

Federal Reserve Bank of Minneapolis  
Research Department

**Reassessing Aggregate Returns to  
Scale With Standard  
Theory and Measurement**

Harold L. Cole and Lee E. Ohanian\*

Working Paper 566

July 1996

**ABSTRACT**

Constant returns to scale is a central construct of neoclassical theory. Previous studies argued that one must adopt a specification of the production function with substantial unobserved service variation to reconcile constant returns with the data. Some economists have argued that this finding has not resolved the size of returns to scale, since factor service variation is unobserved, and there is no generally accepted theory to guide specification of this alternative framework. In this paper we show that the stochastic version of the neoclassical growth model delivers an orthogonality condition which can be used to estimate returns to scale. Rather than the standard finding of increasing returns, we show that standard theory and conventional measures of output and inputs yield estimates of constant returns to scale at the aggregate level. Our estimates also suggest that factor service variation is not an important determinant of output fluctuations.

\*Cole, Federal Reserve Bank of Minneapolis; Ohanian, University of Minnesota and Federal Reserve Bank of Minneapolis. We would like to thank Susanto Basu, Toni Braun, John Geweke, Peter Klenow, Narayana Kocherlakota, Mathew Shapiro, and Victor Rios-Rull for their comments, and Basu for providing his data. We would also like to thank Daniel Houser and Jason Schmidt for research assistance. The views expressed herein are those of the authors and not necessarily those of the Federal Reserve Bank of Minneapolis or the Federal Reserve System.

## 1. Introduction

The size of returns to scale in the aggregate production function has important implications for many questions in macroeconomics. Analyses of business cycles, growth, and the scope of government policy depend fundamentally on whether there are constant returns to scale (CRS) or increasing returns to scale (IRS). For example, with IRS, competitive equilibria are no longer Pareto optimal, and there may be an important, welfare enhancing role for activist fiscal and/or monetary policy. With regard to growth, if the aggregate production technology is CRS and depends on measured inputs of capital and labor, then long-run growth appears to be primarily determined by technological factors, while models which exhibit IRS can imply that growth is largely due to increases in accumulable factors. Finally, while business cycle models driven primarily by demand shocks are difficult to reconcile with aggregate data U.S. data if the aggregate production function is assumed to have CRS, demand shock models with IRS can be reconciled easily with the data.

A number of recent studies have estimated the extent of returns to scale. Two main findings emerge from this literature. The first is that estimation of the conventional production function with standard measures of output and inputs produces estimates of increasing returns to scale (see for example Hall (1988), King and Baxter (1991), and the discussion by Burnside, Eichenbaum and Rebelo (1995)). This finding has motivated researchers to estimate returns to scale using different methods, specifications, and data. The second main finding, which is drawn from these very recent studies, is that constant returns to scale estimates *can* be obtained, provided that an alternative specification of the production function is adopted, with substantial unmeasured variation in capital and/or labor services (Burnside, Eichenbaum and Rebelo (1995), Basu(1995), Basu and Fernald (1995)).<sup>1</sup>

However, the methodologies used in these studies have been questioned, and the findings of constant returns to scale has been discounted. Hall (1995) has criticized the specification of non-standard production functions and the proxies used for the true unmeasured

---

<sup>1</sup>An exception to this is Shapiro (1993) who uses unpublished data from the Census's Survey of Plant Capacity to estimate the workweek of capital between 1977-1988. He finds that when capital hours are taken into account there is no evidence of increasing returns in a panel of four-digit manufacturing industries. Unfortunately, the short duration of the data and the limitations on its coverage precludes its use given our focus on aggregate annual data.

inputs. First, he notes that there are many proxies for labor and/or capital services that can yield constant returns to scale, and that the common feature of these proxies is that they are highly correlated with output. However, input services are unobservable in these models, and theory provides little guidance in choosing alternative specifications of the production function or proxies for factor services. Thus, he argues that it is not clear what type of alternative specification or data proxy are reasonable, or what the appropriate criteria are for choosing among different specifications. He concludes from these observations that estimates of constant returns from these alternative specifications are not compelling.

Given these criticisms of this recent literature, we estimate returns to scale in the aggregate production function using standard theory and measurement, but we use an estimation technique that differs from methods used in earlier studies. In developing our analysis, we note that estimating returns to scale typically requires the use of instrumental variables. This is an important issue, since it is difficult to find variables that are arguably uncorrelated with the "error term", but reasonably correlated with the endogenous variables in the model. We show that the neoclassical growth model implies an orthogonality condition between the capital stock, which is predetermined, and the innovation to the Solow parameter (total factor productivity) *only* at the true value of returns to scale. In addition, we show that capital is correlated with the endogenous variables. This suggests that capital is a good instrument. We exploit this orthogonality condition implied by the growth model to estimate returns to scale using a simple method of moments procedure. Our results are surprising, in that our findings provide evidence against both components of the conventional view. The estimates show that (i) U.S. data appear to be broadly consistent with CRS with standard theory and measurement, and (ii) assuming that returns are not decreasing, there is little unobserved cyclical service variation.

To test the robustness of our results, we show that estimates of returns to scale are roughly one for a variety of different stochastic specifications for the Solow process. These include integrated processes, such as log-random walk with drift, and also autoregressive processes that are stationary around time trends. Moreover, the finding of constant returns is not sensitive to a more complicated specification of the Solow parameter, which in addition to the its own past history, includes other variables such as government spending, capital and

labor tax shocks. We also show that even allowing for plausible amounts of measurement error in capital does not significantly change our findings.

Since part of the motivation of this analysis comes from critiques of studies that have adopted different specifications of the production function with unmeasured factor service variation, we summarize the well known papers in this area. This is discussed in section 2. The economic model we use in our analysis is presented in section 3. Our estimation technique is presented in section 4. Our results are presented in section 5. Section 6 provides a summary and conclusion.

## 2. Literature Review

There is an extensive and old literature that has estimated returns to scale. Despite the large number of empirical studies, there has been lack of consensus on returns to scale at the aggregate level. This is due partially to the fact that different methods and data have yielded different returns to scale estimates, but also reflects the fact that identifying and consistently estimating returns to scale is very difficult, requiring a number of strong, and in some cases questionable, identifying and exogeneity assumptions. This has led economists to question seriously the empirical findings from this literature, and more recently, to try to develop new techniques and use alternative data to produce more robust estimates of returns to scale.

The fundamental problem in estimating returns to scale is that the shock to the production function is unobservable, and that theory predicts that the inputs are correlated with this shock. To understand this problem, consider the standard log-linear representation of the production function  $y_t = \lambda_t + \theta_t k_t + (1 - \theta_t)l_t$ , where  $\lambda_t$  is the log of the technology parameter,  $k_t$  is the log of the capital stock and  $l_t$  is the log of total hours. It is very common in the literature to assume that the technology parameter follows a random walk with drift process and hence to work with this specification in first-difference form:

$$\Delta y_t = \gamma_0 + \gamma_1[\theta \Delta k_t + (1 - \theta) \Delta l_t] + \epsilon_t, \quad (1)$$

where  $\epsilon_t$  is assumed to be the innovation to the technology process. However, estimating a relationship like (1) under ordinary least squares (OLS) would lead to upwardly biased

estimates, since the OLS estimate is  $\hat{\gamma} = \gamma + (X'X)^{-1}X'\epsilon$ , (where the elements of  $X$  are the inputs), and standard theory predicts that some elements of  $X$  and  $\epsilon$  positively covary.

Recent approaches to estimating returns to scale have tried to address this issue by using instrumental variables. For example, Hall (1988) used instrumental variables, and found evidence that was inconsistent with competitive behavior and a constant-returns-to-scale production function in sectoral data. Hall tested whether  $(\Delta y_t - \Delta k_t) - (1 - \theta_t)(\Delta l_t - \Delta k_t)$  was orthogonal to instruments which he argued on a priori grounds were orthogonal to the rate of technological progress, but not to the level of factor inputs. These instruments were the first-differences of the “Ramey instruments” which include: (1) the political party of the president, (2) the nominal price of oil, and (3) the level of military spending. Hall found strong evidence against a zero correlation. Motivated by this work, a number of other studies have found in favor of increasing returns using the standard production function and standard inputs.

An alternative interpretation of this evidence has been to argue that the inputs have been mis-measured, and in particular, that standard measurement of capital and labor significantly understate the true variability of these inputs. This argument suggests that estimates of returns to scale based on conventional input measures will be biased upwards if the ratio of true factor services to measured factor services is procyclical. There are two notable approaches have been used which focus on the mis-measurement of inputs, and how to deal with the unobservability of true factor services. The first strategy, as exemplified by Burnside, Eichenbaum and Rebelo (BER) (1995) and Caballero and Lyons (CL) (1992) has used electricity or energy consumption to measure the cyclical variation in factor services. The second strategy, as exemplified by Basu and Fernald (1994) and Basu(1995) (BF), has used materials as a proxy for true capital and labor services within a gross output production function, rather than value-added.<sup>2</sup>

---

<sup>2</sup>Baxter and King pursue a third strategy. They consider a dynamic general equilibrium model in which there are both preference and productivity shocks. The specification of their model is such that they can use the observed quantities to infer both the preference and productivity shocks given the observed quantities and the returns parameter. They estimate the stochastic process generating both the technology and preference shock process and then simulate their model given these processes. They argue that the model with the returns parameter equal to 1.3 does a better job of matching the data. While their results are interesting, their finding seems quite sensitive to the assumed form of preferences and the preference shock.

Under the first strategy, BER and CL argue that capital services are not proportional to the capital stock, but rather that there are substantial variations in how intensively capital is used. BER construct an alternative measure of capital services which assumes that true capital input is a function of energy. However, this raises the difficulty of seeking to divide the observed change in  $\Delta y_t - (1 - \theta)\Delta l_t$  into an increase in capital services and an increase in technology. BER develop a structural relationships relating true capital input to electricity consumption, and used this structural relationship to substitute for capital services in the production function. Their preferred specification for capital services was  $K^s = \nu E$ , which implies that  $\Delta k_t^s = \Delta e_t$ . This specification yields an equation of the form:

$$\Delta y_t = \gamma_0 + \gamma_1[\theta\Delta e_t + (1 - \theta)\Delta l_t] + \epsilon_t. \quad (2)$$

To understand the implications of use of electricity as a proxy for capital services, we have plotted in figure 1 total factor productivity under the assumption of constant returns to scale and that true capital input is proportional to electricity use, as in the BER specification. The solid line is the log of TFP (up to a constant) under the assumption that the relevant measure of electricity is industrial consumption, which is the measure used by BER. The dashed line is TFP under the assumption that the relevant measure of electricity is industrial plus commercial electric use. The behavior of productivity implied by these measures seems puzzling. Under both measures, productivity drops sharply between 1948 and 1956. With the narrower measure of electricity use there is no productivity growth relative to the initial level for the first 30 years, while the broader measure implies that there has been no productivity growth over the entire postwar period. This figure confirms that on average, electricity use has grown significantly faster than the capital stock, and that its use for measuring total factor productivity may be inappropriate.<sup>3</sup>

BER argue that while the level of total factor productivity may be biased by the electricity proxy, the cyclical implications of this proxy may not suffer from this problem. To explore this issue, we have estimated returns using OLS for the standard production specification in equation (1), and for the BER specification in equation (2). These results are

---

<sup>3</sup>The trend in the electricity/capital ratio, and its implications for productivity measurement, are discussed in the Survey of Current Business, 1972.

presented in table 1 and are over the 1949-1993 period using annual data.<sup>4</sup> These estimates show that returns to scale is sensitive to the inclusion of electricity in the production function. The standard specification generates evidence of increasing returns, though this estimate is likely biased upwards. In fact, it is interesting to note that this estimate is roughly what one would expect from a real business cycle model *with* a CRS production function; use of calculations similar to Aiyagari (1994) suggest the regression coefficient bias may be on the order of 0.15 to 0.4.<sup>5</sup> The BER formulation yields returns to scale of .76, which implies *decreasing* returns. Moreover, since OLS estimates may be biased upwards substantially, .76 can be reasonably viewed as an upper bound on returns to scale<sup>6</sup>.

Table 1: OLS Estimates

cases/parameters	$\gamma_0$	$\gamma_1$	$R^2$
ST	0.009	1.20	.80
	(.002)	(.092)	
BER	0.006	0.76	.83
	(.002)	(.052)	

An alternative strategy has been pursued by Basu, and Basu and Fernald, who specify a sectoral gross output production function, rather than value added. They use materials (intermediate inputs) as a proxy for capital and labor. To gain insight into their approach, assume that the production function for gross output is given by:

$$y_t^i = H(\lambda_t^{1i} F(k_t^i, l_t^i), \lambda_t^{2i} m_t^i).$$

<sup>4</sup>The electricity data was from the Energy Information Administration's *Annual Energy Review 1994*, table 8.6, and is electric utility retail sales by end-use-sector measured in billions of kilowatthours. We discuss the other data in more detail in section 5 when we present our results.

<sup>5</sup>Following Aiyagari (1994) assume detrended labor input is given by  $l_t = \gamma_l \lambda_t + \zeta_t$ , where  $\zeta_t$  represents all the fluctuations in  $l_t$  arising from factors that are orthogonal to the contemporaneous productivity level, and that detrended output is  $y_t = \lambda_t + (1 - \theta)l_t$ . Assume that the values of  $\gamma_l$ ,  $var(\zeta)/var(\lambda)$ , and  $\theta$  are chosen so that the model correctly predicts that: (1) productivity and labor input are uncorrelated, (2) the variability of labor input relative to output is 0.85, and (3) labor's factor share matches its long term average of about 0.7. Condition (1) implies that  $cov(y_t, l_t) = var(l_t)$ , and that a regression of  $y_t$  on  $(1 - \theta)l_t$  recovers a value of 1.4 ( $= 1/(1 - \theta)$ ). If we included capital in our analysis, this would tend to lower the coefficient since capital is less correlated with output than is labor.

<sup>6</sup>BER find roughly constant returns to scale, using the 1977-93 sample period, quarterly data, and instrumental variables.

If  $H$  is linear homogeneous and  $\lambda_t^{1i} = \lambda_t^{2i} = \lambda_t^i$ , then this can be written as:

$$y_t^i = \lambda_t^i H(F(k_t^i, l_t^i), m_t^i),$$

This is the specification adopted by Basu-Fernald. Since they find that the growth rates of materials and output at the sectoral level are virtually perfectly correlated, their preferred specification of  $H$  is Leontieff. Given this specification, they estimate the relationship between materials and gross output using the Ramey instruments.<sup>7</sup> They find in favor of moderate increasing returns to scale.

Materials are a good proxy for capital and labor services under the assumption that  $\lambda_t^{1i} = \lambda_t^{2i}$ . However, this assumption does not seem to be consistent with conventional theories of technical progress, which assume that advances in technology are embodied in labor (human capital) or new capital equipment (vintage capital models). If this assumption was not satisfied, then materials may not be a useful proxy for capital and labor services. We argue that the maintained assumption is inconsistent with U.S. manufacturing data. Our argument is based on the relationship between material and gross output. The BF data (from the manufacturing sector), are presented in figure 2. The most striking feature of these data is that the raw growth rates of these two series are virtually identical. This observation has several important implications. First, it implies that  $y_t^i/m_t^i$  is constant. Thus, either the relative price of materials has also been constant during this period, or  $H$  is Leontieff, which seems to be their preferred specification. However, if  $H$  is Leontieff, then the growth rate in  $\lambda_t^i$  must be zero, since the growth rate in gross output is the sum of the minimum of the growth rates of value added and materials, plus the growth rate of the technology parameter. Thus, the nearly perfect correlation between gross output and materials, with a Leontieff specification, implies no technical progress over the postwar period.

An alternative interpretation is that technological innovations have come in the form of changes to  $\lambda_t^{1i}$  and not to  $\lambda_t^{2i}$ . However, this implies that materials may be a good proxy for  $y$  or  $\lambda_t^{1i}F(k_t^i, l_t^i)$ , but not  $F(k_t^i, l_t^i)$ . Moreover, this suggests that OLS estimation of gross

---

<sup>7</sup>When they assume that  $H$  is homogenous, they estimate the returns parameter along with the elasticity of substitution using the relative price of materials and the Ramey instruments.

output on materials should recover a coefficient of roughly 1 (CRS).

From this perspective, it is surprising that the BF IV estimates show moderately increasing returns to scale, since instruments should correct for the possible upward bias in OLS estimation. Given the presumption that OLS is biased upwards, this suggests that either the instrument is correlated with the true innovation, or that the model is mis-specified.<sup>8</sup> Additionally, the Solow residual backed out by BF and BER is uncorrelated with the inputs. Rather than a theoretical puzzle, this finding may also be due to instruments that are correlated with the true productivity shock.

The main difficulty with seeking to use proxies for unmeasured service variation is that there are three unmeasured components in the production function: capital services, labor services and productivity. As Hall (1995) points out, there is no standard, accepted theory for cyclical factor utilization to help guide the specification of the alternative model. He notes that “it is well established that adding a free variable to a cyclical productivity equation will eliminate evidence of increasing returns to scale. These free variables are highly correlated with output. Research now needs to turn to the issue of whether the role of the free variable makes sense as a matter of theory...” The theoretical and measurement difficulties associated with these alternative specifications lead us to estimate returns to scale using the standard production function with conventional measures of inputs, but with different techniques than those used in previous research.

### 3. Model

The basic model is similar to that used by Benhabib and Farmer (1994) and Baxter and King (1992). Time is discrete, and there is a single capital/consumption good produced each period. There is a measure one number of identical infinitely lived consumers. The individuals’ preferences and budget constraint are given by:

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

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

---

<sup>8</sup>This result may also reflect the fact that the Ramey instruments are not highly correlated with inputs; see for example, John Shea (1993) and BER.

where  $b_t$  and  $r_t$  denote the individuals borrowing level and the gross interest rate respectively, and  $\tau_t$  denotes a lump-sum tax,  $c_t$  is consumption,  $1 - l_t$  is leisure (non-market time), and  $d_t$  is a preference (or home production) shock.<sup>9</sup> Per-capita production is given by:

$$y_t = \lambda_t F(k_t, l_t) Y_t^\phi$$

where  $k_t$  and  $l_t$  denote the individuals levels of capital and labor, and  $Y_t$  is aggregate per-capita output.<sup>10</sup> Following other work in this literature, we assume that the production function  $F(\cdot)$  is a linear homogeneous Cobb-Douglas function with “capital’s share” given by  $\theta$ . The term  $\lambda_t$  is the aggregate technology shock. The parameter  $\phi$  determines the size of the externality. The economy is defined as “neoclassical” (aggregate constant returns to scale) if  $\phi = 0$ . While we pursue our analysis within the context of a model where increasing returns reflects an externality, our set-up is also consistent with increasing returns internal to the organization, provided that pure profits are negligible.

We define  $\epsilon_t$  as a  $4 \times 1$  vector of i.i.d. random variables, with  $\epsilon^t \equiv \{\epsilon_s\}_{s=0}^t$ . Let the  $t$  period realization of the technology shock, the government spending shock, the preference shock, and, where relevant, the sunspot variable be given by  $\lambda_t(\epsilon^t)$ ,  $g_t(\epsilon^t)$ ,  $d_t(\epsilon^t)$  and  $v_t(\epsilon^t)$ .

The government’s budget constraint is

$$g_t = \tau_t$$

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

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

An equilibrium for this economy consists of a set of functions which describe: (i) the consumer’s policy  $\{c_t, l_t, k_{t+1}\}(\epsilon^t)$ , (ii) the gross interest rate  $\{r_t(\epsilon^t)\}$ , and (iii) aggregate per capita output  $\{\tilde{y}_t(\epsilon^t)\}$  such that

---

<sup>9</sup>Since households are identical, there will be no borrowing in equilibrium.

<sup>10</sup>The assumption that the aggregate externality is depends upon the level of per capita as opposed to aggregate output is driven by the observation that large countries do not seem to be more productive than small countries.

1. Consumer's maximize with  $\{c_t, l_t, k_{t+1}\}(\epsilon^t)$  and  $b_{t+1} = 0$ , given  $\{r_t(\epsilon^t)\}$  and  $\{\tilde{y}_t(\epsilon^t)\}$ .
2. The equation of motion for aggregate output is consistent with individual production decisions, or

$$\tilde{y}_t = [\lambda_t F(k_t, l_t)]^{\frac{1}{1-\phi}}.$$

#### 4. Estimating the Size of the Aggregate Externality Parameter

To identify total returns to scale, which in this model we denote as  $\psi \equiv 1/(1 - \phi)$ , we exploit the fact that the stochastic version of the Cass-Koopmans model implies that the current period capital stock will be uncorrelated with the unpredictable component (innovation) to the technology shock process  $\hat{\lambda}$  *only for the true*  $\psi$ ,  $\psi^*$ . This orthogonality condition comes from the fact that rational, optimizing individuals will take into account relevant information when making investment decisions. Since these decisions depend upon expected future events, individuals must make forecasts of these events, and one property of a good forecast is that the capital stock will be uncorrelated with the unpredictable component of the technology shock. For all other values for this parameter, the capital stock will be correlated with the innovation to the measured technology shock, since the measured process at a *wrong* value of  $\psi$  will be a function of not only the true process, but also of capital and labor<sup>11</sup>. Given the orthogonality between the innovation to the technology shock process and the capital stock at  $\psi^*$ , we use method of moments to estimate this parameter. In particular, the estimate of  $\psi$  in the method of moments framework is that which sets the sample covariance between the innovation to the technology shock and capital to zero. The maintained identifying assumption in our analysis, as in any IV estimation, is that we specify a stochastic process for the technology shock.<sup>12</sup> Rather than choosing one process, we will conduct our analysis under a number of alternative models for the technology shock.

---

<sup>11</sup>This is true as long as  $\theta var(k) \neq (\theta - 1)cov(k, l)$ .

<sup>12</sup>It is easy to show that given a  $\hat{\phi} \neq \phi$ , one can construct an alternative technology shock process  $\hat{\lambda}_t(\epsilon^t)$  such that

$$\hat{\lambda}_t(\epsilon^t)y_t(\epsilon^t)^{\hat{\phi}} = \lambda_t(\epsilon^t)y_t(\epsilon^t)^{\phi},$$

and hence the economy with this new shock process and externality parameter would be observationally equivalent to the first economy. However, this second economy will typically involve assuming a high order process for the technology shock since capital will typically depend in a complicated way on the history of past shocks.

### A. The First Order Integrated Case:

We first consider estimating  $\psi$  under the assumption that the true technology shock process ( $\lambda$ ) is a log-random walk with drift:

$$\ln(\lambda_t) = \ln(\lambda_{t-1}) + \mu + \epsilon_t^\lambda$$

Given the assumption of an integrated process, the moment conditions are  $E\{h(\chi, w_t)\} = 0$ , where  $\chi = (\psi, \mu)$ ,

$$h(\chi, w_t) = \begin{bmatrix} \Delta \ln(K_t) \epsilon_t^\lambda \\ \epsilon_t^\lambda \end{bmatrix},$$

and  $w_t$  denotes a sample realization. The sample moment conditions for estimating the parameters  $\psi$  and  $\mu$  are given by:

$$\frac{1}{T} \sum (\Delta \ln(K_t) \hat{\epsilon}_t^\lambda) = 0 \quad (3)$$

$$\frac{1}{T} \sum \hat{\epsilon}_t^\lambda = 0 \quad (4)$$

where  $T$  is the effective sample size.

Choosing parameters to satisfy these moment conditions can be handled in the standard method of moments framework of Hansen (1982). Given the nature of the problem, it reduces to a simple linear case, and standard errors of parameters can be computed using the usual formulas<sup>13</sup>. Substituting for  $\hat{\epsilon}_t^\lambda = \Delta y_t / \hat{\psi} - \theta \Delta k_t - (1 - \theta) \Delta l_t - \hat{\mu}$  yields two equations in two unknowns from which we can derive our estimates of  $\hat{\phi}$  and  $\hat{\mu}$ :

$$g(\hat{\chi}; Z_T) = \begin{bmatrix} \frac{1}{T} \sum (\Delta k_t (\Delta y_t / \hat{\psi} - \theta \Delta k_t - (1 - \theta) \Delta l_t - \hat{\mu})) \\ \frac{1}{T} \sum (\Delta y_t / \hat{\psi} - \theta \Delta k_t - (1 - \theta) \Delta l_t - \hat{\mu}) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

where  $\hat{\chi} \equiv (\hat{\psi}, \hat{\mu})$  denotes our vector of parameters, and  $Z_T$  denotes our set of sample obser-

---

<sup>13</sup>Christiano and den Haan (1995) note that estimation of standard errors may be misleading in sample sizes typically encountered in macroeconomics for dependent processes. Under our theory, the innovation to the technology shock process is an i.i.d. random variable, and thus should not suffer as much from these problems.

variations. The asymptotic standard errors of the parameter estimates are given by:

$$\left\{ \left[ \frac{dg(\chi; Z_T)}{d\chi} \right] S^{-1} \left[ \frac{dg(\chi; Z_T)}{d\chi} \right]' \right\}^{-1}$$

where  $S$  is the asymptotic variance of the sample mean of the moment conditions:

$$S \equiv \sum_{v=-\infty}^{\infty} E\{[h(\chi, w_t)][h(\chi, w_{t-v})]'\}.$$

Here

$$E\{[h(\chi, w_t)][h(\chi, w_{t-v})]'\} = \begin{bmatrix} (\Delta k_t)^2 \sigma_e^2 & \Delta k_t \sigma_e^2 \\ \Delta k_t \sigma_e^2 & \sigma_e^2 \end{bmatrix} \text{ if } v = 0,$$

and

$$E\{[h(\chi, w_t)][h(\chi, w_{t-v})]'\} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \text{ if } v \neq 0,$$

since given our theory  $\hat{\varepsilon}_t^\lambda$  is i.i.d. Thus our sample estimate of  $S$ ,  $\hat{S}$ , is given by the variance-covariance matrix of the sample moment conditions or

$$\hat{S} = (1/T) \sum \begin{bmatrix} (\Delta k_t \hat{\varepsilon}_t^\lambda)^2 & \Delta k_t (\hat{\varepsilon}_t^\lambda)^2 \\ \Delta k_t (\hat{\varepsilon}_t^\lambda)^2 & (\hat{\varepsilon}_t^\lambda)^2 \end{bmatrix}.$$

The matrix of first-partial derivatives is given by

$$\frac{dg(\hat{\chi}; Z_T)}{d\hat{\chi}} = \begin{bmatrix} -(1/T) \sum \Delta y_t \Delta k_t / \hat{\psi}^2 & -(1/T) \sum \Delta y_t / \hat{\psi}^2 \\ -(1/T) \sum \Delta k_t & -1 \end{bmatrix}$$

Thus the sample estimate for the standard errors is given by

$$\sqrt{\frac{1}{T} \left( \left[ \frac{dg(\hat{\chi}; Z_T)}{d\hat{\chi}} \right] \hat{S}^{-1} \left[ \frac{dg(\hat{\chi}; Z_T)}{d\hat{\chi}} \right]' \right)^{-1}}$$

## B. Higher Order Shock Specifications:

It is straightforward to estimate  $\psi$  using this technique for *any* technology shock process that is a stationary and finite stochastic process. For example, suppose that the

true process was  $AR(Q)$  around a deterministic trend, rather than the first-order integrated process described above. After detrending, this process is:

$$\ln(\lambda_t) = \sum_{i=1}^Q \rho_i \ln(\lambda_{t-i}) + \epsilon_{\lambda_t}$$

In this case, the  $Q$  autoregressive parameters  $\{\rho_i\}$  would be estimated with  $Q$  additional moment conditions of the form:

$$\frac{1}{T} \sum (\ln(\lambda_{t-i}) \hat{\epsilon}_{\lambda_t}) = 0, \quad i = 1, \dots, Q$$

Moreover, given that the lag order of the technology shock process typically will be unknown, it is straightforward to test for lag-length truncation. More general stochastic processes for the technology shock also could be handled in this framework.

### C. Time Varying Drift Term:

In our analysis, returns to scale are identified by a moment condition requiring orthogonality between capital and the unpredictable component in the technology shock. Given the extensive discussion of the productivity slowdown in the growth literature, a natural feature to incorporate into our specification of the shock process is a trend break in the technology shock process. It is straightforward to do this in our model. For example, assume that the drift term follows a finite state Markov process, and that the drift term in period  $t$ ,  $\mu_t$ , is known as of period  $t - 1$ .

In this case the unanticipated innovation to productivity parameter is still  $\epsilon_t$ . Conditional on the econometrician being able to identify the date of the drift break, our moment conditions in the integrated case, (3) & (4), change slightly. Assume that there is a single drift break during our sample in period  $T'$ . Then our moment condition becomes

$$g(\hat{X}; Z_T) = \begin{bmatrix} \frac{1}{T} \sum_{t=1}^T (\Delta k_t (\Delta y_t / \hat{\psi} - \theta \Delta k_t - (1 - \theta) \Delta l_t - d_t \hat{\mu}_1 - (1 - d_t) \hat{\mu}_2)) \\ \frac{1}{T} \sum_{t=1}^T (\Delta y_t / \hat{\psi} - \theta \Delta k_t - (1 - \theta) \Delta l_t - \hat{\mu}_1) d_t \\ \frac{1}{T} \sum_{t=1}^T (\Delta y_t / \hat{\psi} - \theta \Delta k_t - (1 - \theta) \Delta l_t - \hat{\mu}_2) (1 - d_t) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

where  $\hat{\chi} \equiv (\hat{\psi}, \hat{\mu}_1, \hat{\mu}_2)$ , and  $d_t$  is a dummy variable which takes on a value of one for all  $t < T'$  and a value of zero otherwise. The specification of the sample standard errors for this drift break specification are given in the Appendix.

An alternative specification would be to assume that the current drift term is known, but that the future term is not. In particular, consider the two-state specification:

$$\Pr(\mu_{t+1} = \mu_i | \mu_t = \mu_j) = \alpha_{ij}$$

with Markov transition matrix:

$\alpha_{11}$	$\alpha_{12}$
$\alpha_{21}$	$\alpha_{22}$

If we assume that  $\alpha_{ii}$  and  $\alpha_{jj}$  are near 1, and the  $\mu_j$  and  $\mu_i$  are not much different, then the conditional expectation of the drift term is roughly equal to its current level. In this case the unanticipated innovation in the technology parameter is approximately equal to  $\epsilon_t$  in periods in which no change in the drift term occurs, and approximately equal to  $\epsilon_t + \mu_t - \mu_{t-1}$  in periods in which a break occurs. In this case, the first equation in the moment condition is

$$\frac{1}{T} \sum_{t=1}^T (\Delta k_t (\Delta y_t / \hat{\psi} - \theta \Delta k_t - (1 - \theta) \Delta l_t - \bar{d}_t \hat{\mu}_1 - (1 - \bar{d}_t) \hat{\mu}_2)) = 0,$$

where  $\bar{d}_t$  takes on a value of one for all  $t \leq T'$ , and zero otherwise.

## 5. Findings

Given our focus on the aggregate production function, we conduct our analysis using the standard measures of aggregate output and inputs. This also facilitates comparison of our findings with other estimates in the literature. The output data is real GDP. In our sensitivity analysis, we also examine somewhat narrower measures of aggregate output. The source for the aggregate capital stock was Fixed Reproducible Tangible Wealth, and includes residential, nonresidential, government and private business capital. The source for the aggregate labor input series was the Survey of Current Business, Total Hours Worked by Full-Time and Part-Time Employees (Table 6.9C). The sample period is over 1948–1993, and the frequency of

the data is annual<sup>14</sup>. We assume that capital's share of income ( $\theta$ ) is equal to the average of the ratio of unambiguous capital income (depreciation, dividends, corporate profits, rental income of persons, net interest) to the sum of unambiguous income of capital and labor (compensation of employees) over this period, which is approximately .30 in the data.

### A. Specification 1: Log Random Walk with Drift Process

We consider a number of alternative specifications of this stochastic process to evaluate the robustness of our findings. Our first specification is that the true productivity shock is a log random walk with drift:

$$\ln(\lambda_t) = \ln(\lambda_{t-1}) + \mu + \varepsilon_{\lambda t}$$

$$E(\varepsilon_{\lambda t}) = 0, \quad E(\varepsilon_{\lambda t})^2 = \sigma_{\varepsilon_{\lambda}}^2$$

Denoting  $k = \ln(K)$ , and  $\hat{\varepsilon}_{\lambda t}$  as the sample estimate of the disturbance term, the sample moment conditions that are used to estimate the parameters  $\psi$  and  $\mu$  are:

$$\frac{1}{T} \sum (\Delta k \hat{\varepsilon}_{\lambda t}) = 0$$

$$\frac{1}{T} \sum \hat{\varepsilon}_{\lambda t} = 0$$

This specification is a natural starting point, since it is the most commonly employed specification in the literature (Basu and Fernald, Caballero and Lyons, Baxter and King, Burnside, Eichenbaum, and Rebelo, Hall, and Evans ). The common use of this statistical model reflects two observations. First, under the assumption of constant returns to scale, it is a reasonably good characterization for the measured process (Prescott (1986)). Second, one can take a log-linear approximation to a general, homogeneous, production function that yields a log-differenced equation to be estimated that relates the percent change in output

---

<sup>14</sup>We use annual data, since the pure measurement quality of the data is probably better at this frequency than at the quarterly frequency, and also because seasonal adjustment is not an issue at the annual frequency.

to income-weighted percent changes in inputs. A key issue that is often unstated in this literature is that this log-differenced specification implicitly assumes a log-shock process that is serially uncorrelated in first differences. The log random walk with drift process satisfies this restriction.

In this specification, the relevant moment condition is to set the covariance between the *percentage change in capital* and the innovation to the technology shock to zero. This is useful, because if the capital stock has persistent measurement error, as would be the case if depreciation rates are measured with error, the log-differenced specification would tend to eliminate some of that mis-measurement. Thus, under the assumption of persistent measurement error, the measurement of log-differenced capital will be more accurate than the measurement of the raw capital stock<sup>15</sup>.

To gain insight into how our moment procedure works, we present in Figure 1 the correlation between the percentage change in capital, and the innovation to the technology shock for values of returns to scale ranging from 0.1 to 2.0 over the full postwar sample. This figure reveals a number of interesting features that provide intuition about our technique, and that also shed new light on externality values that have been used in the recent literature. Given our theory, the covariance between (de-measured) log-differenced capital and the implied innovation to the technology shock at any conjectured  $\hat{\psi}$  is:

$$\frac{1}{T} \sum (\Delta \ln(K) \left[ \frac{\psi}{\hat{\psi}} \varepsilon_\lambda + \frac{\psi - \hat{\psi}}{\hat{\psi}} (\theta \Delta \ln(K) + (1 - \theta) \Delta \ln(L)) \right])$$

First, this covariance is zero *only* for the true externality parameter,  $\psi = \hat{\psi}$ , which occurs in U.S. data near 1 (.984). Second, our theory predicts that the covariance between these variables should be: (1) positive for  $\hat{\psi} < \psi$ , (2) negative for  $\hat{\psi} > \psi$ , and (3) that this covariance should decline monotonically as  $\hat{\psi}$  rises<sup>16</sup>. These predictions from our theory are exactly consistent with postwar U.S. observations: the correlation is positive for values of

<sup>15</sup>We discuss measurement error in detail in section 6.

<sup>16</sup>These implications of our theory hold provided that  $\theta \text{var}(\Delta \ln(K)) + (1 - \theta) \text{cov}(\Delta \ln(K), \Delta \ln(L)) > 0$ , which is satisfied in the data. If this expression was negative, rather than positive, monotonicity of the covariance would still hold, but it would fall as  $\hat{\psi}$  rises, rather than decline. If this expression was exactly equal to 0, which is not the case, then identification fails.

$\hat{\psi} < .984$ , and is negative for values of  $\hat{\psi} > .984$ .

This figure shows clearly the dependence between the percentage change in capital and the estimated innovation to the technology shock for the different values of  $\psi$ . In addition, this figure indicates that estimates of the externality parameter that have been used in earlier literature (e.g. Benhabib and Farmer), such as 1.4, are at variance with the aggregate data. At  $\psi = 1.4$ , the correlation between the percent change in capital and the estimated innovation to the technology shock is -.37.

The second column of Table 2a presents the Method of Moments estimates and standard errors for this basic case. As noted above, the estimate of returns to scale is .984, with a standard error of .155. The estimated drift in the log random walk is .01, with standard error of .004. It turns out that the estimates are somewhat sensitive to the first two observations in the sample. The third column of Table 2 presents estimates over the 1950-1993 period. The estimate of returns to scale rises from just under 1, to 1.13. The estimate of the drift term is not much affected. As we will show below, this higher estimate of returns to scale is very sensitive to the assumption that there is no change in the stochastic process for the technology shock.

### **The Effect of the Productivity Slowdown on Estimated Returns to Scale**

Our analysis involves estimating the unpredictable component of the technology shock. As we noted earlier, it is well known that productivity growth in the U.S. has slowed considerably since the early 1970s. This is seen clearly in Figure 2, which displays total factor productivity (under the assumption of constant returns to scale) over the entire postwar period. The average growth rate of TFP between 1948 and 1972 is 1.5 percent, while the average growth rate since 1973 has been only 0.6 percent. Within the context of our model, this observation suggests that the stochastic process should be specified to accommodate a trend (drift) break at the time of the productivity slowdown. Therefore, we consider an alternative specification for the technology shock:

$$\ln(\lambda_t) = \ln(\lambda_{t-1}) + \mu_1 D_t + \mu_2(1 - D_t) + \varepsilon_{\lambda t}$$

where  $D_t$  is a dummy variable that is unity prior to the break, and is zero otherwise. We

analyze break dates in two ways. First, given the prevailing view that there was a technology slowdown in the early 70's, we allow for a drift break beginning in 1973.<sup>17</sup> Second, we use the critical values constructed by Andrews (1993), search over all possible break dates, and choose the break on the basis of the date of the maximum test statistic. We find that this occurs with a break after 1965. Therefore, we estimated returns to scale, allowing for a single break to occur (i) beginning in 1966, or (ii), beginning in 1973.

The estimates for the drift break specifications for both the full sample and the restricted sample are presented in Tables 2b and 2c. The final row of table 2b reports the J-test statistic for no break. For both the full and sub-samples with the break occurring after 1965, estimates of returns to scale are near 1, and the standard errors are somewhat lower than under the single drift specification. We also find in each sample that there is significant difference in the estimates of the two drift terms. If the break is assumed to have occurred after 1972, we find a full sample estimate of returns to scale of .81, while the sub-sample estimate is .91. Thus, there is no point-estimate evidence in favor of increasing returns when a break is allowed in the drift term of the technology shock process.

## **B. Specification 2: Trend Stationary with I.I.D. Shocks**

An alternative process that we consider is a trend stationary process with i.i.d. shocks. Unlike the standard real business cycle model, this specification features serially uncorrelated shocks to technology. This process implies that the serial correlation in measures of the shock obtained by Prescott (1986) is an artifact of assuming constant returns to scale. Briefly, this specification of the shock process is not supported by the data. In figure 5, we plot the correlation of the innovation with detrended capital. We find that at no reasonable estimate of returns to scale is first-order serial correlation below .75, and moreover, that the correlation between detrended capital and the innovation is zero only if returns to scale are around 4, a highly implausible number.

---

<sup>17</sup>The findings aren't sensitive to the precise year that we choose the break in the early 1970's.

### C. Specification 3: Trend Stationary Autoregressive Process

We next consider a specification that is a trend-stationary autoregressive process in logs:

$$\ln(\lambda_t) = \rho \ln(\lambda_{t-1}) + \gamma t + \mu + \varepsilon_{\lambda t}$$

where  $\rho$  is the serial correlation coefficient,  $\gamma$  is the deterministic growth rate, and  $\mu$  is the constant term. Given the uncertainty over whether aggregate time series are best represented as an integrated process, or a persistent trend stationary process, this specification is a natural alternative to our first model for the shock process. The parameters  $\gamma$ ,  $\rho$ ,  $\mu$ , and  $\psi$  can be estimated using the method of moments.

$$\frac{1}{T} \sum (1 - \rho L) \tilde{k}_t \hat{\varepsilon}_{\lambda t}$$

$$\frac{1}{T} \sum \hat{\varepsilon}_{\lambda t} = 0$$

$$\frac{1}{T} \sum \tilde{\lambda}_{t-1} \hat{\varepsilon}_{\lambda t}$$

$$\frac{1}{T} \sum t \hat{\varepsilon}_{\lambda t}$$

The findings from the trend-stationary AR(1) specification with no trend break are presented in Table 3a. The second and third columns of this table presents the findings with a single trend coefficient for the full sample and the 1950-1993 sample. In both samples, estimated returns to scale is about 1.1, and the serial correlation is above .8, with a small standard error. We also conduct the analysis with a trend break. When we again follow Andrews's procedure, we find the maximum break statistic to occur beginning 1964. We therefore estimated the model allowing for a break date (i) beginning in 1964 and (ii)

beginning in 1973.

Using these two break dates, we find, as in the integrated case, that allowing for a trend break eliminates estimates of increasing returns to scale. For both break dates, and for both samples, we find returns to scale between .97 and .99. Moreover, the amount of serial correlation in the shock falls considerably.

## Discussion

Why do estimated returns to scale fall when we allow for a trend break in the technology shock process? To understand this finding, consider the moment condition in the integrated case that restricts the covariance of log-differenced capital and the innovation to the shock to be 0:

$$\frac{1}{T} \sum \left\{ \Delta k \left( \frac{1}{\psi} \Delta y - (\theta \Delta k + (1 - \theta)l) \right) \right\} = 0 \quad (5)$$

This expression indicates how the estimate of returns to scale,  $\psi$ , depends on (1) the variance of capital, (2) the covariance between capital and output, and (3) the covariance between capital and labor. We find that the difference in the estimate of returns to scale between the single drift specification and the drift break specification is primarily due to differences in the covariance between capital and labor under these two specifications. In particular, the growth rate of capital tends to be high prior to 1973, and low afterwards (2.1 percent prior to 1973, and 1.3 percent after). However, the growth rate of per-capita labor has the *opposite* pattern: it is low prior to 1973, and higher afterwards (no growth before 1973, 0.4 percent after.) Given these different patterns in growth rates between these series, this implies that the covariance between these two variables will be biased downward if the means of the growth rates are restricted to be the same over the entire sample. Of course, this restriction is imposed in the specification of the shock process with a constant drift parameter. Moreover, if this covariance is biased downward, we see from (5) that it will tend to *increase* the returns to scale parameter,  $\psi$ .

We in fact find that the covariance between capital and labor is significantly lower in the single drift specification for the 1973 break date. Table 4 presents a correlation matrix for the variables  $\Delta y$ ,  $\Delta k$ , and  $\Delta l$ . Note that the correlation between capital and labor is

.46 conditional on a single mean, but is .56 with a mean break in 1973. This issue is also important for the time trend specification. With the single trend, the correlation between capital and labor is only -.1, while allowing for a trend break in 1973 results in a correlation of .72! Given these different trends, specifying the stochastic process to accommodate for a break in the drift term at the start of the productivity slowdown is important.

#### D. Specification 4: Vector Shock Processes

Our findings based on the log-differenced and trend stationary specifications of the technology shock indicate constant returns to scale. Of course, there are other alternative specifications for the shock process. Among these alternative processes are those which include other variables in the processes.

Hall (1990) and Evans (1992) find that the Solow residual is correlated with other macroeconomic variables, such as government spending. Two possible interpretations of this finding are as follows. First, there could be a component of the technology shock that is forecasted by government policy variables. Thus, rational agents would use that information when making investment decisions, which implies that their forecasting equation for the technology shock would include those variables that are useful in prediction. A second interpretation of this finding is the argument put forward by Hansen and Prescott (1993), who interpret implied residuals from production function broadly. In particular, they suggest that variations in implied technology shocks reflect not only “true” advances in technical progress, but also are functions of *any* exogenous influence that affects the efficiency of an organization’s ability to produce final output. This observation suggests that changes in government regulation, such as OSHA and EPA rules, will affect computed productivity residuals.

We pursue this empirically by specifying a shock process in which the technology shock depends on government policy. For the integrated case, we consider the following process (with a drift break):

$$\ln(\lambda_t) = \ln(\lambda_{t-1}) + \mu_1 D_t + \mu_2 (1 - D_t) + \sum_{i=1}^n \alpha_i g_{it-1} + \varepsilon_{\lambda t}$$

where  $g$  is a vector of government policy variables (specified as percent changes). The variables

we consider are government spending, labor tax rates, capital tax rates, and the monetary base. All variables are lagged one period. These variables are chosen since we have consistent measures available over the entire postwar period, and because other authors have found some of these variables to be correlated with the Solow residual.<sup>18</sup>

Table 5 presents our findings for the vector integrated case. Our finding of constant returns to scale is not changed when we include these additional variables in the technology shock process for the integrated case. We find that returns to scale are .99 and .95 for the entire sample and the sub-sample with the trend break specification. Table 6 presents the results for the trend stationary case. These findings are also similar to our earlier results, with an estimate of 1.12 for the full sample, and 1.02 for the 1950-1993 sub-sample.

### Alternative Instruments

Our analysis uses a single instrument, capital. We also have considered three other instruments that seem reasonable: lagged capital, labor input, lagged two years, and output, lagged two years. We do not consider labor input or output lagged one just year, since with time averaged data, it is likely that the informational assumption we require in our analysis will not hold. This is not an issue with capital, since the planning horizon is quite long. We find that the incremental  $R^2$  from adding these three other variables is very low, which implies that they add little in the way of explanatory power beyond the contemporaneous capital stock. The adjusted  $R^2$  are typically negative. This also suggests that our estimate of the innovation to the technology shock is uncorrelated with these other variables. Given recent work on the effects of irrelevant instruments in small sample IV estimation, we do not include them in the analysis.

## 6. Extensions

### Higher Order Processes

We have also considered higher order processes for the integrated and trend stationary cases, including AR(1) in log-differences, and AR(2) in the trend-stationary case. In all

---

<sup>18</sup>The capital and labor tax rates were constructed by Joines (1981), and updated by McGratten (1993).

the cases we considered, the data never support including any additional lags at the 10% marginal significance level. Moreover, we find that returns to scale are largely unaffected by including these additional lags. Since the findings are similar to those in tables 2-3, we have not presented them here.

### **Alternative Measures of Output**

We have also examined a measure of aggregate output that subtracts from total GDP (i) services from residential capital and (ii) government output. We exclude these measures, since much of government output is valued at input cost, and because there is no labor input associated with the service flow from residential capital. As inputs, we use aggregate hours of private labor, and the nonresidential capital stock. We include government capital, since it has been argued by Aschauer and others that government capital is an important input in private production. These findings for the trend stationary case are presented in Table 6a. The findings are very similar to those obtained with the broader measure of output and inputs. Returns to scale is .95 with no break, .97 if a break is included, and 1.03 if the government policy variables are included in the equation for the technology shock. We also conducted the analysis with the integrated assumption, and found that returns to scale were around .7.

#### **A. Instrument Quality**

These findings broadly suggest constant returns to scale at the aggregate level, using the standard production function and standard measures of output and inputs. In this section we extend the analysis by evaluating the overall quality of capital as an instrument. As Nelson and Startz (1992) and others have pointed out, an important issue in any finite sample IV analysis is the extent to which the instrument is correlated with the endogenous regressor(s). In both the integrated and trend stationary specifications, we find that the correlation between capital and labor is high. For example, in the integrated case with the break in 1973, the correlation between log-differenced capital and log-differenced labor is .56. Based on the work of Shea, and Burnside, Eichenbaum and Rebelo, this suggests that capital dominates the Ramey instruments along this dimension.

A second issue associated with instrumental variables is whether it is reasonable to

argue that the instrument is uncorrelated with the shocks. In our analysis, this orthogonality will be true under the assumption that the date “ $t$ ” capital stock is determined prior to the date “ $t$ ” realization of the innovation to the technology shock. In our view, this is a reasonable assumption, given the standard law of motion of the capital stock, and in particular, the substantial lead time between appropriations for investment and investment expenditures. This is one reason why we have not included consumer durables in our measure of the capital stock. Thus, in our analysis, we are exploiting the *pre-determined* nature of the capital stock as an instrument, rather than choose a contemporaneous variable, and arguing that it is *exogenous*, as is done in much of the other literature. Although other lagged variables could be used as instruments, we found that they did not add much incremental explanatory power.

Another potential issue is measurement error in the capital stock. For example, if depreciation is measured with error, the capital stock will tend to have persistent measurement error. Although this type of measurement error would be a problem in the *levels* of the capital stock, first-differencing the data may tend to reduce its influence. To gain more insight into this issue, suppose that the measured percent change in capital is equal to the true percent change plus classical measurement error:

$$\Delta k_t = \Delta k_t^* + \eta_t$$

$$\text{var}(\eta) = \sigma_\eta^2$$

$$\text{cov}(\Delta k^*, \eta) = 0$$

Classical measurement error implies that the variance of measured capital is higher than the variance of true capital. Defining  $r$  to be the correlation coefficient, this implies that the correlation between true capital and labor is higher than the correlation between measured capital and labor:

$$r(\Delta k^*, \Delta l) > r(\Delta k, \Delta l)$$

$$r(\Delta k, \Delta l) = .56$$

To understand how big measurement error in the percent change in capital might be, and how it would affect our estimates of returns to scale, we conduct the following experiment. Under the assumption of classical measurement error, the correlation between capital and labor of .56 is biased downward. Since the bias is a function of the variance of the measurement error, we can estimate the variance of the measurement error for any conjectured correlation between true capital and labor. Therefore, we conduct a sensitivity analysis by assuming different values for the correlation between true capital and labor. We use this assumed correlation and the sample correlation between capital and labor to estimate the variance of the measurement error, and then re-calculate returns to scale accounting for measurement error.

We consider two alternatives to our estimated value of .56. The first is .64, which is the sample correlation between the percent change in labor and the percent change in investment. Assuming this value as the correlation between the true percent change in capital and the percent change in labor, we find our estimates of returns to scale in the integrated drift break specification rise from .91 to about .94. The second value we consider is .74, which is the correlation between the percent change in labor and the percent change in output. Given that the capital stock is the sum of current and all past levels of investment, it would seem reasonable to assume that the correlation between true capital and labor is not higher than the correlation between labor and output. Using this value, we find returns to scale rise from .91 to about 1.01. This analysis suggests that our finding of constant returns is robust to allowing for plausible levels of classical measurement error.

## **B. Capital and Labor Service Variation**

In our analysis, we have used measured capital and labor as inputs. Given the importance that the recent literature has placed on unobserved service variation in analyzing returns to scale, it is natural to ask how abstracting from this issue might effect our estimates. We consider two simple cases. In the first case, the ratio of services to the measured input is a

function of the current technology shock. Under this assumption, we show that our estimate of returns to scale is unaffected. We then consider an alternative case in which this ratio is, in equilibrium, a function of the level of aggregate output. In this case we show that our estimate of returns to scale depends on the service fluctuation parameters.

First, assume that:

$$k_t^s/k_t = \Gamma_k(\varepsilon_\lambda)$$

$$l_t^s/l_t = \Gamma_l(\varepsilon_\lambda)$$

In this case, our estimated of  $\psi$  is unaffected by factor service variation. To see this, consider our simple integrated example. If  $\phi$  is the true value and  $\hat{\phi}$  is the value that we take to be correct in constructing the productivity series, we can derive the measured productivity series in terms of the true series as follows:

$$Y_t = \left[ \lambda_t (\Gamma_k(\lambda) K_t)^\theta (\Gamma_l(\lambda) L_t)^{1-\theta} \right]^{\frac{1}{1-\phi}}$$

Taking logs yields

$$y_t = \frac{1}{1-\phi} [\lambda_t + \Gamma(\lambda_t) + \theta k_t + (1-\theta)l_t],$$

where  $\Gamma(\lambda_t) = \theta \Gamma_k(\lambda_t) + (1-\theta) \Gamma_l(\lambda_t)$ , and  $\gamma, \lambda$  are taken to denote the logs of the original values. This implies that

$$(1 - \hat{\phi})y_t - [\theta k_t + (1 - \theta)l_t] = \frac{1 - \hat{\phi}}{1 - \phi} [\lambda_t + \Gamma(\lambda_t)] + \frac{\phi - \hat{\phi}}{1 - \phi} [\theta k_t + (1 - \theta)l_t].$$

But this is just our original formulation with  $\lambda_t + \Gamma(\lambda_t)$  being a combination of the productivity parameter and the relative services from the inputs. Hence we can still recover consistent estimates of  $\phi$  from initial capital given an assumed process for  $\lambda_t + \Gamma(\lambda_t)$ .

In the second case, we assume that the ratio of the log of capital and labor services to their measured inputs are given by

$$k_t^s - k_t = \omega_k + v_k(y_t - \eta_k t)$$

and

$$l_t^s - l_t = \omega_l + v_l(y_t - \eta_l t)$$

then, since

$$y_t = \lambda_t + \theta k_t^s + (1 - \theta)l_t^s + \phi y_t,$$

this implies that  $y_t = (\omega_k + \omega_l + \lambda_t + \theta k_t + (1 - \theta)l_t - \eta t) + (\phi + \theta v_k + (1 - \theta)v_l)y_t$ , where  $\eta = \eta_k + \eta_l$ .

Thus, if we do not account for service variation, our estimate of the aggregate externality parameter is actually equal to

$$\psi = \frac{1}{1 - (\phi + \theta v_k + (1 - \theta)v_l)}.$$

Thus we can indirectly evaluate how much the maximum variation in services from capital and labor given our assumption that  $\phi^* \geq 0$ . This is given by our estimate of  $\hat{\psi}$ .

Given that our estimates for returns to scale based on measured inputs are near, but less than one, our evidence suggests little unobserved variation in factor services.

## 7. Conclusion

Constant returns to scale is a central construct of neoclassical theory. The theoretical underpinnings of a constant-returns-to-scale aggregate production technology relies simply on either CRS at the plant level, or replication and a sufficiently small optimum firm size. Given the minimal nature of these assumptions, and the widespread use of neoclassical theory, one would be reluctant to discard the theory lightly. Previous studies argued that one must accept significant factor service variation to reconcile constant returns with the data. Some economists have argued that this finding has not resolved the size of returns to scale, since factor service variation is unobserved, and there is no generally accepted theory to guide specification of this alternative framework. As Hall has noted, “it is well established that adding a free variable to the production function will eliminate evidence of increasing returns. Research needs to turn to whether the role of the free variables...makes sense as a matter of theory.”

In this paper, we show that the stochastic version of the neoclassical (Cass-Koopmans)

growth model delivers an orthogonality condition which can be used to estimate returns to scale. Rather than finding point estimates in favor of increasing returns, we show that standard theory and conventional measures of output and inputs yield estimates of constant returns to scale at the aggregate level. Our estimates are robust to various specifications of the stochastic process for the technology shock, to an alternative measure of output that abstracts from government product and service flows from residential capital, and to plausible measurement error. Our estimates also suggest that factor service variation (at the annual frequency) may not be an important determinant of output fluctuations.

## References

- [1] Aiyagari, R. (1994) "On the Contribution of Technology Shocks to Business Cycles", *Quarterly Review of the Federal Reserve Bank of Minneapolis*, 18 (1), 22-34.
- [2] Basu, S. (1995) "Procyclical Productivity: Increasing Returns or Cyclical Utilization?", NBER Working Paper # 5336.
- [3] Basu, S. and J. Fernald (1994) "Constant Returns and Small Markups in U.S. Manufacturing", Discussion Paper, Board of Governors of the Federal Reserve System.
- [4] Baxter, M. and R. King (1992) "Productive Externalities and Business Cycles", Federal Reserve Bank of Minneapolis Working Paper # 53.
- [5] Benhabib, J. and R. Farmer (1994) "Indeterminacy and Increasing Returns", *Journal of Economic Theory*, 63 (1) 19-41.
- [6] Burnside, Eichenbaum, and Rebelo (1995) "Capital Utilization and Returns to Scale", 1995 NBER Macroeconomics Annual.
- [7] Caballero, R. and R. Lyons (1992) "External Effects in U.S. Procyclical Productivity", *Journal of Monetary Economics*, 29 (2), 209-226.
- [8] Christiano, L. and W. den Haan (1995) "Small Sample Properties of GMM for Business Cycle Analysis", Federal Reserve Bank of Minneapolis Staff Report # 199.

- [9] Evans, C. (1992) "Productivity Shocks and Real Business Cycles", *Journal of Monetary Economics*, 29 (2), 191-208.
- [10] Hall, R. (1988) "The Relationship between Price and Marginal Cost in U.S. Industry," *Journal of Political Economy*, vol 114, 921-947.
- [11] Hall, R. (1990) "Invariance Properties of Solow's Productivity Residual", in *Growth, Productivity, Employment*, Peter Diamond, ed. MIT Press, Cambridge.
- [12] Hall, R. (1995) "Comment", 1995 NBER Macroeconomics Annual.
- [13] Hansen, L. (1982) "Large Sample Properties of Generalized Method of Moments Estimators", *Econometrica*, 50, 1269-1286.
- [14] Joines, D. (1981) "Estimates of Effective Tax Rates on Factor Incomes", *Journal of Business*, 54, 191-226.
- [15] Lucas, R. (1988) "On The Mechanics of Economic Development", *Journal of Monetary Economics*, 22, 1-42.
- [16] McGratten, E. (1994) "The Macroeconomic Effects of Distortionary Taxation", *Journal of Monetary Economics*, 33, 573-601.
- [17] Prescott, E. (1986) "Theory Ahead of Business Cycle Measurement", *Quarterly Review of the Federal Reserve Bank of Minneapolis*, 10 (1), 9-22.
- [18] Romer, P. (1986) "Increasing Returns and Long Run Growth", *Journal of Political Economy*, 94, 1002-1073.
- [19] Shaprio, M. (1993) "Cyclical Productivity and the Workweek of Capital," *American Economic Review*, vol. 83 no. 2, 229-233.

## 8. Appendix

### A. Estimation with Drift Break:

$$g(\hat{\chi}; Z_T) = \begin{bmatrix} \frac{1}{T} \sum_{t=1}^T (\Delta k_t (\Delta y_t / \hat{\psi} - \theta \Delta k_t - (1 - \theta) \Delta l_t - d_t \hat{\mu}_1 - (1 - d_t) \hat{\mu}_2)) \\ \frac{1}{T} \sum_{t=1}^T (\Delta y_t / \hat{\psi} - \theta \Delta k_t - (1 - \theta) \Delta l_t - \hat{\mu}_1) d_t \\ \frac{1}{T} \sum_{t=1}^T (\Delta y_t / \hat{\psi} - \theta \Delta k_t - (1 - \theta) \Delta l_t - \hat{\mu}_2) (1 - d_t) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Our sample estimate of  $S$ ,  $\hat{S}$ , is given by

$$\hat{S} = \begin{bmatrix} \frac{1}{T} \sum (\Delta k_t \hat{\varepsilon}_t^\lambda)^2 & \frac{1}{T} \sum \Delta k_t (\hat{\varepsilon}_t^\lambda)^2 d_t & \frac{1}{T} \sum \Delta k_t (\hat{\varepsilon}_t^\lambda)^2 (1 - d_t) \\ \frac{1}{T} \sum \Delta k_t (\hat{\varepsilon}_t^\lambda)^2 d_t & \frac{1}{T} \sum (\hat{\varepsilon}_t^\lambda)^2 d_t & 0 \\ \frac{1}{T} \sum \Delta k_t (\hat{\varepsilon}_t^\lambda)^2 (1 - d_t) & 0 & \frac{1}{T} \sum (\hat{\varepsilon}_t^\lambda)^2 (1 - d_t) \end{bmatrix}$$

And,

$$\frac{dg(\hat{\chi}; Z_T)}{d\hat{\chi}} = \begin{bmatrix} -\frac{1}{T} \sum \Delta y_t \Delta k_t / \hat{\psi}^2 & -\frac{1}{T} \sum \Delta y_t d_t / \hat{\psi}^2 & -\frac{1}{T} \sum \Delta y_t (1 - d_t) / \hat{\psi}^2 \\ -\frac{1}{T} \sum d_t \Delta k_t & -N/T & 0 \\ -\frac{1}{T} \sum (1 - d_t) \Delta k_t & 0 & -(T - N)/T \end{bmatrix}$$

The sample estimate for the standard error for  $\hat{\psi}$  is given by

$$\sqrt{\left( \left[ \frac{dg(\hat{\chi}; Z_T)}{d\hat{\chi}} \right] \hat{S}^{-1} \left[ \frac{dg(\hat{\chi}; Z_T)}{d\hat{\chi}} \right]' \right)^{-1}}$$

### B. Estimation AR(1) Specification:

Assume that the stochastic process generating the technology parameter was given by

$$\lambda_t = \rho \lambda_{t-1} + \gamma t + \tilde{\mu} + \varepsilon_t,$$

where  $0 \leq \rho < 1$ . In this case, if we first detrend the inputs and outputs, then

$$\lambda_t = y_t / \psi - \theta k_t - (1 - \theta) l_t,$$

and

$$\epsilon_t = (1 - \rho L)(\lambda_t - \mu),$$

where in an abuse of notation, we have used  $\lambda$ ,  $y$ ,  $k$  and  $l$  to denote the detrended values of these variables. (Note that simply demeaning the inputs and outputs does not ensure that the  $\epsilon_t$ 's are mean zero since we are dropping at least one observation and hence we retain the constant  $\mu$ .) Then, our moment conditions are

$$E\{(1 - \rho L)k_t\epsilon_t\} = 0,$$

$$E\{\epsilon_t\lambda_{t-1}\} = 0.$$

$$E\{\epsilon_t\} = 0$$

Consequently,

$$\frac{dg(\hat{\chi}; Z_T)}{d\hat{\psi}} = \begin{bmatrix} -\frac{1}{T} \sum [(1 - \rho L)k_t(1 - \rho L)y_t/\hat{\psi}^2] \\ -\frac{1}{T} \sum [(1 - \rho L)y_t\lambda_{t-1}/\hat{\psi}^2 + \epsilon_t y_{t-1}/\hat{\psi}^2] \\ -\frac{1}{T} \sum [(1 - \rho L)y_t/\hat{\psi}^2] \end{bmatrix}'.$$

$$\frac{dg(\hat{\chi}; Z_T)}{d\hat{\rho}} = \begin{bmatrix} -\frac{1}{T} \sum [k_{t-1}\epsilon_t + (1 - \rho L)k_t(\lambda_{t-1} - \mu)] \\ -\frac{1}{T} \sum (\lambda_{t-1} - \mu)\lambda_{t-1} \\ -\frac{1}{T} \sum (\lambda_{t-1} - \mu) \end{bmatrix}'.$$

$$\frac{dg(\hat{\chi}; Z_T)}{d\hat{\mu}} = \begin{bmatrix} -\frac{1}{T} \sum (1 - \rho L)k_t(1 - \rho) \\ -\frac{1}{T} \sum \lambda_{t-1}(1 - \rho) \\ -(1 - \rho) \end{bmatrix}'.$$

The  $\hat{S}$  matrix is once again given by the covariance matrix of the sample moment conditions.

**Table 1: Univariate Log Random Walk with Drift Specification****Table 1a: Single Drift**

parameters/cases	Full Sample 1948-1993	Sub Sample 1950-1993
$\psi$ (returns to scale)	0.984 (.155)	1.132 (.158)
$\mu$ (drift)	0.011 (.004)	0.008 (.003)

**Table 1b: Drift Break 1965/66**

parameters/cases	Full Sample 1948-1993	Sub Sample 1950-1993
$\psi$ (returns to scale)	0.940 (.126)	1.041 (.145)
$\mu_1$ (drift)	0.018 (.005)	0.014 (.004)
$\mu_2$ (second drift)	0.008 (.005)	0.006 (.003)
$J$ statistic	7.62	7.05

**Table 1c: Drift Break 1972/73**

parameters/cases	Full Sample 1948-1993	Sub Sample 1950-1993
$\psi$ (returns to scale)	0.811 (.127)	0.912 (.143)
$\mu_1$ (drift)	0.020 (.006)	0.015 (.004)
$\mu_2$ (second drift)	0.010 (.004)	0.008 (.004)
$J$ statistic	4.86	4.60

**Table 2: Univariate Trend Stationary AR(1) Specification****Table 2a: Single Trend**

parameters/cases	Full Sample 1948-1993	Sub Sample 1950-1993
$\psi$ (returns to scale)	1.120 (.076)	1.097 (.081)
$\rho$ (AR coefficient)	0.846 (.044)	0.881 (.056)
$\gamma$ (time trend)	0.001 (.0004)	0.001 (.0005)
$\mu$ (constant)	0.228 (.063)	0.185 (.090)

**Table 2b: Trend Break 1963/64**

parameters/cases	Full Sample 1948-1993	Sub Sample 1950-1993
$\psi$ (returns to scale)	0.987 (.134)	0.983 (.136)
$\rho$ (AR coefficient)	0.557 (.092)	0.485 (.109)
$\gamma_1$ (time trend)	0.005 (.001)	0.006 (.002)
$\gamma_2$ (time trend)	0.003 (.001)	0.004 (.002)
$\mu_1$ (constant)	0.731 (.216)	0.867 (.275)
$\mu_2$ (constant)	0.784 (.226)	0.920 (.284)
$J$ statistic	19.88	23.18

**Table 2c: Trend Break 1972/73**

parameters/cases	Full Sample 1948-1993	Sub Sample 1950-1993
$\psi$ (returns to scale)	0.998 (.100)	0.973 (.103)
$\rho$ (AR coefficient)	0.616 (.095)	0.612 (.105)
$\gamma_1$ (time trend)	0.005 (.002)	0.005 (.002)
$\gamma_2$ (time trend)	0.003 (.001)	0.003 (.001)
$\mu_1$ (constant)	0.625 (.199)	0.663 (.245)
$\mu_2$ (constant)	0.666 (.218)	0.701 (.264)

**Table 3: Correlation Matrix of Inputs and Output 1948-1993**

**Table 3a: First Differenced with Single Mean**

	$\Delta y$	$\Delta k$	$\Delta l$
$\Delta y$	1		
$\Delta k$	0.43	1	
$\Delta l$	0.89	0.46	1

**Table 3b: First Differenced with Mean Break in 1972/73**

	$\Delta y$	$\Delta k$	$\Delta l$
$\Delta y$	1		
$\Delta k$	0.40	1	
$\Delta l$	0.92	0.56	1

**Table 3c: Detrended Levels**

	$y$	$k$	$l$
$y$	1		
$k$	0.63	1	
$l$	0.43	-0.1	1

**Table 3d: Detrended Levels with Trend Break in 1972/73**

	$y$	$k$	$l$
$y$	1		
$k$	0.57	1	
$l$	0.95	0.72	1

Table 4: Vector Integrated Process Specification<sup>19</sup>

parameters	Full Sample 1948-1993	Sub Sample 1950-1993
$\psi$ (returns to scale)	0.987 (.134)	0.954 (.174)
$\mu$ (drift)	0.016 (.004)	0.016 (.005)
$\alpha_g$ (govt. spending)	0.058 (.036)	0.061 (.040)
$\alpha_{\tau_k}$ (capital tax)	0.000 (.001)	0.000 (.001)
$\alpha_{\tau_l}$ (labor tax)	-0.005 (.004)	-0.007 (.004)
$\alpha_{slb}$ (St. Louis Base)	-0.131 (.052)	-0.106 (.048)

Table 5: Vector Trend Stationary  $AR(1)$  Specification<sup>20</sup>

parameters/cases	Single Trend	Break 1963/64
$\psi$ (returns to scale)	1.123 (.075)	1.022 (.087)
$\rho$ ( $AR$ coefficient)	0.544 (.120)	0.410 (.139)
$\gamma_1$ (time trend)	0.003 (.001)	0.004 (.002)
$\gamma_2$ (time trend)		0.005 (.002)
$\mu_1$ (constant)	0.645 (.162)	0.955 (.266)
$\mu_2$ (constant)		0.963 (.276)
$\alpha_g$ (govt. spending)	0.016 (.024)	0.004 (.023)
$\alpha_{\tau_k}$ (capital tax)	-0.009 (.021)	-0.003 (.029)
$\alpha_{\tau_l}$ (labor tax)	-0.027 (.024)	-0.038 (.028)
$\alpha_{slb}$ (St. Louis Base)	-0.053 (.015)	-0.066 (.046)

<sup>19</sup>For the case in which the sample was 1950-1993, the  $\bar{R}^2$  from regressing the errors on lagged first differenced capital and twice lagged first differenced output and labor was  $-.02$ .

<sup>20</sup>The sample was 1950-1993, though the results were essentially unchanged over the whole sample. The  $\bar{R}^2$  from regressing the errors on lagged detrended capital and twice lagged detrended output and labor was  $-.05$  and  $-.004$  respectively.

**Table 6: Nonresidential Private Output, Private Labor and Nonresidential Capital Stock**

**Table 6a: Trend Stationary AR(1) Specification<sup>21</sup>**

parameters	Univariate		Break in 1968/69		Vector	
$\psi$ (returns to scale)	0.948	(.133)	0.969	(.110)	1.030	(.119)
$\rho$ (AR coefficient)	0.727	(.102)	0.535	(.103)	0.517	(.126)
$\gamma$ (time trend)	0.002	(.001)	0.006	(.001)	0.004	(.001)
$\gamma_2$ (time trend)			0.004	(.001)		
$\mu$ (constant)	0.517	(.247)	0.830	(.201)	0.820	(.221)
$\mu_2$ (constant)			0.855	(.208)		
$\alpha_g$ (govt. spending)					-0.005	(.040)
$\alpha_{\tau_k}$ (capital tax)					-0.020	(.032)
$\alpha_{\tau_l}$ (labor tax)					-0.060	(.040)
$\alpha_{slb}$ (St. Louis Base)					-0.044	(.014)

**Table 6b: Correlations Detrended Levels**

	$y$	$k$	$l$
$y$	1		
$k$	0.32	1	
$l$	0.57	-0.7	1

**Table 6c: Correlations Detrended Levels with Trend Break in 1968/69**

	$y$	$k$	$l$
$y$	1		
$k$	0.58	1	
$l$	0.94	0.59	1

<sup>21</sup>The sample was 1950-1993. The  $\bar{R}^2$  from regressing the errors on lagged detrended capital and twice lagged detrended output and labor was .01 and -.01 for the trend break and vector cases. The maximum  $J$  statistic on the trend break was in 1968/69 with a value of 18.85. With private nonresidential capital the returns-to-scale estimates are somewhat lower; 0.739 for the univariate case and 0.917 for the vector specification.

Figure 1: Total factor Productivity with Electricity as Measure of Capital Services

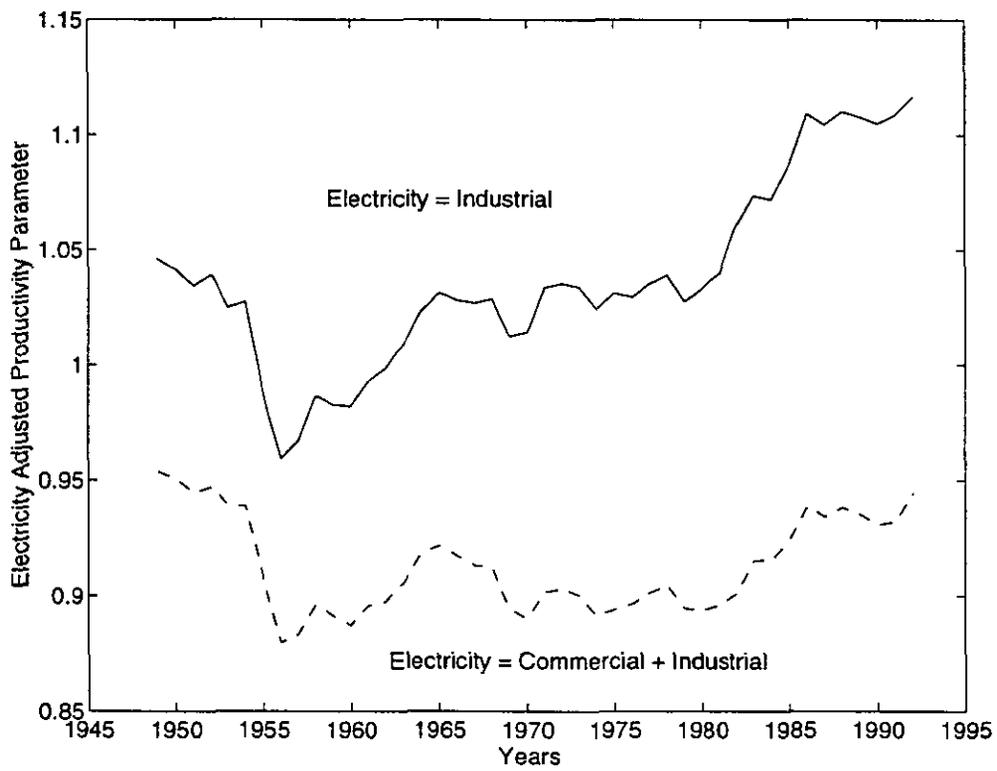


Figure 2: Growth Rates of Gross Output and Materials in Manufacturing

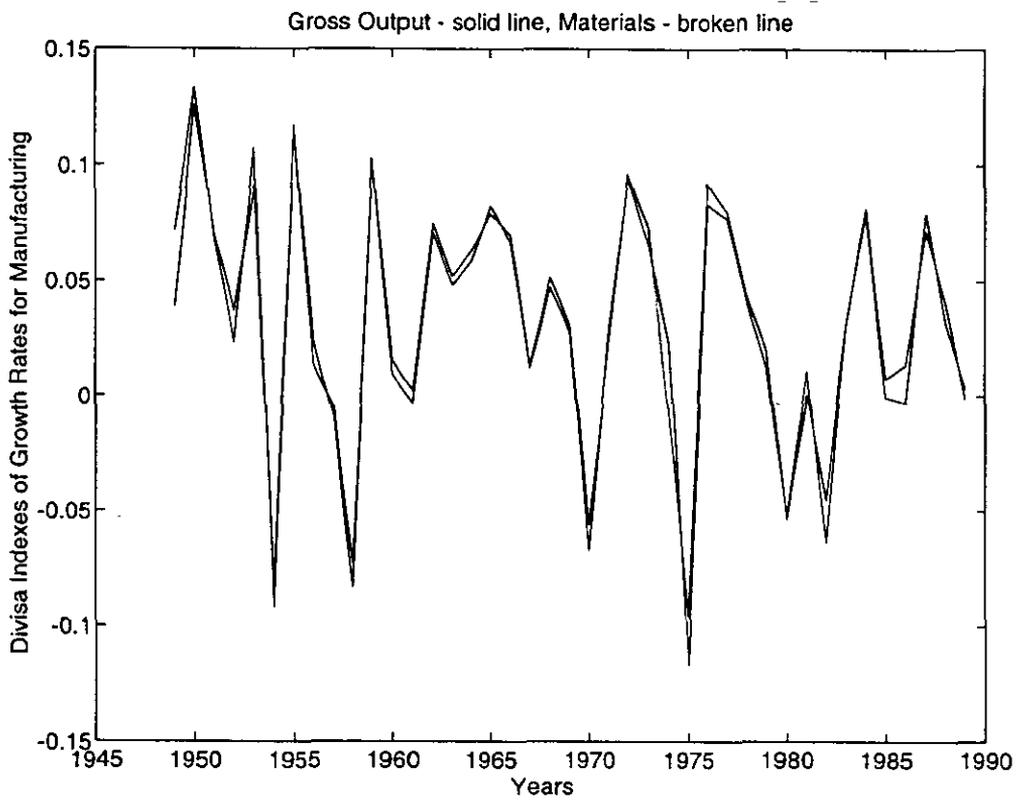


Figure 3: Random Walk with Drift Specification

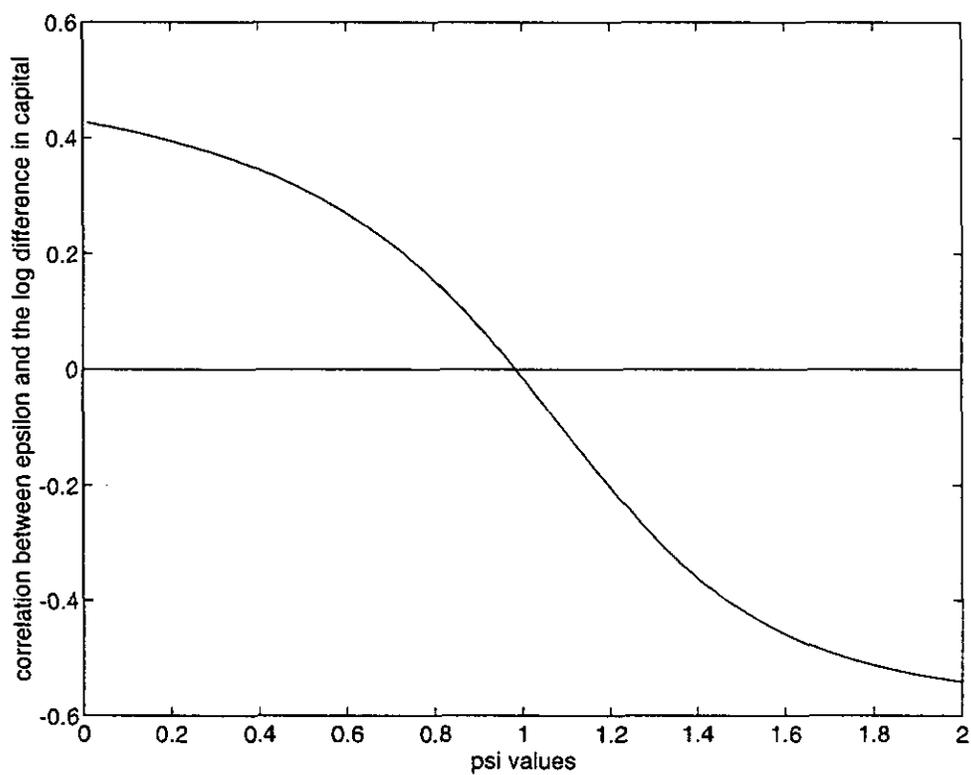


Figure 4: Total Factor Productivity

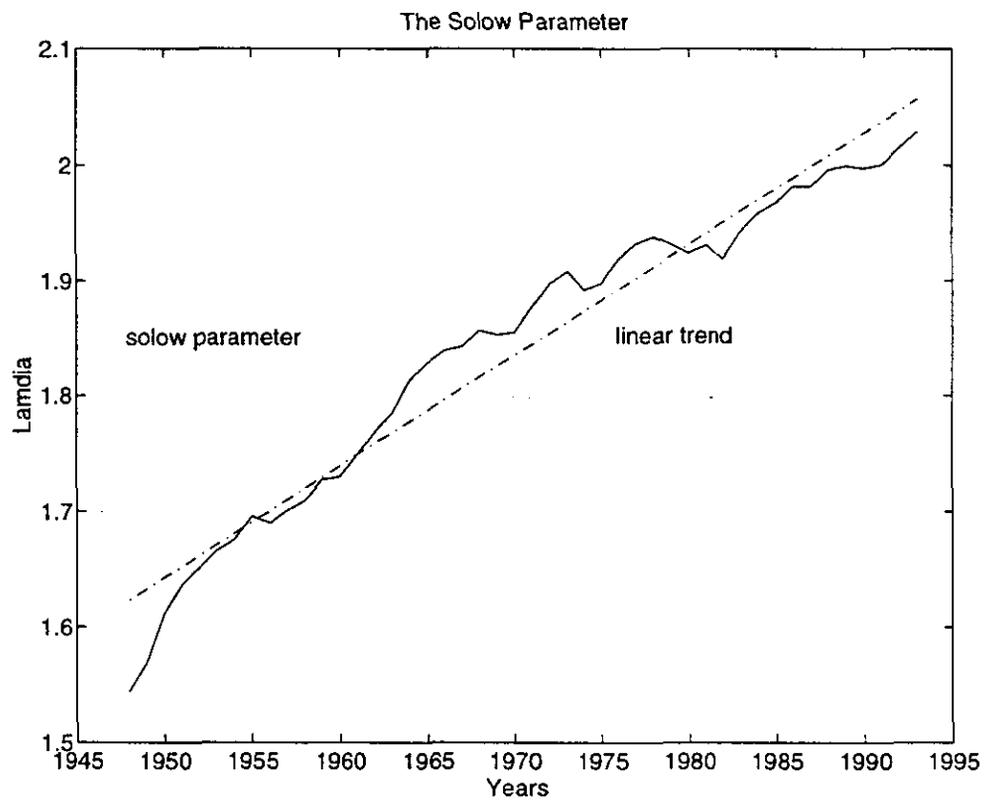


Figure 5: Correlations with Trend and I.I.D. Shocks

