chicago's *Economic Perspectives* in 2001.

³⁰ From mfp2ddot.txt, obtained from <ftp://146.142.4.23/pub/special.requests/opt/mp/mfptable.zip>, downloaded May 18, 2001.

Instruments

We use the following as demand-side instruments:

1. The price of oil, measured as the PPI for crude petroleum deflated by the GDP deflator.
2. Ramey-Shapiro (1998) military buildup dates. These dates are a dummy variable, equal to one in 1950, 1965, and 1980, and zero otherwise. It measures upward shifts in government defense spending.
3. Money shocks from an identified VAR, using an updated version of Christiano, Eichenbaum, and Evans (1999) and Christiano, Eichenbaum, and Evans (1996). We sum the four quarterly policy shocks in year $t-1$ as instruments for year t . These shocks are available from 1960 through 1999.

References

- Abel, A.B., (1999). The Effects of a Baby Boom on Stock Prices and Capital Accumulation in the Presence of Social Security. Unpublished, University of Pennsylvania.
- Baily, M.N. and Lawrence, R.Z., (2001). Do We Have a New E-conomy? *American Economic Review Papers and Proceedings*, **91**: 308-312. May.
- Basu, S., (1996). Procyclical Productivity: Increasing Returns or Cyclical Utilization? *Quarterly Journal of Economics*, **111**: 719-51.
- Basu, S. and Fernald, J.G., (1997). Aggregate Productivity and Aggregate Technology. International Finance Discussion Paper 593, Board of Governors of the Federal Reserve System.
- Basu, S. and Fernald, J.G., (2001). Why is Productivity Procyclical? Why Do We Care? *New Directions in Productivity Research*, eds. C. Hulten, E. Dean, and M. Harper. Chicago: University of Chicago Press.
- Basu, S., Fernald, J.G., and Kimball, M.S., (1999). Are Technology Improvements Contractionary? Unpublished, University of Michigan.
- Basu, S. and Kimball, M.S., (1997). Cyclical Productivity with Unobserved Input Variation. NBER Working Paper 5915.
- Blank, R.M. and Shapiro, M.D., (2001). Labor and the Sustainability of Output and Productivity Growth. Prepared for *Sustainable Employment*, eds. A.B. Krueger and R.M. Solow. New York: Russell-Sage Foundation/Century Foundation, in process.
- Burnside, C., Eichenbaum, M.S., and Rebelo, S.T., (1995). Capital Utilization and Returns to Scale. *NBER Macroeconomics Annual* eds. B.S. Bernanke, and J.J. Rotemberg, **10**: 67-110.
- Carlton, D.W., (1983). Equilibrium Fluctuations When Price and Delivery Lag Clear Markets. *Bell Journal of Economics*, **14**: 562-572.
- Christiano, L.J., Eichenbaum, M.S., and Evans, C.E., (1996). The Effects of Monetary Policy Shocks: Some Evidence from the Flow of Funds. *Review of Economics and Statistics*, **78**: 16-34.
- Christiano, L.J., Eichenbaum, M.S., and Evans, C.E., (1999). Monetary Shocks: What Have We Learned and To What End? *Handbook of Macroeconomics*, eds. J. Taylor and M. Woodford. Amsterdam: Elsevier.
- Gordon, R.J., (2000). Does the “New Economy” Measure up to the Great Inventions of the Past?

- Journal of Economic Perspectives* **14**: 49-74.
- Greenwood, J., Hercowitz, Z., and Huffman, G.W., (1988). Investment, Capacity Utilization, and the Real Business Cycle. *American Economic Review*, **78**: 402-17.
- Hall, R.E., (1980). Employment Fluctuations and Wage Rigidity. *Brookings Papers on Economic Activity*, 91-123.
- Hall, R.E., (1990). Invariance Properties of Solow's Productivity Residual. *Growth/Productivity/Unemployment: Essays to Celebrate Bob Solow's Birthday*, Ed. P.A. Diamond, Cambridge, MA: MIT Press.
- Hall, R.E., (2001a). Struggling to Understand the Stock Market (Ely Lecture). *American Economic Review Proceedings*, **91**: in process.
- Hall, R.E., (2001b). The Stock Market and Capital Accumulation. *American Economic Review*, forthcoming.
- Jorgenson, D.W., Gollop, F. and Fraumeni, B., (1987). *Productivity and U.S. Economic Growth*. Cambridge: Harvard University Press.
- Jorgenson, D.W. and Stiroh, K.J., (2000). Raising the Speed Limit: U.S. Economic Growth in the Information Age. *Brookings Papers on Economic Activity*, 125-211.
- Kiley, M.T., (2001). Computers and Growth with Costs of Adjustment: Will the Future Look Like the Past? *Carnegie-Rochester Conference Series on Public Policy*, this issue.
- Nadiri, M.I. and Prucha, I.R., (2001). Dynamic Factor Demand Models and Productivity Analysis. *New Directions in Productivity Analysis*, eds. C. Hulten, E. Dean, and M. Harper. Chicago: University of Chicago Press.
- Nordhaus, W.D., (2000). Productivity and the New Economy. Cowles Foundation Discussion Paper No. 1284. New Haven, CT: Yale University.
- Oliner, S.D. and Sichel, D.E., (2000). The Resurgence of Growth in the Late 1990s: Is Information Technology the Story? *Journal of Economic Perspectives*, **14**: 3-22.
- Ramey, V.A. and Shapiro, M.D., (1998). Costly Capital Reallocation and the Effects of Government Spending. *Carnegie-Rochester Conference Series on Public Policy*, **48**: 145-94.
- Rotemberg, J.J. and Woodford, M., (1995). Dynamic General Equilibrium Models with Imperfectly Competitive Product Markets. *Frontiers of Business Cycle Research* ed. T.F. Cooley. Princeton: Princeton University Press.
- Shapiro, M.D., (1986a). The Dynamic Demand for Capital and Labor. *Quarterly Journal of*

- Economics*, **101**: 513-42.
- Shapiro, M.D., (1986b). Capital Accumulation and Capital Utilization: Theory and Evidence. *Journal of Applied Econometrics*, **1**: 211-34.
- Shapiro, M.D., (1993). Cyclical Productivity and the Workweek of Capital. *American Economic Review Papers and Proceedings* **83**: 229-33.
- Shapiro, M.D., (1996). Macroeconomic Implications of Variation in the Workweek of Capital. *Brookings Papers on Economics Activity*, 79-133.
- Shiller, R.J., (2000). *Irrational Exuberance*. Princeton: Princeton University Press.
- Sims, C.A., (1974). Output and Labor Input in Manufacturing, *Brookings Papers on Economic Activity*, 695-728.
- Solow, R.M., (1957). Technological Change and the Aggregate Production Function. *Review of Economics and Statistics*, **39**: 312-320.
- Solow, R.M., (1990). Reactions to Conference Papers. *Growth/Productivity/Unemployment: Essays to Celebrate Bob Solow's Birthday* ed. Peter Diamond. Cambridge, MA: MIT Press.
- Stiroh, K.J., (2001). Information Technology and the U.S. Productivity Revival: What Do the Industry Data Say? Federal Reserve Bank of New York Staff Reports, Number 115. January.
- Whelan, K., (2000a). Computers, Obsolescence, and Productivity. FEDS working paper 2000-6.
- Whelan, K., (2000b). A Guide to the Use of Chain Aggregated NIPA Data. FEDS working paper 2000-35.
- Department of Commerce, Bureau of Economic Analysis (1999). *Fixed Reproducible Tangible Wealth in the United States, 1925-94*. Washington, D.C: Government Printing Office.
- Economic Report of the President*, (1994). Washington, D.C: Government Printing Office

Table 1. Parameter Estimates

$$dy_i - \mathbf{f}_i di_i = \mathbf{g}_i^* dx_i + \mathbf{x}_i dh_i + dz_i$$

	Durable	Non-durable	
	Manufacturing	Manufacturing	Non-manufacturing
Returns to scale, \mathbf{g}_i^*	1.03	0.78	1.00
	(0.04)	(0.04)	(0.09)
Utilization, \mathbf{x}_i	0.74	1.21	1.33
	(0.17)	(0.21)	(0.55)

Note: Estimation by three-stage least squares pooling across industries within sectors. Data are from Jorgenson and Stiroh and are at an industry level from 1965 to 1996. Instruments are monetary policy shocks, growth in PPI for crude oil, and Ramey-Shapiro (1998) military buildup dates. Sectoral estimates include industry fixed effects (not reported). Standard errors are in parentheses.

Table 2. Estimated Growth in Technology

(Percent per Year)

	1987- 1999	1990- 1995	1995- 1999
Private Sector	2.0	1.2	3.1
Durable Manufacturing	4.6	3.4	6.7
Non-durable Manufacturing	1.2	1.4	0.9
Non-manufacturing	1.7	0.9	2.7

BEA data. Estimates of technological change based on theoretical framework of Section 2 and parameter estimates in Table 1. Conceptually, these estimates represent an appropriately weighted average of estimated technology growth in 29 sectors that span the private non-farm business economy, after controlling for utilization, capital-adjustment, and deviations from constant returns and perfect competition.

Table 3. Output, Input, and Productivity Growth and Corrections

(Percent per Year)

A. Private Sector

	1987-	1990-	1995-
	1999	1995	1999
Output	3.5	2.6	5.1
- Inputs	1.8	1.4	2.5
= Solow residual	1.6	1.2	2.5
- Utilization correction	0.0	0.2	0.1
- Adjustment correction	-0.3	-0.2	-0.6
- Reallocation and scale correction	-0.1	0.0	-0.1
= Technology	2.0	1.2	3.1

B. Durable Manufacturing

	1987-	1990-	1995-
	1999	1995	1999
Output	4.8	3.9	7.6
- Inputs	0.6	0.0	1.9
= Solow residual	4.2	3.9	5.7
- Utilization correction	-0.1	0.7	-0.6
- Adjustment correction	-0.4	-0.3	-0.6
- Reallocation and scale correction	0.1	0.1	0.2
= Technology	4.6	3.4	6.7

Table 3 (continued)

C. Non-durable manufacturing

	1987-	1990-	1995-
	1999	1995	1999
Output	0.7	2.0	-0.7
- Inputs	0.7	0.7	0.5
= Solow residual	0.0	1.3	-1.2
- Utilization correction	-0.1	0.0	0.0
- Adjustment correction	-0.3	-0.1	-0.4
- Reallocation and scale correction	-0.8	0.0	-1.6
= Technology	1.2	1.4	0.9

D. Non-manufacturing

	1987-	1990-	1995-
	1999	1995	1999
Output	3.3	2.4	4.7
- Inputs	2.0	1.5	2.6
= Solow residual	1.3	0.8	2.1
- Utilization correction	0.0	0.1	0.1
- Adjustment correction	-0.3	-0.2	-0.5
- Reallocation and scale correction	-0.1	0.0	-0.1
= Technology	1.7	0.9	2.7

Appendix Table 1. Estimated Growth in Technology: Jorgenson-Stiroh Data

(Percent per Year)

	1965- 1996	1965- 1973	1973- 1979	1979- 1990	1990- 1995	1995- 1996
Private Sector	0.6	1.5	0.3	0.2	-0.4	1.2
Durable Manufacturing	2.1	1.7	3.2	1.5	2.2	7.7
Non-durable Manufacturing	2.2	3.6	2.1	1.1	1.2	1.4
Non-manufacturing	2.2	1.1	-1.0	-0.2	-1.1	-0.2

Note: Estimates of technological change based on parameter estimates in Table 1. The Jorgenson-Stiroh data contains adjustments for the quality of inputs, so estimates of growth in technology based on those data should be smaller on average than from the BEA data.

Figure 1. Ratio of Nonresidential Fixed Investment to GDP (Nominal)

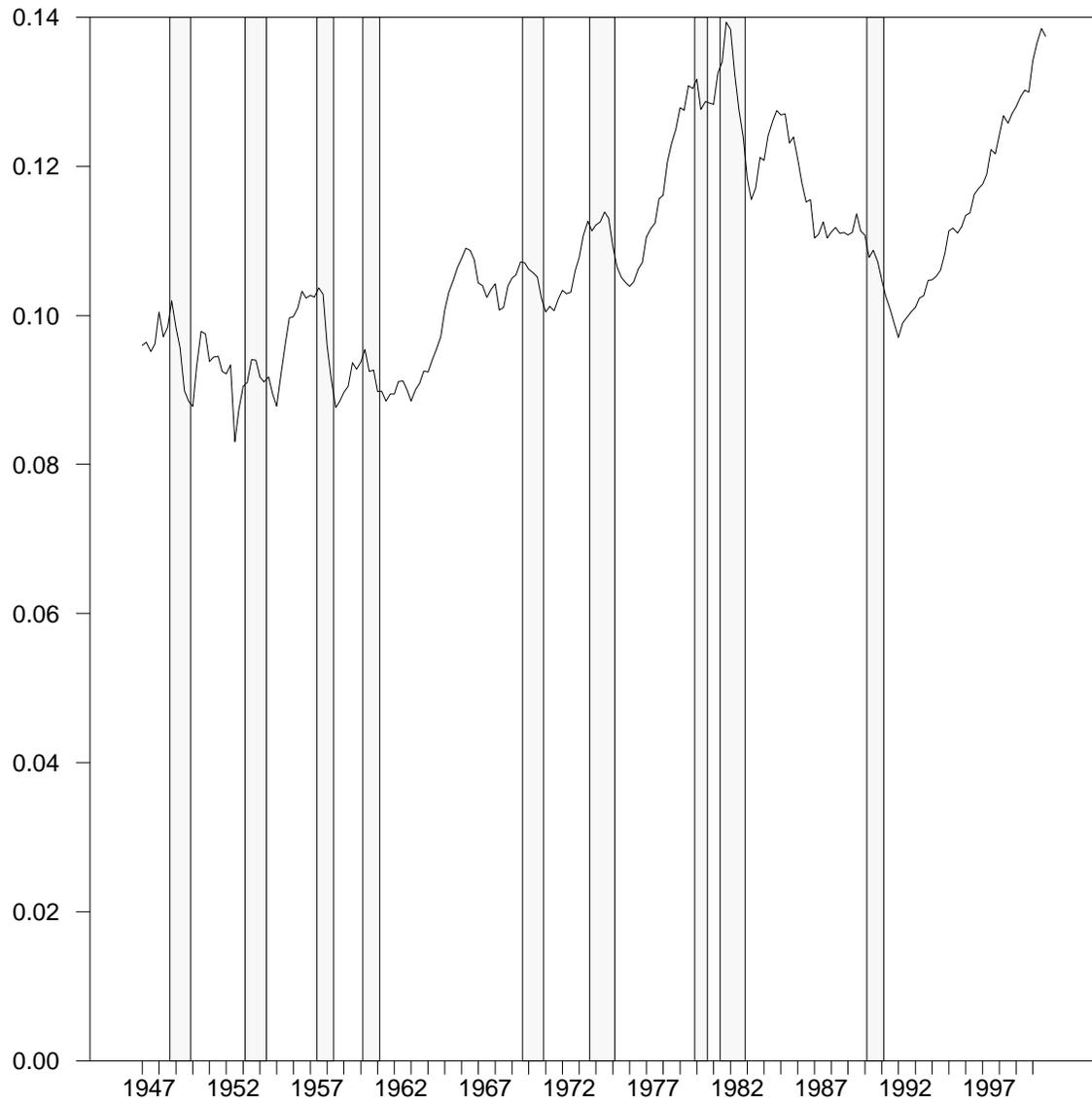


Figure 2. Ratio of Nonresidential Fixed Investment to GDP (Nominal)

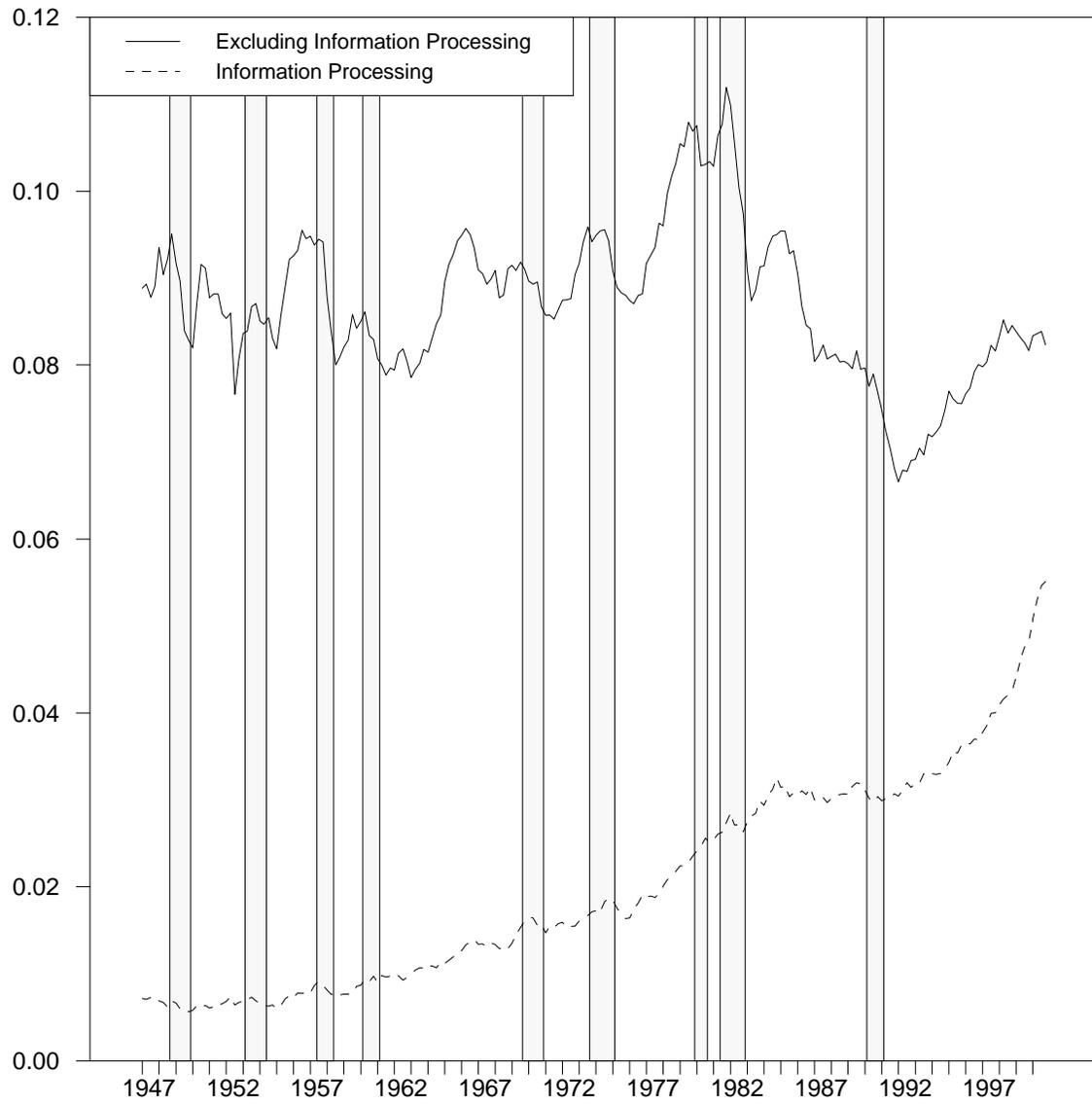


Figure 3. Average Weekly Hours of Production Workers

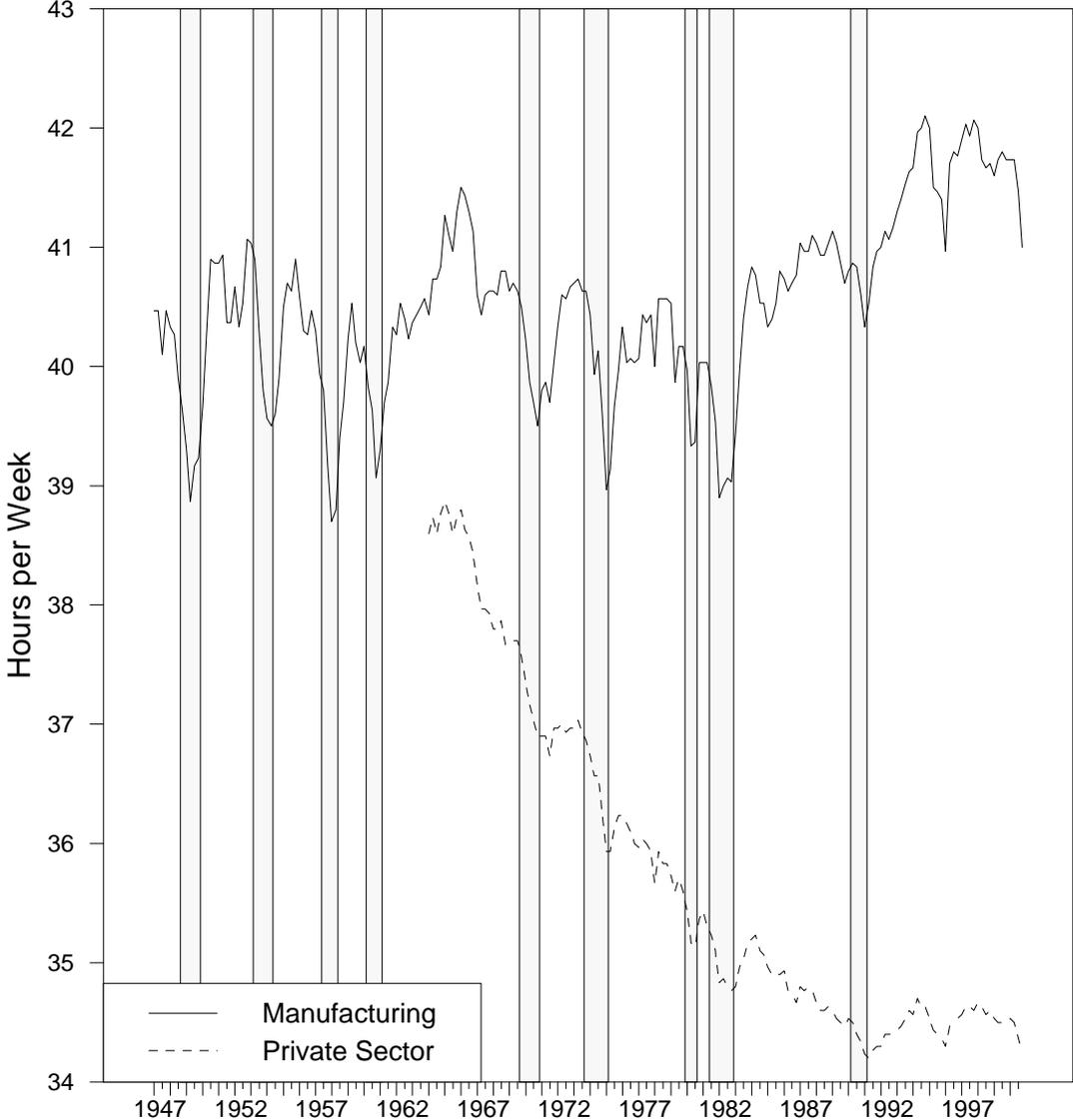


Figure 4. Private Sector

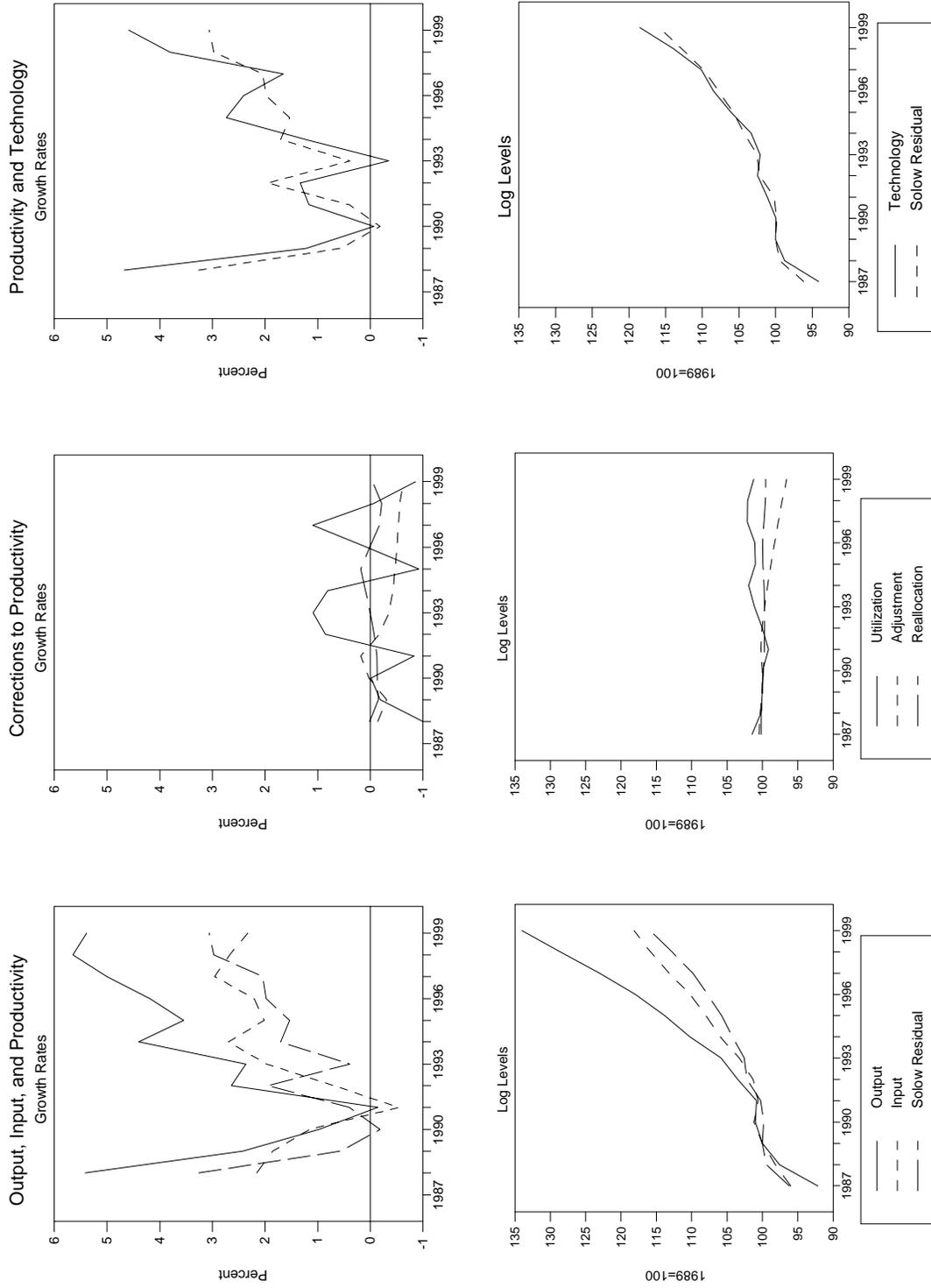


Figure 5. Durable Manufacturing

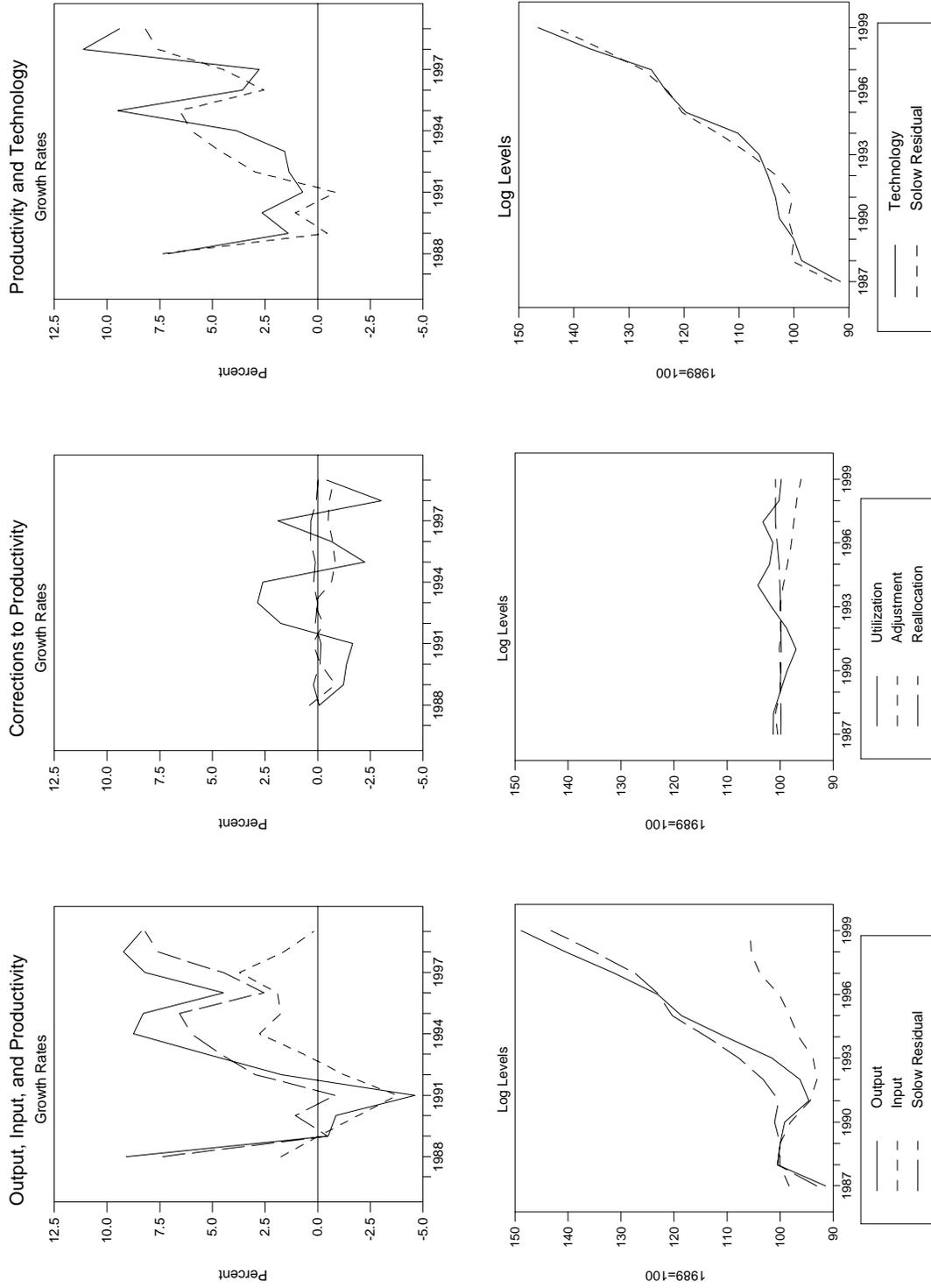


Figure 6. Non-durable Manufacturing

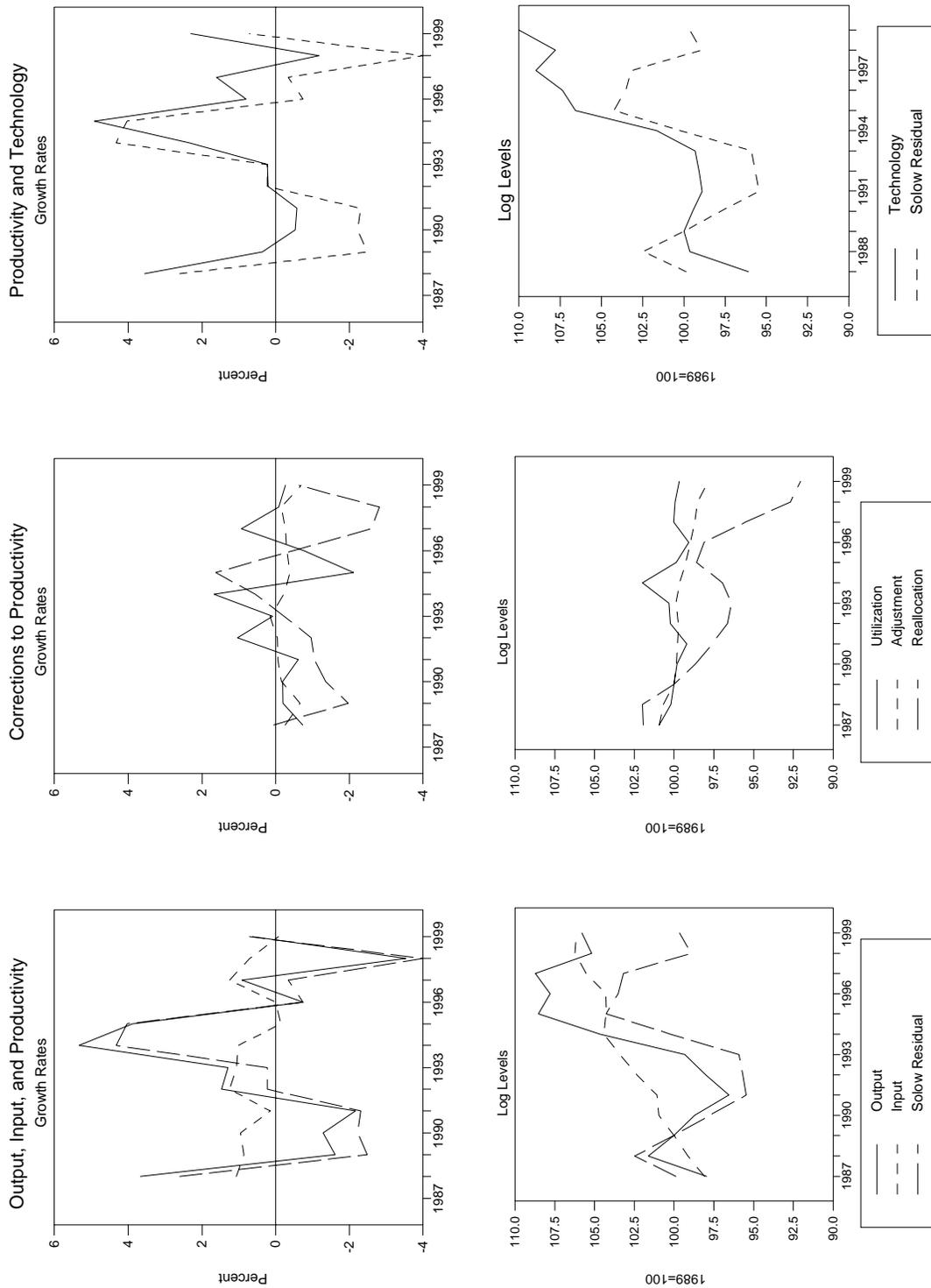


Figure 7. Non-manufacturing

