Federal Reserve Bank
of Minneapolis

Summer 1999

Quarterly Review

Federal Reserve Bank of Minneapolis

Quarterly Review Vol. 23, No. 3

ISSN 0271-5287

This publication primarily presents economic research aimed at improving policymaking by the Federal Reserve System and other governmental authorities.

The *Quarterly Review* is published by the Research Department of the Federal Reserve Bank of Minneapolis. Subscriptions are available free of charge.

Electronic files of *Quarterly Review* articles are available through the Minneapolis Fed's home page on the World Wide Web: http://www.minneapolisfed.org.

Comments and questions about the *Quarterly Review* may be sent to

Quarterly Review
Research Department
Federal Reserve Bank of Minneapolis
P.O. Box 291
Minneapolis, Minnesota 55480-0291
(Phone 612-204-6455 / Fax 612-204-5515).

Subscription requests may also be sent to the circulation assistant at elaine.reed@mpls.frb.org; editorial comments and questions, to the managing editor at ksr@res.mpls.frb.fed.us.

Federal Reserve Bank of Minneapolis
Quarterly Review Summer 1999

Aggregate Returns to Scale: Why Measurement Is Imprecise

Harold L. Cole*
Senior Economist
Research Department
Federal Reserve Bank of Minneapolis

Lee E. Ohanian*
Economist
Research Department
Federal Reserve Bank of Minneapolis
and Associate Professor of Economics
University of Minnesota

The value of aggregate *returns to scale*—the percentage change in output from a given percentage change in factor inputs—has important implications for the sources of shocks that lead to business cycle fluctuations. With constant or decreasing returns to scale, business cycle models driven largely by technology shocks are consistent with a number of business cycle facts, in particular, with procyclical labor productivity. In contrast, with constant or decreasing returns to scale, business cycle models driven primarily by monetary shocks are inconsistent with procyclical productivity. With constant or decreasing returns to scale, the marginal product of labor is diminishing; therefore, an increase in labor input brought about by a monetary shock alone drives down productivity. With increasing returns, however, monetary shocks can generate procyclical productivity in otherwise standard models. Moreover, if the value of returns to scale is sufficiently large, equilibria may not be unique, and as Benhabib and Farmer (1994) show, self-fulfilling beliefs, or *animal spirits,* alone can generate fluctuations that are difficult to distinguish from fluctuations in the standard real business cycle model driven by technology shocks. In fact, business cycle fluctuations in economies with a sufficiently large returns to scale value can be due to virtually any shock that moves factor inputs. Thus, our ability to evaluate the importance of the sources of business cycle fluctuations depends on the value of aggregate returns to scale.

While the value of returns to scale is important for evaluating the sources of business cycle shocks, measuring returns is difficult. First, there is an identification problem: model economies with any value of returns to scale are observationally equivalent if the unobserved stochastic process generating the shocks is unrestricted. Second, although researchers have come up with various ways of confronting the identification problem to measure returns to scale, the resulting estimates often cover a wide range of values, including significant decreasing and large increasing returns to scale. Moreover, the estimates often have large standard errors and corresponding wide confidence intervals. Consequently, firm conclusions about the value of returns to scale are hard to draw; thus, the importance of various sources of business cycle shocks are hard to evaluate.

In this article, we analyze the measurement of aggregate returns to scale. We first show why there is an identification problem and discuss what assumptions are required to solve that problem. We then conduct a simple analysis that sheds light on how precisely returns to scale can be measured. In this analysis, we compare the technology shocks inferred from aggregate production functions that are identical except for the value of returns to

*The authors thank Tom Holmes, Narayana Kocherlakota, Ed Prescott, and Warren Weber for helpful comments and Jenni Schoppers for expert editorial advice.

scale. If we can measure returns to scale precisely, then the technology shocks should be sensitive to changes in the value of returns to scale. With precise measurement, the technology shock inferred from the assumption of decreasing returns should be different from the technology shock inferred from the assumption of increasing returns. Alternatively, with imprecise measurement, the shocks should be insensitive to changes in the value of returns to scale.

We conduct our analysis for values ranging from significant decreasing returns up to substantial increasing returns. Our main finding is that the technology shocks inferred from this range of values are nearly identical. For this range of values of returns to scale, the correlation of the shocks is close to one and the shocks have the same serial correlation properties and similar variances. Unfortunately, these results have negative implications for how precisely we can measure aggregate returns using standard measures of inputs and output. The similarity of the series suggests that the likelihood functions researchers use to estimate returns to scale are insensitive to variation in this parameter and, consequently, that measurement of returns to scale will not be precise.

To conduct our analysis, we construct a model economy similar to models used in the literature, derive an observational equivalence result, and show how researchers have restricted the stochastic process for the shocks to identify returns to scale. We then show how restricting the shock process also implies, in principle, sharp restrictions on the covariance properties of the technology shocks. We go on to show that the covariance properties of the innovations to the technology shock are nearly the same for a wide range of returns to scale values.

The Model

In this section, we construct our basic model economy and characterize a set of functions that constitute an equilibrium.

Our basic model is similar to the one used by Benhabib and Farmer (1994). Our model has a measure 1 number of identical households. The households' preferences are given by

$$(1) \qquad E\left\{\sum_{t=0}^{\infty}\beta^t u(c_t, 1-l_t, d_t)\right\}$$

where β is the discount factor, u is the utility function, c_t is consumption of the single physical good produced in the economy, $1 - l_t$ is *leisure* (nonmarket time), and d_t is a *preference* (home production) shock.[1]

The household's budget constraint is given by

$$(2) \qquad y_t + (1-\delta)k_t = c_t + k_{t+1} + b_{t+1} - r_t b_t + \tau_t$$

where b_t and r_t denote the household's borrowing level and the gross interest rate, respectively, and τ_t denotes a lump-sum tax.

Per capita production of the household is given by

$$(3) \qquad y_t = \lambda_t F(k_t, l_t) Y_t^{\phi}$$

where k_t and l_t denote the household's levels of capital and labor and Y_t is aggregate per capita output.[2] Following other work in the literature, we assume that the production function $F(\cdot)$ is a linear homogeneous Cobb-Douglas function.[3] The term λ_t is the aggregate technology shock. The parameter ϕ determines the value of the externality and, consequently, aggregate returns to scale. The economy is defined as *neoclassical* if $\phi = 0$.

For generality, we allow for different sources of shocks that might lead to business cycle fluctuations—technology shocks, government spending shocks, preference shocks, and shocks to extraneous factors (*sunspots*). We define ε_t as a 4×1 vector of independent and identically distributed random variables, with $\varepsilon^t \equiv \{\varepsilon_s\}_{s=0}^t$. We let the period t realization of the technology shock, the government spending shock, the preference shock, and, where relevant, the sunspot variable be given by $\lambda_t(\varepsilon^t)$, $g_t(\varepsilon^t)$, $d_t(\varepsilon^t)$, and $v_t(\varepsilon^t)$.

The government's budget constraint is

$$(4) \qquad g_t = \tau_t.$$

The aggregate level of (per capita) output is given by

$$(5) \qquad Y_t = [\lambda_t F(K_t, L_t)]^{1/(1-\phi)} = [\lambda_t K_t^{\theta} L_t^{1-\theta}]^{1/(1-\theta)}$$

where K_t is the period t capital stock and L_t is the period t labor input.

---

[1]Because households are identical, there is no borrowing in equilibrium in this economy.

[2]The assumption that the aggregate externality depends on the level of per capita output, rather than aggregate output, is motivated by the observation that large countries do not seem to be systematically more productive than small countries.

[3]Our results do not depend on the Cobb-Douglas functional form; we only require that F be homogeneous.

An equilibrium for this economy consists of a set of functions which describe the household's policy $\{c_t, l_t, k_{t+1}\}(\varepsilon^t)$, the gross interest rate $\{r_t\}(\varepsilon^t)$, and aggregate per capita output $\{Y_t\}(\varepsilon^t)$ such that the following conditions are satisfied:

i. Households maximize with $\{c_t, l_t, k_{t+1}\}(\varepsilon^t)$ and $b_{t+1} = 0$, given $\{r_t\}(\varepsilon^t)$ and $\{\tilde{y}_t\}(\varepsilon^t)$.
ii. The equation of motion for aggregate output is consistent with households' production decisions:

$$(6) \quad y_t = [\lambda_t k_t^{\theta} k_t^{1-\theta}] Y_t^{\phi} = [\lambda_t F(K_t, L_t)]^{1/(1-\phi)}.$$

Observational Equivalence

Here we derive an observational equivalence result and show the implications of restricting the technology shock process.

Note that in constructing our model, we made no restriction on the stochastic process for the technology shock, λ_t. The following proposition, which is drawn from Cole and Ohanian 1996 and is similar to results in Kamihigashi 1996, shows that if this stochastic process is left unrestricted, an observational equivalence result has important implications for measuring returns to scale:

PROPOSITION. *Given ϕ and a corresponding equilibrium $\{c_t^{\phi}, l_t^{\phi}, k_{t+1}^{\phi}, r_t^{\phi}, y_t^{\phi}\}(\varepsilon^t)$, then for any $\hat{\phi}$, a stochastic process for the shocks can be constructed, $\hat{\lambda}_t(\varepsilon_t)$, such that the original equilibrium is an equilibrium for the economy with $\hat{\phi}$ and $\hat{\lambda}_t(\varepsilon_t)$.*

Proof. Define the new stochastic process:

$$(7) \quad \hat{\lambda}_t(\varepsilon^t) = \lambda_t(\varepsilon^t) Y_t(\varepsilon^t)^{(\phi-\hat{\phi})}.$$

We can easily verify that with this new stochastic process for the technology shock, $\hat{\lambda}_t$,

$$(8) \quad \hat{\lambda}_t(\varepsilon^t) Y_t(\varepsilon^t)^{\hat{\phi}} = \lambda_t(\varepsilon^t) Y_t(\varepsilon^t)^{\phi}.$$

This implies that the household's original policy functions still solve the household's problem and that condition (i) of a competitive equilibrium is satisfied. Given that the household's maximization condition (i) is satisfied, the aggregate consistency condition (ii) is satisfied as well. Q.E.D.

Empirically, this proposition indicates that observed time series of consumption, investment, output, capital, and labor input can be generated by this economy with any value of returns to scale. The nature of the observational equivalence is that the technology shock and the returns to scale value are both unrestricted and both play similar roles in the production function: given the inputs, K and L, the technology shock *multiplicatively* scales the Cobb-Douglas function of the inputs, while the returns to scale value *exponentially* raises that scaled function. If the shock process is left unrestricted, then a new technology shock can be constructed that is the product of the original shock and output raised to the appropriate power.

Thus, observed business cycle fluctuations may be due to relatively large technology shocks with constant returns to scale, as measured by Prescott (1986). However, fluctuations may also be due to small technology shocks that are amplified by increasing returns or by any shock that, with sufficiently large returns to scale, increases output and, consequently, productivity. These include monetary shocks, government spending shocks, preference shocks, or sunspots. Without restrictions on the shock process, the data shed no light on the relative importance of these various sources for business cycle fluctuations.

Restricting the Shock Process

In this section, we examine two approaches to restricting the shock process: restricting the order of the process and using instrumental variables.

The proposition in the preceding section suggests that returns to scale cannot be identified without restricting the shock process. To solve this problem, we need an identifying assumption that restricts the shock process. Ideally, identifying assumptions are derived from economic theory, because theory makes transparent how identification is achieved and requires no assumptions outside the theory used to construct the model economy. Unfortunately, there is no generally accepted theory of these unobserved technology shocks that restricts the shock process. Without identifying assumptions derived explicitly from theory, identifying assumptions are sometimes based on prior knowledge. This approach, however, is hard to implement because the technology shock is a latent variable.

If neither theory nor prior knowledge can be used to derive identifying assumptions, researchers must resort to using ad hoc assumptions to achieve identification. However, choosing restrictions that are not derived from economic theory or at least motivated by strong prior knowledge is not entirely satisfactory. Nevertheless, as the proposition shows, some restrictions on the stochastic process are required to break the observational equivalence.

The standard approach, known in the literature as the *instrumental variables approach,* consists of two restrictions. First, a particular stochastic process for the shock is specified. Second, an instrumental variable is chosen. This variable has two properties: it is assumed to be orthogonal to the *innovation* (the unpredictable component) to the technology shock, and it is correlated with the right side variables in the equation, in this case, the factor inputs. Before we examine the instrumental variables approach, we will first examine the implications of specifying a shock process.

Restricting the Order . . .

Here we consider the implications of restricting the order of the stochastic process. Assume that the stochastic process for the technology shock at any value of returns to scale other than the correct value is of order *s,* where *s* is the length of the history of the shocks. Assume also that at the correct value of returns to scale, the stochastic process theoretically will be of the order specified by the researcher. This is usually an order much lower than *s*—typically, a process of order 1 or 2 is chosen. Before we describe how this restriction is useful for understanding imprecise measurement, we will first illustrate how this restriction on the order of the stochastic process arises.

To understand this restriction, assume that the shocks follow a first-order log-linear Markov process. Now consider our model with $\phi = \phi^* \neq 0$. Define the shock vector $\mu_t \equiv [\lambda_t\ g_t\ d_t\ \nu_t]$. The conditional probability distribution of μ_t is given by $F(\mu_t;\mu_{t-1})$. A Markov equilibrium for this economy, conditional on ϕ^*, is given by a set of functions $\{c^{\phi^*},l^{\phi^*},k^{\phi}_{t+1}\}(s_t,k)$ and $\{r^{\phi^*},Y^{\phi^*}\}(s_t)$, where the state vector $s_t \equiv (\mu_t,K_t)$.

Now consider an alternative, neoclassical ($\hat{\phi}$=0) economy. Given a Markov equilibrium for the ϕ^* economy with the restricted shock process, we can't construct an equivalent Markov specification for the shock process for the neoclassical economy.

To see this, define $\hat{\mu}_t \equiv [\hat{\lambda}_t\ g_t\ d_t\ \nu_t]$. Define $G(\hat{\mu}_t,s_{t-1})$ as the probability distribution over $\hat{\mu}_t$. We can take

$$G^i(\hat{\mu}_t;s_{t-1}) = F^i(\mu_t;\mu_{t-1}) \tag{9}$$

for i = 2, 3. Next define

$$\lambda(\hat{\mu},s_{t-1}) = \hat{\lambda}_t Y^{\phi}(\mu,s_{t-1}). \tag{10}$$

We can't construct an equivalent Markov equilibrium because the shock process in the ϕ^* economy is a first-order Markov process, but the constructed shock process in the neoclassical economy depends on the level of output. Because output depends on labor and capital and because capital is a function of all past shocks, we can show that the constructed shock process will not be a first-order process, as was the original shock process, but rather will be a process of order *s,* where *s* is the length of the history of the shocks. The difference between the order of these shock processes is the key implication of this restriction.

. . . And Using Instrumental Variables

With the instrumental variables approach, returns to scale can be measured by choosing a value such that the covariance between the innovation to the technology shock and the instrument is zero. We use an instrument because theory implies that at least some of the inputs are not orthogonal to the innovation to the technology shock; consequently, some other variable must be chosen as an instrument. Burnside, Eichenbaum, and Rebelo (1995) discuss of some of the instruments that have been used.

To understand the instrumental variables estimation formally, consider estimating returns to scale in our model economy with an instrumental variable. To do this, we need to make an assumption for the stochastic process of the technology shock (λ_t). For simplicity, we assume that the process is exogenous and that it is a log-random walk with drift:

$$\ln(\lambda_t) = \ln(\lambda_{t-1}) + \mu + \varepsilon_t. \tag{11}$$

Using lowercase letters as natural logs of variables, we define x_t as a weighted sum of the logged inputs:

$$x_t \equiv [\theta k_t + (1-\theta)l_t]. \tag{12}$$

Given this specification and abstracting from a constant term, we have that the innovation to the log of the technology shock is given by

$$e_t = \Delta y_t - \psi \Delta x_t \tag{13}$$

where $\psi \equiv 1/(1-\phi)$ and $e_t = \psi\varepsilon_t$.

Estimating returns to scale ψ by ordinary least squares is not appropriate because standard business cycle models imply that the innovation and the inputs will be correlated. In particular, a positive technology shock will increase the marginal product of labor and lead households to increase labor input. (See, for example, Prescott 1986 and McGrattan 1994.) To see this, note that the estimate of ψ, $\hat{\psi}$, is the solution to

$$(14) \quad \sum \Delta x_t(\Delta y_t - \hat{\psi}\Delta x_t) = 0$$

which implies that

$$(15) \quad \hat{\psi} = \sum x_t \Delta y_t \Big/ \sum (\Delta x_t)^2 = \psi + \left[\sum x_t e_t \Big/ \sum (\Delta x_t)^2\right].$$

The instrumental variables approach assumes that we can find an instrumental variable, z_t, that is uncorrelated with the innovation to the technology shock, but is correlated with the inputs. Given an instrument, we can use this identifying assumption to estimate $\hat{\psi}$. We do this by choosing the value of returns to scale that sets the sample covariance between z_t and the innovation to zero:

$$(16) \quad \sum z_t(\Delta y_t - \psi\Delta x_t) = 0$$

which implies that

$$(17) \quad \hat{\psi} = \sum z_t \Delta y_t \Big/ \sum z_t \Delta x_t.$$

Thus, the instrumental variables estimate of returns to scale is obtained by setting the sample covariance between the instrument and output divided by the sample covariance between the instrument and the weighted sum of the inputs to zero. This condition is often called an *orthogonality condition*. (Note that the ordinary least squares estimate is a special case of the instrumental variables estimate in which the instrument and the right side variable are the same.)

Investigating Imprecise Measurement

In this section, we investigate why returns to scale measurement is imprecise, and we evaluate our empirical results.

Large standard errors and, as a consequence, wide confidence intervals are common in measuring aggregate returns to scale. For example, Burnside, Eichenbaum, and Rebelo (1995, p. 93), using standard measures of inputs and output, report a standard error for economywide returns to scale of 0.34. With this standard error and a point estimate of returns to scale of 0.98, a 95 percent confidence interval for returns to scale ranges from about 0.3 up to about 1.66. Baxter and King (1991) estimate aggregate returns to scale of 1.53 with a standard error of 0.56, which yields a 95 percent confidence interval from about 0.4 to about 2.65. These estimates show that the value of aggregate returns is measured imprecisely. In what follows, we will investigate why.

Our investigation exploits the fact that the constructed technology shock is a function of returns to scale. If returns to scale measurement is precise, the shocks will differ considerably depending on the value of returns to scale: the shock processes will look different, their serial correlation properties will be different, and their variances will be different. If returns to scale measurement is imprecise, however, then the shocks will be insensitive to variations in returns to scale. In this case, for various values of returns to scale, the shock processes will look similar, their serial correlation properties will be similar, and their variances will be similar. Consequently, detecting changes in the shocks for various values of returns to scale will be more difficult.

Our analysis consists of restricting the order of the stochastic process of the technology shock, inferring the innovations to the shock process for various values of returns to scale, and comparing the shocks and their stochastic properties. We then ask, How similar are the innovations to the technology shock for different values of returns to scale?

We conduct the analysis under the widely used assumption that the technology shock process (λ_t) is a log-random walk with drift (γ):

$$(18) \quad \ln(\lambda_t) = \ln(\lambda_{t-1}) + \gamma + \varepsilon_t.$$

When we define y as the natural log of output, k as the natural log of the capital stock, and l as the natural log of labor input, we can infer the innovation to the shock for any returns to scale value as follows. Consider a value for returns to scale, denoted by ψ^i, where $\psi^i = 1/(1-\phi_i)$. Next construct the difference between output growth and a weighted sum of the growth of inputs that has been scaled by ψ^i:

$$(19) \quad \eta_t^i \equiv \Delta y_t - \psi^i[\theta\Delta k_t + (1-\theta)\Delta l_t].$$

With the log-random walk assumption for the technology shock, demeaning x_t^i yields the innovation to the shock:

$$(20)\quad e_t^i = \eta_t^i - \bar{\eta}_t^i$$

where $e_t^i = \psi^i \varepsilon_t^i$. When we denote the true value of ψ by ψ^*, we see that the relationship between the constructed innovation to the technology shock (e_t^i) and the true innovation (e_t^*) is

$$(21)\quad e_t^i = e_t^* - (\psi^i - \psi^*)[\theta\Delta\hat{k}_t + (1-\theta)\Delta\hat{l}_t]$$

where $\Delta\hat{k}_t$ and $\Delta\hat{l}_t$ are the demeaned values of the percentage change in capital and labor, respectively.

If the value considered for returns to scale, ψ^i, is near the true value, then the implied innovation e_t^i should be similar to the true innovation and should also be approximately a *white noise*[4] process:

$$(22)\quad E(e_t^i e_{t-j}^i) = 0$$

for all $j \neq 0$.

If the considered value is not near the true value, then, as we have already noted, the constructed innovation e_t^i will differ from the true innovation and, consequently, will not be a white noise process:

$$(23)\quad E(e_t^i e_{t-j}^i) \neq 0$$

for all $j \leq s$. This is because $\theta\Delta\hat{k}_t + (1-\theta)\Delta\hat{l}_t$ will depend on the history of the shocks through the effect of the level of $\hat{k}_t$ on investment and labor effort decisions.

Empirical Results

To evaluate the differences in the technology shock at various values of returns to scale, we use annual U.S. data from 1949 to 1998. The data, from the Bureau of Labor Statistics (BLS) of the U.S. Department of Labor, consist of gross domestic product (GDP), adjusted for inflation, minus compensation of government workers; *labor input,* defined as private hours worked from the BLS establishment survey; and the stock of private capital. We need a value for the parameter θ to construct the shock. In our model, this parameter corresponds to capital's share of income. We use 0.35, which is the ratio of *capital income* (corporate profits, depreciation, rental income of persons, and net interest) to real GDP minus indirect business taxes and minus proprietor's income.[5]

The first part of our analysis is a simple visual comparison of the constructed innovations to the technology shock from various returns to scale values. The accompanying chart shows the constructed innovations to the technology shock for significant decreasing returns (0.8), constant returns (1.0), and relatively large increasing returns (1.2). The most striking aspect of this chart is the similarity of the three series. At many points, all three lie nearly on top of each other. This similarity suggests that the driving innovations to technology shocks are insensitive to the value of returns to scale.

We turn next to formally assessing what our assumption that the shock process follows a random walk with drift can tell us about returns to scale. We start by asking for which values of returns to scale the innovations appear to be white noise. We then calculate the autocorrelations of the innovations for the following values of returns to scale: {0.8, 0.9, 1.0, 1.1, 1.2}. Table 1 shows the standard deviation (σ) and the first six autocorrelations of each innovation series. The sample autocorrelation at lag k is defined as the sample covariance between the innovations with displacement k divided by the product of the standard deviations of these terms:[6]

$$(24)\quad \sum e_t e_{t-k} \Big/ (T-k) \div \left\{ \left[\left(\sum e_t\right)^2 \Big/ T\right]\left[\left(\sum e_{t-k}\right)^2 \Big/ (T-k)\right]\right\}^{1/2}.$$

The standard error for each of these autocorrelations is 0.07.

The autocorrelations in Table 1 appear to be both fairly low, which we would expect if the innovations were white noise, and quite similar. To formally test for white noise, we use the Ljung-Box test. Under the null hypothesis of white noise, this statistic is distributed as a χ^2 random variable with degrees of freedom equal to the number of autocorrelations being examined. Using 12 autocorrelations, we fail to reject the hypothesis of white noise for each of these series. The values for the test statistic for these five values are 7.53, 7.30, 6.94, 6.48, and 6.07. These values are all below the critical value of 21.55 for a 5 percent test. This implies that we can't use our assumption that the shock process is a random walk with drift to infer the value of returns to scale.

---

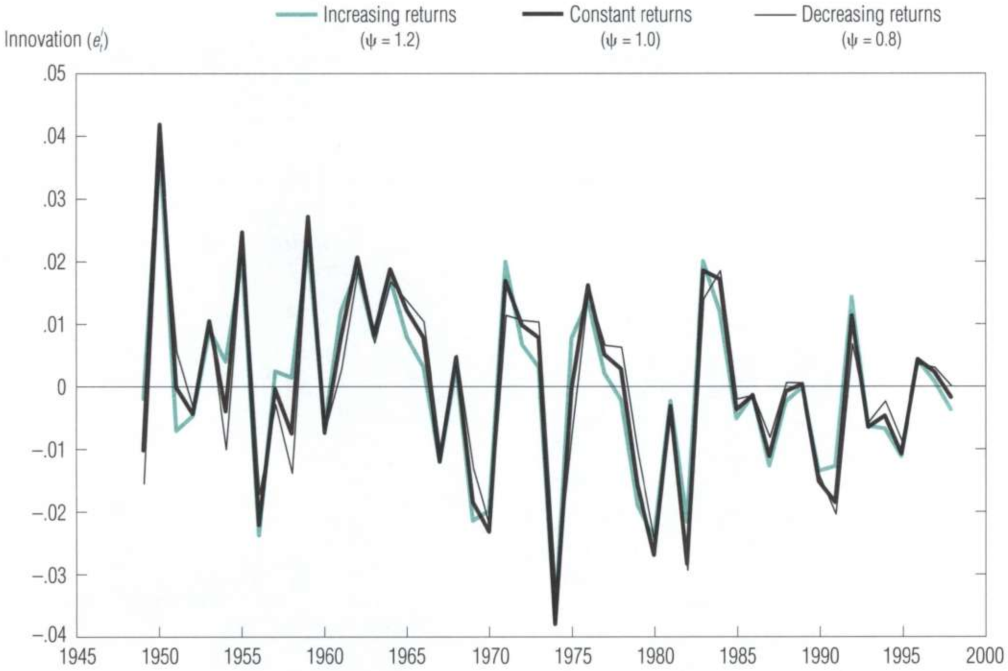
[4]The random variable ε_t is a white noise process if $E(\varepsilon_t \varepsilon_{t-k}) = 0$ for all k, $k \neq 0$.

[5]Our results are not sensitive to small changes in θ.

[6]Note that the autocorrelations of e_t^i are the same as those for ε_t^i, because ψ^i drops out. Furthermore, the correlations between e_t^i and e_t^j are the same as those between ε_t^i and ε_t^j for the same reason.

A Look at the Insensitivity of Technology Shocks to the Value of Returns to Scale

Technology Shock Innovations for Various Values of Returns to Scale;
Based on Stochastic Process for Technology Shocks and Annual U.S. Data for 1949–98



Sources of basic data: U.S. Department of Commerce, Bureau of Economic Analysis;
U.S. Department of Labor, Bureau of Labor Statistics

Instrumental variables estimation takes advantage of both a restriction on the stochastic process and an orthogonality assumption. However, if the innovations are too similar at various values of returns to scale, then finding appropriate instruments is likely to prove difficult. To shed more light on the similarities between innovations to the technology shock at various values of returns to scale, we calculate the correlations between these series for these returns to scale values. Table 2 shows these correlations. We find that all these series are highly correlated. For example, the correlation between the innovations for constant returns and aggregate returns of 1.2 is 0.97. Moreover, we find a strong similarity even for large differences in returns to scale: the correlation between the innovations at 0.8 and 1.2 is about 0.90.[7]

The message that emerges from these simple comparisons is that not only is a white noise process supported over a wide range of returns to scale values, but the innovations to the technology shock are insensitive to the value of returns to scale. These data show the problem with measuring returns to scale precisely: the innova-

[7] We also conducted this analysis using quarterly measures of GDP and labor input and constructing a quarterly measure of the capital stock. To do this, we used the annual measures of the capital stock compiled by the Bureau of Economic Analysis of the U.S. Department of Commerce and the quarterly investment flows from the national income and product accounts. Given these data, we solved for the annual depreciation rate of capital such that the sequence of stocks was consistent with the quarterly flows. We used the derived depreciation rate plus the quarterly investment series to construct a quarterly measure of the capital stock. The correlations between the innovations based on the quarterly data were even higher than those for the annual data.

tions to the technology shocks are virtually the same over a wide range of values. The shocks have similar variances and autocorrelations, and all the innovations are highly correlated. Because our model implies that the log of any constructed technology shock process is equal to the log of the original shock process plus the scaled log of output, the similarity of these series suggests that the shock and output are highly correlated. Thus, the theoretical difference that arises between the shock processes from the inclusion of capital is not quantitatively important.

From an applied perspective, the fact that the residuals from the production function are so insensitive to changes in returns to scale values is bad news for measuring returns to scale precisely. To understand this intuitively, consider the high correlations between the innovations shown in Table 2. If returns to scale were unidentified, then the correlation between the innovations at various values of returns to scale would be exactly one, and the covariance between the instrument and the innovation would be identical for all values of returns to scale. Now consider what happens as the correlation between the innovations approaches one. Although returns to scale can be identified theoretically in this case, returns will be hard to measure precisely, particularly in small samples, because if the innovations are highly correlated at various values of returns to scale, then the sample covariance between the instrument and the innovation at these values can also be similar. Thus, the covariance between the instrument and the innovation can be *flat;* that is, the covariance changes very little as the returns to scale value changes. This suggests that many values of returns to scale are about equally likely, given the data.

This flat slope of the orthogonality condition is important, because it can make ruling out all but extreme values of returns to scale difficult. To see this formally, consider the sample variance of the instrumental variables estimate of returns to scale. When the variance of the innovation to the technology shock is denoted by σ^2, the variance of the estimate of returns to scale, σ^2_ψ, is

$$(25) \quad \sigma^2_\psi = \sigma^2 \sum \Delta x_t^2 \Big/ \left(\sum z_t \Delta x_t\right)^2.$$

The variance of the estimate depends on two components: the variance of the innovation and the sample variance of the inputs relative to the squared covariance between the inputs and the instrument. To understand clearly how the slope of the orthogonality condition relates to the variance of the estimator, consider the case for the lowest possible variance for the instrumental variables estimator of returns to scale. This is the case in which the correlation of the instruments with the inputs approaches one. In this case, the orthogonality condition (16) used to estimate returns to scale becomes

Tables 1–2

Assessing Returns to Scale Using Statistical Properties of Technology Shock Innovations

Table 1 Autocorrelations of Innovations: Low and Similar

	Value at Returns to Scale of				
	.8	.9	1.0	1.1	1.2
Autocorrelation					
ρ (1)	−.067	−.061	−.051	−.036	−.015
ρ (2)	−.028	−.021	−.020	−.029	−.048
ρ (3)	−.109	−.104	−.098	−.092	−.086
ρ (4)	−.016	−.002	.016	.038	.064
ρ (5)	.136	.149	.161	.172	.183
ρ (6)	−.112	−.112	−.112	−.112	−.112
Standard Deviation (σ)	.016	.015	.014	.014	.013

Table 2 Correlations Between Innovations: High

	Value at Returns to Scale of				
Returns to Scale	.8	.9	1.0	1.1	1.2
.8	1.000	.995	.979	.949	.902
.9	.995	1.000	.994	.975	.939
1.0	.979	.994	1.000	.993	.970
1.1	.949	.975	.993	1.000	.992
1.2	.902	.939	.970	.992	1.000

(26) $\sum \Delta x_t(\Delta y_t - \psi \Delta x_t) = 0.$

The variance of this estimate of returns to scale approaches the following ratio:

(27) $\sigma_\psi^2 = \sigma^2 \Big/ \sum \Delta x_t^2.$

This simpler measure of the variance depends on just the innovation variance and the variance of the inputs. To see how the variance is related to the slope of the orthogonality condition, differentiate the orthogonality condition (26) with respect to the returns to scale value to get an estimate of the slope:

(28) $\partial \sum \Delta x_t(\Delta y_t - \psi \Delta x_t)/\partial \psi = -\sum \Delta x_t^2.$

This derivative measures how the covariance changes as the returns to scale value changes. If the slope is flat—that is, if the covariance is insensitive to changes in the returns to scale value—then $-\Sigma \Delta x_t^2$ is small. The key implication of this derivative for the variance of the instrumental variables estimate is that the derivative is the denominator of the variance. Thus, if the slope of the orthogonality condition is flat, then the slope is near zero, and the variance of returns to scale will be large. Thus, an important reason for imprecise measurement of returns to scale is that there is insufficient variation in the inputs. This feature of the data has negative implications for measuring aggregate returns to scale more precisely.

Conclusion

The value of aggregate returns to scale has key implications for the sources of various shocks that lead to business cycle fluctuations. Many economists have measured returns to scale with instrumental variables techniques. However, it is hard to find instruments that can be considered exogenous and that are also correlated with factor inputs. A lot of economic research has been devoted to finding instruments that have these two properties. In this article, we analyze a more basic problem with measuring returns to scale: insufficient variation in the factor inputs. We have shown that residuals inferred from a standard aggregate production function over a wide range of return to scale values are nearly identical. This problem is independent of the characteristics of the instruments. Consequently, the likelihood function in this range of values is flat, and the standard error is large.

Our conclusion is not optimistic: given observed variation in standard measures of inputs, we don't think we can measure aggregate returns to scale precisely. To measure returns more precisely, our analysis indicates that more variation is needed in the inputs. Some economists have argued that the inputs, particularly capital, vary more over the business cycle than is observed in the standard measures of factor services. Applications of this idea include those of Burnside, Eichenbaum, and Rebelo (1995), who measure returns to scale in overall manufacturing using alternative input measures for capital, and Basu (1996), who measures returns to scale at the two-digit level in manufacturing using variations in materials in a gross output production function. While this interesting work may advance our ability to measure aggregate returns to scale precisely, measuring returns to scale generically with alternative input measures is a tall order. Generic measurement requires a generally accepted theory of variable factor utilization and a generally accepted alternative measure of factor services. Moreover, with latent shocks and latent factor services, the identification problem may be even more troublesome.

In the meantime, our knowledge of aggregate returns to scale based on standard measures of inputs is limited. Because distinguishing between constant returns and increasing returns sufficiently large to generate sunspot equilibria is difficult, our ability to determine whether business cycle fluctuations are due primarily to technology shocks, monetary shocks, preference shocks, or other shocks is also limited.

References

Basu, Susanto. 1996. Procyclical productivity: Increasing returns or cyclical utilization? *Quarterly Journal of Economics* 111 (August): 719–51.

Baxter, Marianne, and King, Robert G. 1991. Productive externalities and business cycles. Institute for Empirical Macroeconomics Discussion Paper 53. Federal Reserve Bank of Minneapolis.

Benhabib, Jess, and Farmer, Roger E. A. 1994. Indeterminacy and increasing returns. *Journal of Economic Theory* 63 (June): 19–41.

Burnside, Craig; Eichenbaum, Martin; and Rebelo, Sergio. 1995. Capital utilization and returns to scale. In *NBER Macroeconomics Annual 1995*, ed. Ben S. Bernanke and Julio J. Rotemberg, Vol. 10, pp. 67–110. Cambridge, Mass.: MIT Press.

Cole, Harold L., and Ohanian, Lee E. 1996. Reassessing aggregate returns to scale with standard theory and measurement. Research Department Working Paper 566. Federal Reserve Bank of Minneapolis.

Kamihigashi, Takashi. 1996. Real business cycles and sunspot fluctuations are observationally equivalent. *Journal of Monetary Economics* 37 (February): 105–17.

McGrattan, Ellen R. 1994. A progress report on business cycle models. *Federal Reserve Bank of Minneapolis Quarterly Review* 18 (Fall): 2–16.

Prescott, Edward C. 1986. Theory ahead of business cycle measurement. *Federal Reserve Bank of Minneapolis Quarterly Review* 10 (Fall): 9–22.