

# *Contributions to Macroeconomics*

---

*Volume 4, Issue 1*

2004

*Article 4*

---

## Is the U.S. Aggregate Production Function Cobb-Douglas? New Estimates of the Elasticity of Substitution

Pol Antràs\*

\*Harvard University, pantras@fas.harvard.edu

Copyright ©2004 by the authors. All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, without the prior written permission of the publisher, bepress. *Contributions to Macroeconomics* is one of *The B.E. Journals in Macroeconomics*, produced by The Berkeley Electronic Press (bepress). <http://www.bepress.com/bejm>.

# Is the U.S. Aggregate Production Function Cobb-Douglas? New Estimates of the Elasticity of Substitution\*

Pol Antràs

## **Abstract**

I present new estimates of the elasticity of substitution between capital and labor using data from the private sector of the U.S. economy for the period 1948-1998. I first adopt Berndt's (1976) specification, which assumes that technological change is Hicks neutral. Consistently with his results, I estimate elasticities of substitution that are not significantly different from one. I next show, however, that restricting the analysis to Hicks-neutral technological change necessarily biases the estimates of the elasticity towards one. When I modify the econometric specification to allow for biased technical change, I obtain significantly lower estimates of the elasticity of substitution. I conclude that the U.S. economy is not well described by a Cobb-Douglas aggregate production function. I present estimates based on both classical regression analysis and time series analysis. In the process, I deal with issues related to the nonsphericity of the disturbances, the endogeneity of the regressors, and the nonstationarity of the series involved in the estimation.

**KEYWORDS:** Capital-Labor Substitution, Technological Change

---

\*Harvard University, NBER and CEPR. Email: pantras@fas.harvard.edu

# 1 Introduction

The elasticity of substitution between capital and labor is a central parameter in economic theory. Models investigating the sources of economic growth and the determinants of the aggregate distribution of income have been found to deliver substantially different implications depending on the particular value of the elasticity of substitution.

Perhaps to a larger extent than in other fields, the elasticity of substitution between capital and labor is central in growth theory, both traditional and new. On the one hand, in the framework of the Neoclassical growth model, the sustainability of long-run growth in the absence of technological change depends crucially on whether the elasticity of substitution is greater than or smaller than one.<sup>1</sup> On the other hand, the recent literature on induced technical change has developed models that deliver very different implications depending on the particular value of the elasticity of substitution. For instance, Acemoglu (2003) builds on the assumption of a lower-than-one elasticity of substitution to construct a model that rationalizes the coexistence of both purposeful labor- and capital-augmenting technological change in the transitional dynamics of an economy that converges to a balanced growth path in which technical change is purely labor-augmenting. Furthermore, as pointed out by Hsieh (2000), the value of the elasticity of substitution is also relevant to the empirical debate on the sources of economic growth (cf., Mankiw, Romer and Weil, 1992).<sup>2</sup>

In the field of public finance, the value of the elasticity of substitution constitutes an important determinant of the response of investment behavior to tax policy. The late 1960's witnessed a lively debate between those who perceived fiscal policy as an effective tool for influencing investment behavior (e.g., Hall and Jorgenson, 1967) and those who recognized only minor benefits from tax incentives (e.g., Eisner and Nadiri, 1968). Their debate revolved around the issue of whether the elasticity of substitution between capital and labor was significantly below one, with a lower elasticity being associated with

---

<sup>1</sup>If the elasticity of substitution is greater than one, the marginal product of capital remains bounded away from zero as the capital stock goes to infinity. Under certain parameter restrictions, this violation of the Inada condition can yield long-run endogenous growth even in the absence of technological progress (cf. Barro and Sala-i-Martin, 1995, p. 44).

<sup>2</sup>Following Mankiw, Romer and Weil (1992), most studies trying to disentangle the relative role of technological change and factor accumulation in explaining cross-country income differences have assumed the elasticity of substitution to be equal to one. Hsieh (2000), shows that relaxing this assumption and allowing for biased technical change may alter substantially the results of these studies.

a lower response of investment to tax benefits.<sup>3</sup>

Soon after the explicit derivation of the Constant Elasticity of Substitution (CES) production function by Arrow et al. (1961), a wealth of articles appeared trying to estimate this elasticity for the U.S. manufacturing sector.<sup>4</sup> Cross-sectional studies at the two-digit level tended to find elasticities insignificantly different from one (e.g., Dhrymes and Zarembka, 1970). Lucas (1969), however, discussed several biases inherent in the use of cross-sectional data in the estimation of the elasticity. He suggested the use of time series data instead. Time series studies generally provided much lower estimates of the elasticity. Lucas (1969) himself estimated the elasticity of substitution to be somewhere between 0.3 and 0.5, while Maddala (1965), Coen (1969) and Eisner and Nadiri (1968) also computed estimates significantly below one.

In a widely cited contribution, Berndt (1976) illustrated how the use of higher quality data translated into considerably higher time-series estimates of the elasticity, thus leading to a reconciliation of the time-series and cross-sectional studies. In particular, a careful construction of the series involved in the estimation led him to obtain time series estimates for the period 1929-1968 insignificantly different from one. It has become customary in the literature to cite Berndt's paper as providing evidence in favor of the assumption of a Cobb-Douglas functional form for the aggregate production function (e.g., Judd, 1987, Trostel, 1993).

In this paper, I will start by following closely the approach suggested by Berndt (1976), which assumes that technological change is Hicks neutral. Using time-series data from the private sector of the U.S. economy for the period 1948-1998, the first result of this paper will be a confirmation of Berndt's finding of a unit elasticity of substitution between capital and labor when technological change is assumed to be Hicks neutral. The second and more substantive contribution of this paper will consist in demonstrating that in the presence of non-neutral technological change, Berndt's approach leads to estimates of the elasticity that are necessarily biased towards one. When the econometric specification is modified to allow for biased technical change, I generally obtain significantly lower estimates of the elasticity of substitution.

The source of the bias is rather simple to illustrate. Suppose that U.S. aggregate output can be represented by a production function of the form:

$$Y_t = A_t F(K_t, L_t),$$

characterized by constant returns to scale in the two inputs, capital and labor.

---

<sup>3</sup>See Chirinko (2002) for more details.

<sup>4</sup>For a thorough review of this literature see Nerlove (1967) and Berndt (1976, 1991).

The parameter  $A_t$  is an index of technological efficiency, which is neutral in Hicks' sense, i.e., in the sense that it has no effect on the ratio of marginal products for a given capital-labor ratio. Profit maximization by firms in a competitive framework delivers two optimality conditions equalizing factor prices with their marginal products. Combining these conditions delivers

$$\frac{r_t K_t}{w_t L_t} = \frac{f'(k_t)k_t/f(k_t)}{1 - f'(k_t)k_t/f(k_t)},$$

where  $f(k)$  is output per unit of labor,  $k$  is the capital-labor ratio, and  $r$  and  $w$  are the rental prices of capital and labor, respectively. As is well-known, in the United States, the value of the left-hand side of this expression has been remarkably stable throughout the post-WWII period, while the capital-labor ratio has steadily increased. It follows that this equation can be consistent with U.S. data only if  $f'(k_t)k_t/f(k_t)$  is not a function of  $k_t$ , i.e., only if  $F(K_t, L_t)$  is a Cobb-Douglas production function.<sup>5</sup> In words, when technological change is Hicks neutral and the capital-labor ratio grows through time, the only aggregate production function consistent with constant factor shares is one featuring a unit elasticity of substitution between capital and labor. As I will discuss in section 2, the approach of Berndt (1976) consists of running log-linear specifications closely related to the expression above. In light of this discussion, his finding of a unit elasticity of substitution should not be surprising.

The main problem with Berndt's approach is that when technological change is allowed to affect the ratio of marginal products, the Cobb-Douglas production function ceases to be the only one consistent with stable factor shares. In particular, a well-known theorem in growth theory states that, in the presence of exponential labor-augmenting technological change, any well-behaved aggregate production function is consistent with a balanced growth path in which factor shares are constant. In a similar vein, Diamond, McFadden and Rodriguez (1978) formally proved the impossibility of identifying the separate roles of factor substitution and biased technological change in generating a given time series of factor shares and capital-labor ratios. The literature has generally circumvented this impossibility result by imposing some type of structure on the form of technological change. As I will discuss in section 5, when technological efficiency grows exponentially these two effects can be separated and the elasticity of substitution can be recovered from the available data. Furthermore, my empirical results below suggest that allowing for biased technological change leads to estimates of the elasticity of substitution that

---

<sup>5</sup>Solving the differential equation  $f'(k_t)k_t/f(k_t) = \alpha$  yields  $y = Ck_t^\alpha$ , where  $C$  is a constant of integration.

are, in general, significantly lower than one. I conclude from my results that the U.S. aggregate production function does *not* appear to be Cobb-Douglas.

This is not the first paper to estimate the elasticity of substitution while taking into account the presence of biased technological change. Among others, David and van de Klundert (1965) and Kalt (1978) ran regressions analogous to those in section 5 below, and estimated elasticities equal to 0.32 and 0.76, respectively. This paper adds to this literature in at least three respects. First, by explicitly discussing and correcting the bias inherent in the assumption of Hicks-neutral technological change, I am able to reconcile the traditional low estimates of Lucas (1969) and others with the widely cited ones of Berndt (1976).<sup>6</sup> Second, by focusing on a more recent period, I am able to benefit from the higher-quality data made available by the work of Herman (2000), Krusell et al. (2000), and Jorgenson and Ho (2000). Finally, my empirical analysis incorporates recent developments in the econometric analysis of time series that permit a better treatment of the nonstationary nature of the series involved in the estimation.<sup>7</sup>

The rest of the paper is organized as follows. In section 2, I follow Berndt (1976) in deriving six alternative specifications for the estimation of the elasticity of substitution under the assumption of Hicks-neutral technological change. Section 3 discusses the data used in the estimations. Section 4 presents estimates of the elasticity based on both classical econometrics techniques and modern time series analysis. Section 5 discusses the crucial misspecification in Berndt's (1976) contribution and presents estimates that correct for it by allowing for biased technical change. Section 6 concludes.

## 2 Model specification

I begin by assuming that aggregate production in the U.S. private sector can be represented by a constant returns to scale production function characterized by a constant elasticity of substitution between the two factors, capital and labor.

---

<sup>6</sup>Kalt (1978), for instance, incorrectly dismissed Berndt's results by claiming that he had estimated the elasticity "without regard to technological change" (p. 762).

<sup>7</sup>Following the dual cost function approach pioneered by Nerlove (1963) and Diewert (1971), a separate branch of the literature has provided estimates of the elasticity based on first-order conditions derived from cost minimization rather than profit maximization. For instance, Berndt and Christensen (1973) fitted a translog cost function to the U.S. manufacturing sector for the period 1929-68, obtaining elasticities of substitution between capital equipment and labor and between capital structures and labor slightly higher than one. Nevertheless, their estimates should be treated with caution because, like Berndt (1976), the authors failed to deal properly with technological change.

Arrow et al. (1961) showed that the assumption of a constant elasticity of substitution implied the following functional form for the production function:

$$Y_t = A_t \left[ \delta K_t^{\frac{\sigma-1}{\sigma}} + (1-\delta)L_t^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}},$$

where  $Y_t$  is real output,  $K_t$  is the flow of services from the real capital stock,  $L_t$  is the flow of services from production and nonproduction workers,  $A_t$  is a Hicks-neutral technological shifter,  $\delta$  is a distribution parameter, and the constant  $\sigma$  is the elasticity of substitution between capital and labor.<sup>8</sup> Following Berndt (1976), it is useful to define the aggregate input function  $F_t \equiv Y_t/A_t$ , which given the assumption of Hicks-neutral technological change is independent of  $A_t$ . Profit maximization by firms in a competitive framework implies two first-order conditions, equating real factor prices to the real value of their marginal products. These conditions can be rewritten and expanded with an error term to obtain:

$$\log(F_t/K_t) = \alpha_1 + \sigma \log(R_t/P_t) + \varepsilon_{1,t} \quad (1)$$

$$\log(F_t/L_t) = \alpha_2 + \sigma \log(W_t/P_t) + \varepsilon_{2,t}, \quad (2)$$

where  $R_t$ ,  $W_t$ , and  $P_t$  are the prices of capital services, labor services, and aggregate input  $F_t$ , respectively, and  $\alpha_1$  and  $\alpha_2$  are constants that depend on  $\delta$ .<sup>9</sup> A third alternative specification can be obtained by subtracting (1) from (2)

$$\log(K_t/L_t) = \alpha_3 + \sigma \log(W_t/R_t) + \varepsilon_{3,t}. \quad (3)$$

Following Berndt (1976) one can also rearrange equations (1) through (3) to obtain the following three reverse regressions:

$$\log(R_t/P_t) = \alpha_4 + (1/\sigma) \log(F_t/K_t) + \varepsilon_{4,t} \quad (4)$$

$$\log(W_t/P_t) = \alpha_5 + (1/\sigma) \log(F_t/L_t) + \varepsilon_{5,t} \quad (5)$$

$$\log(W_t/R_t) = \alpha_6 + (1/\sigma) \log(K_t/L_t) + \varepsilon_{6,t}. \quad (6)$$

I hereafter denote the estimates of  $\sigma$  based on equations (1) through (6) by  $\sigma_i$ ,  $i = 1, \dots, 6$ .<sup>10</sup> As pointed out by Berndt (1976), in this bivariate setting,

<sup>8</sup>The elasticity of substitution between capital and labor is defined as  $\sigma = d \log(K/L) / d \log(F_L/F_K)$ , where  $F_K$  and  $F_L$  are the marginal products of capital and labor, respectively.

<sup>9</sup>A simple way to justify these disturbance terms is to appeal to optimization errors on the part of firms (cf., Berndt, 1991, p. 454).

<sup>10</sup>In the presence of imperfect competition in the product market, the markup becomes

the following equalities will necessarily hold for the OLS estimates:

$$\frac{\sigma_1}{\sigma_4} = R_1^2 = R_4^2; \frac{\sigma_2}{\sigma_5} = R_2^2 = R_5^2; \frac{\sigma_3}{\sigma_6} = R_3^2 = R_6^2, \quad (7)$$

where  $R_i^2$  refers to the R-square in equation  $i$ . These equalities in turn imply the inequalities  $\sigma_1 \leq \sigma_4$ ,  $\sigma_2 \leq \sigma_5$  and  $\sigma_3 \leq \sigma_6$ . More importantly, it follows from (7) that the larger the R-square in the OLS regressions, the closer will the standard and reverse estimates be. It should be emphasized, however, that these results hold only for the OLS estimates.<sup>11</sup>

On the other hand, nothing can be predicted on statistical grounds about the relative size of the estimates  $\sigma_1$ ,  $\sigma_2$  and  $\sigma_3$ , although previous studies led Berndt (1976) to point out that estimates based on the marginal product of labor equation (2) seem to yield higher estimates of the elasticity of substitution than estimates based on the marginal product of capital equation (1). One could therefore expect the estimates to satisfy  $\sigma_2 > \sigma_1$ .<sup>12</sup>

### 3 Data Construction and Sources

Estimation of equations (1) through (6) requires data on the flow of labor services  $L_t$ , the nominal price of these labor services  $W_t$ , the flow of capital services  $K_t$ , the rental price of capital  $R_t$ , and the aggregate input index  $F_t$ , as well as its associated price  $P_t$ . To illustrate the effect of data quality on the estimates of the elasticity, I experiment with different methods in the construction of these variables.

I initially assume that labor services are proportional to employment and an omitted variable in equations (1), (2), (4), and (5), but equations (3) and (6) remain valid. On the other hand, in the presence of imperfect competition in the factor markets, even equations (3) and (6) may produce biased estimates if the wedge between marginal products and factor prices is different for different factors.

<sup>11</sup>As pointed out by a referee, the error terms  $\varepsilon_{i,t}$ ,  $i = 1, \dots, 6$ , are likely to be correlated across equations. I have experimented with running equations (1) through (3) and (4) through (6) as a seemingly unrelated regression (SUR) system. Because little efficiency is gained by doing so, I only present single-equation estimates, which are easier to compare with previous studies.

<sup>12</sup>One point that was not explicitly described in Berndt (1976) is the derivation of the standard errors for  $\sigma_4$ ,  $\sigma_5$  and  $\sigma_6$ . By a direct application of the Delta Method, the estimated variance of these elasticities can be computed as follows:

$$Est.Var(\sigma_i) = \left(\frac{-1}{\sigma_i}\right)^2 \cdot Est.Var\left(\frac{1}{\sigma_i}\right) \cdot \left(\frac{-1}{\sigma_i}\right)^2 \quad i = 4, 5, 6$$



proxy the flow of these services by total private employment, defined as the sum of the number of employees in private domestic industries and the number of self-employed workers.<sup>13</sup> For the regressions including the public sector (data configuration A below), the total number of government employees was added to the labor input measure. Jorgenson has argued repeatedly that total employment is not an appropriate measure of the flow of labor services because it ignores significant differences in the quality of the labor services provided by different workers. Jorgenson and collaborators have also provided quality-adjusted measures of labor services in several contributions by combining individual data from the Censuses of Population and from the Current Population Survey. Their measure of labor input reflects characteristics of individuals workers, such as age, sex and education, as well as class of employment and industry. In particular, their measure is a weighted sum of the supply of the different types or categories of labor input, where the weights are the share of overall labor compensation captured by a particular type. I consider here the most recent series reported in Jorgenson and Ho (2000), which considers 168 different categories of workers.

I take the nominal price of labor services to equal the total compensation of employees divided by  $L_t$ . Compensation of employees was obtained from the National Income and Production Accounts (NIPA) and includes wage and salary accruals, as well as supplements to wages and salaries (e.g., employer contributions for social insurance). Following the approach in Krueger (1999), I next correct this wage measure by adding two-thirds of proprietors' income to the overall compensation of employees.<sup>14</sup>

As is standard in the literature, I assume that the flow of capital services is proportional to the U.S. capital stock.<sup>15</sup> The nominal capital stock data was obtained from Herman (2000) and is defined as the sum of nonresidential private fixed assets and government assets, the latter being left out when only the private sector is considered. The real capital stock  $K_t$  is simply defined as the nominal capital stock divided by the price of capital.<sup>16</sup> I first construct

---

<sup>13</sup>These series were obtained from the Bureau of Economic Analysis website.

<sup>14</sup>Gollin (2002) suggests treating *all* proprietors's income as labor income. This alternative adjustment turns out to have only a marginal effect on the estimates (details available upon request).

<sup>15</sup>An interesting literature (Burnside et al., 1995; Basu, 1996) casts some doubts on this assumption by emphasizing the importance of variations in capital utilization for explaining the procyclical nature of productivity. An explicit correction for factor utilization is beyond the scope of this paper.

<sup>16</sup>As a robustness test, I employed the perpetual inventory method to construct an alternative measure of the real private capital stock using investment data from NIPA and depreciation data from Fraumeni (1997). The resulting capital stocks were remarkably sim-

the price of capital using nonresidential private investment deflators obtained from the NIPA. The NIPA deflator for capital equipment has been criticized for not adjusting for the increasing quality of capital goods, thereby systematically overstating the price of capital equipment. Krusell et al. (2000) have constructed an alternative deflator for equipment, building on previous work by Gordon (1990). They also suggest the use of the implicit price deflators for nondurable consumption and services when deflating the nominal stock of capital structures. I employ their price indices to construct an alternative deflator for private nonresidential fixed assets using Tornqvist's discrete approximation to the continuous Divisia index.<sup>17</sup> In particular, letting  $P_t^E$  be the adjusted price of equipment and  $P_t^S$  the price of structures, the price of capital  $P_t^K$  is constructed as follows:

$$\log P_t^K - \log P_{t-1}^K = \overline{s_t^E} (\log P_t^E - \log P_{t-1}^E) + (1 - \overline{s_t^E}) (\log P_t^S - \log P_{t-1}^S),$$

where  $\overline{s_t^E}$  is the arithmetic mean of the expenditure shares in capital equipment in the two periods, i.e.,

$$\overline{s_t^E} = \frac{1}{2} \frac{P_{t-1}^E K_{t-1}^E}{P_{t-1}^E K_{t-1}^E + P_{t-1}^S K_{t-1}^S} + \frac{1}{2} \frac{P_t^E K_t^E}{P_t^E K_t^E + P_t^S K_t^S}.$$

Capital income is defined as the sum of corporate profits, net interest, and rental income of persons, and is taken from the NIPA. The rental price of capital services  $R_t$  is computed as the ratio of total capital income to the real capital stock  $K_t$ .<sup>18</sup> As shown by Hulten (1986), if capital is the sole quasi-fixed input in production and there is perfect competition, this approach yields unbiased estimates of the unobserved shadow rental rate of capital, whereas the alternative Hall and Jorgenson (1967) formulae produce biased estimates.<sup>19</sup>

Finally, we are left with the construction of the aggregate input index  $F_t$  and its price  $P_t$ . Unfortunately, there is no clear counterpart for these variables in the data. Given that  $F_t = Y_t/A_t$ , one alternative would be to construct  $F_t$  by deflating value added  $Y_t$  by some index of Hicks-neutral technical efficiency. Berndt (1976) instead suggested constructing a measure of  $F_t$  based on the available data on capital and labor services. In particular, he computed the

ilar to those obtained by Herman (2000).

<sup>17</sup>The deflator for structures is also constructed as a Tonqvist index using the NIPA implicit price deflators for nondurable consumption and services.

<sup>18</sup>Assuming instead that the rental price of capital is proportional to the price of capital  $P_t^K$  leads to very similar results.

<sup>19</sup>I am grateful to an anonymous referee for pointing this out.

price of aggregate input  $P_t$  as a Tornqvist price index of the rental prices of capital ( $R_t$ ) and labor ( $W_t$ ). The aggregate input index  $F_t$  is then constructed as  $(R_t K_t + W_t L_t) / P_t$ . In the first part of the paper and in order to facilitate comparison with his results, I will follow the approach in Berndt (1976). In section 5, I will show that this approach is infeasible in the presence of biased technical change, and will discuss alternative specifications that make use of time series on value added.

**Table 1. Data Configurations**

	A	B	C	D	E
Only Private Sector	No	Yes	Yes	Yes	Yes
Include Proprietors' Income	No	No	Yes	Yes	Yes
Krusell et al. Deflator	No	No	No	Yes	Yes
Jorgenson and Ho labor input	No	No	No	No	Yes

Table 1 summarizes the six different data configurations for which I computed estimates of the elasticity. I started with specification A which (i) includes the public sector, (ii) does not add any fraction of proprietors' income to total compensation of employees, (iii) uses the NIPA deflators to construct the value of capital services and their price, and (iv) uses employment as a measure of labor input services. The public sector is excluded in data procedure B, while proprietors' income is added in C. Columns D and E incorporate sequentially the quality-adjusted price of capital indices of Krusell et al. (2000) and the quality-adjusted labor input series of Jorgenson and Ho's (2000). I interpret data configurations A through E as employing successively more refined data.

## 4 Estimates under Hicks-Neutral Technological Change

In this section, I present estimates of the elasticity of substitution between capital and labor based on both classical regression analysis and modern time series analysis. I start by reporting simple Ordinary Least Squares estimates of equations (1) through (6) for different data configurations. I later refine these estimates by dealing with issues related to autocorrelation of the disturbances, endogeneity of the regressors, and nonstationarity of the series.

### Ordinary Least Squares Estimation

Table 2 presents OLS estimates of equations (1) through (6) for the different data configurations described in the previous section. The results are striking in that all the estimates of the elasticity are remarkably close to one. Furthermore, the R-square in the regressions tends to increase with the quality and precision of the data, implying that the standard and reciprocal specifications of each first-order condition yield increasingly similar estimates. Furthermore, the more refined the data procedure, the closer the estimates become *across* first order conditions. For instance, with the least preferred data configuration (column I), the estimates range from 0.924 to 0.962, whereas with the most preferred data procedure (column V) this range collapses to the interval [1.002, 1.022]. A comparison of the standard errors of the estimates with those obtained by Berndt (and reported in column VIII of Table 2) reveals that my estimates are four to five times more precise than his. Despite this fact, the null hypothesis of a unit elasticity of substitution cannot be rejected at the 5% significance level for any of the six specifications.<sup>20</sup> Table 2 also reports the Durbin-Watson statistic for each estimation. The highest Durbin-Watson statistic in columns I through V is 0.625, indicating a clear rejection of the null hypothesis of no serial autocorrelation in the residuals.<sup>21</sup>

### Feasible Generalized Least Squares Estimation

The OLS Durbin-Watson statistics indicate the existence of serial correlation in the residuals but are not informative about the specific autocorrelation structure. A natural candidate is a standard AR(1) process, i.e.,  $\varepsilon_t = \rho\varepsilon_{t-1} + u_t$ , with  $u_t$  being white noise. Richer ARMA processes could potentially provide a better fit of the residuals, but this would leave us with fewer observations for the estimation of the parameters of interest. In order to study the plausibility of the assumption of an AR(1) process, I ran the regression  $\hat{\varepsilon}_t = \rho\hat{\varepsilon}_{t-1} + u_t$ , where  $\hat{\varepsilon}_t$  is the vector of OLS residuals in column V.<sup>22</sup> Ljung-Box tests at up to five lags were performed for each of the six specifications leading to no rejections of the null hypothesis of the estimated residuals  $\hat{u}_t$  being white noise. These results favor the use of an AR(1) process to parameterize the structure of the disturbances in equations (1) through (6).

Column VI of Table 2 then presents FGLS estimates of the elasticity obtained by applying the two-step Prais-Winsten procedure to the preferred data

<sup>20</sup>In Berndt (1976), the hypothesis is rejected in three of the six cases.

<sup>21</sup>For each of the six specifications in column V, I also performed a Ljung-Box test for autocorrelation. The null hypothesis of no autocorrelation up to order  $k$  was rejected in all six regressions for all  $k \leq 30$ .

<sup>22</sup>Hereafter, I limit the analysis to the most refined data configuration E.

Table 2. Estimates with Hicks-Neutral Technological Change

		OLS					FGLS	GIV	Berndt (1976)	
		A	B	C	D	E	E	E	OLS	2SLS
		I	II	III	IV	V	VI	VII	VIII	IX
<b>1</b>	$\sigma_1$	<b>0.924</b>	<b>0.929</b>	<b>1.000</b>	<b>1.005</b>	<b>1.002</b>	<b>0.944</b>	<b>0.992</b>	<b>0.967</b>	<b>1.148</b>
	<i>S.E.</i>	(0.027)	(0.022)	(0.025)	(0.020)	(0.020)	(0.039)	(0.048)	(0.082)	(0.098)
	$R^2$	0.961	0.974	0.971	0.981	0.981	0.994	0.960	0.785	0.757
	$D - W$	0.521	0.607	0.602	0.609	0.605	1.565	0.692	2.162	2.656
<b>2</b>	$\sigma_2$	<b>0.922</b>	<b>0.927</b>	<b>0.999</b>	<b>1.004</b>	<b>1.001</b>	<b>0.937</b>	<b>0.989</b>	<b>0.960</b>	<b>1.165</b>
	<i>S.E.</i>	(0.027)	(0.022)	(0.025)	(0.020)	(0.021)	(0.040)	(0.050)	(0.084)	(0.103)
	$R^2$	0.960	0.973	0.970	0.980	0.980	0.995	0.960	0.776	0.740
	$D - W$	0.529	0.600	0.582	0.589	0.584	1.533	0.702	2.119	2.668
<b>3</b>	$\sigma_3$	<b>0.924</b>	<b>0.929</b>	<b>1.000</b>	<b>1.005</b>	<b>1.002</b>	<b>0.942</b>	<b>0.991</b>	<b>0.966</b>	<b>1.151</b>
	<i>S.E.</i>	(0.027)	(0.022)	(0.025)	(0.020)	(0.020)	(0.039)	(0.049)	(0.082)	(0.098)
	$R^2$	0.961	0.974	0.971	0.981	0.981	0.983	0.960	0.784	0.755
	$D - W$	0.523	0.605	0.597	0.604	0.600	1.558	0.693	2.160	2.663
<b>4</b>	$\sigma_4$	<b>0.962</b>	<b>0.954</b>	<b>1.030</b>	<b>1.025</b>	<b>1.022</b>	<b>1.022</b>	<b>1.017</b>	<b>1.231</b>	<b>1.233</b>
	<i>S.E.</i>	(0.028)	(0.022)	(0.026)	(0.020)	(0.020)	(0.042)	(0.050)	(0.103)	(0.108)
	$R^2$	0.961	0.974	0.971	0.981	0.981	0.991	0.858	0.785	0.785
	$D - W$	0.544	0.625	0.622	0.622	0.620	1.720	1.996	2.699	2.699
<b>5</b>	$\sigma_5$	<b>0.961</b>	<b>0.952</b>	<b>1.030</b>	<b>1.024</b>	<b>1.021</b>	<b>1.022</b>	<b>1.015</b>	<b>1.238</b>	<b>1.245</b>
	<i>S.E.</i>	(0.028)	(0.023)	(0.026)	(0.021)	(0.021)	(0.043)	(0.052)	(0.108)	(0.113)
	$R^2$	0.960	0.973	0.970	0.980	0.980	0.993	0.850	0.776	0.775
	$D - W$	0.552	0.618	0.602	0.602	0.600	1.696	1.915	2.705	2.705
<b>6</b>	$\sigma_6$	<b>0.961</b>	<b>0.954</b>	<b>1.030</b>	<b>1.024</b>	<b>1.022</b>	<b>1.022</b>	<b>1.017</b>	<b>1.232</b>	<b>1.235</b>
	<i>S.E.</i>	(0.028)	(0.022)	(0.026)	(0.020)	(0.021)	(0.042)	(0.050)	(0.104)	(0.108)
	$R^2$	0.961	0.974	0.971	0.981	0.981	0.973	0.857	0.784	0.784
	$D - W$	0.546	0.623	0.617	0.617	0.615	1.715	1.979	2.704	2.704
No. Obs.		51	51	51	51	51	51	50	40	40

configuration E.<sup>23</sup> The FGLS estimates are overall very similar to the OLS ones and all lie in the small interval [0.937, 1.022]. The FGLS standard errors are substantially higher than the OLS ones and lead again to no rejections of the null hypothesis of the U.S. aggregate production function being Cobb-Douglas.

### Generalized IV Estimation

Berndt's (1976) approach of estimating equations (1) through (3) *and* their reverse specifications clearly exposes the existence of an endogeneity problem. Equations (1) through (6) were derived from the first order conditions of profit maximization and, hence, they can be readily interpreted as the aggregate private sector demand for capital and labor services. From the theory of simultaneous equations models, it is well known that these demand equations will not be identified unless the estimation makes use of some set of exogenous variables that shift the supply of capital and labor (cf., Hausman, 1983). Consequently, the OLS and FGLS estimates presented above are likely to be biased, with the direction of the bias being uncertain.

Berndt (1976) acknowledged the same simultaneous equation bias and proposed a simple two-stage least squares (2SLS) procedure to resolve it. In his first stage regressions, Berndt (1976) introduced a rather large number of instruments.<sup>24</sup> The use of such large set of instruments can be detrimental in at least two respects. On the one hand, if any of these instruments is in fact endogenous, 2SLS estimates will be inconsistent. On the other hand, if some of the instruments are only weakly correlated with the regressors, even if the exogeneity requirement is met, small sample biases will arise (cf., Bound, Jaeger and Baker, 1995). For these reasons, I instead focus on a smaller set of instruments. In particular I take the following three variables to be exogenous to the model but correlated with the regressors: (1) U.S. population, (2) wages in the government sector, and (3) real capital stock owned by the government.<sup>25</sup> I interpret these variables as different types of supply shifters. It is clear that the size of the U.S. population is likely to have a significant effect on the supply of both capital and labor services. Government wages are also likely to affect the supply of labor in the private sector, with government

---

<sup>23</sup>Iterating the coefficients to convergence had only a minor effect on the results. Since nothing is gained asymptotically by iterating the process, I only report the two-step estimates, which are more comparable to the Generalized IV estimates reported in Table 2 and discussed below.

<sup>24</sup>See Berndt and Christensen (1973) or Antràs (2003) for a complete list.

<sup>25</sup>Wages in the government sector are computed as labor income accruing to government employees divided by their total number, and deflated by the aggregate input price index  $P_t$ . To construct the real stock of capital owned by the government I divide the nominal figures of Herman (2000) by the price of capital index  $P_t^K$ .

capital formation having an analogous effect on the supply of capital in the private sector. The exogeneity of these instruments also seems plausible. As it is standard in macroeconomics, I take the fertility choice to be exogenous to the model, while government variables are assumed not to respond (at least contemporaneously) to market prices and quantities.

Having discussed the choice of instruments, I next turn to the selection of an appropriate estimation technique. One alternative would be to run equations (1) through (6) using a standard 2SLS procedure. Nevertheless, there is no reason to believe that instrumenting would solve the autocorrelation problem discussed above. I choose instead to implement a generalized instrumental variable (GIV) procedure developed by Fair (1970) and which I summarize in Appendix A. Column VIII of Table 2 presents estimates of the elasticity of substitution obtained by applying this technique to our preferred data configuration E. The estimates are contained in the interval (0.989, 1.017). The standard errors are again higher than the OLS ones and the null hypothesis of a unit elasticity of substitution cannot be rejected for any of the six estimates. A quick comparison of columns VI and VII also reveals that, relative to the FGLS estimates, the GIV estimates are slightly higher in equations (1), (2), and (3), but slightly lower in equations (4), (5), and (6). I interpret this as an indication that the instruments are not only dealing with the simultaneous equation bias, but might also be correcting for a latent errors-in-variables bias (remember that equations (4), (5), and (6) deliver an estimate of  $1/\sigma_i$ ).<sup>26</sup>

### Time Series Estimation

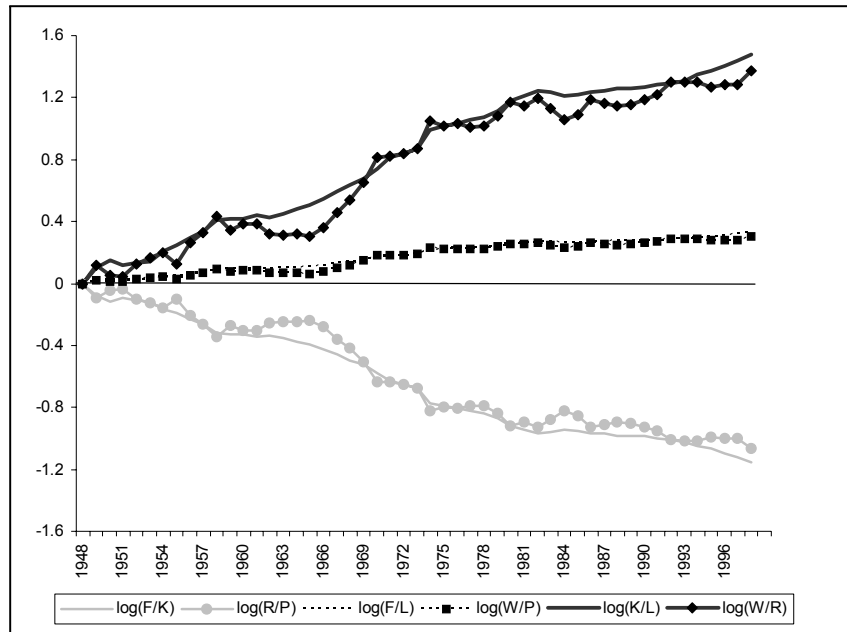
Up to this point, I have followed closely the approach proposed by Berndt (1976), the major variation being in the explicit treatment of serial correlation in the disturbances. I now turn to a whole set of different issues that arise when considering the nonstationary nature of the series involved in the estimation.

Figure 1 graphs the six series that form the basis of our estimates, where the logarithm of the variables has been normalized to equal 0 in 1948. Two facts emerge from the figure. First, the graph uncovers potential nonstationarities in each one of the series: the logarithm of  $W_t/P_t$ ,  $F_t/L_t$ ,  $W_t/R_t$  and  $K_t/L_t$  all clearly trend upwards, while  $R_t/P_t$  and  $F_t/K_t$  show a downward trend. Second, the two variables in each of the specifications (1) through (6) follow similar

---

<sup>26</sup>Conditional on the process followed by the disturbances being AR(1) and under the null hypothesis of exogeneity of the regressors, FGLS provides consistent and asymptotically efficient estimates of the elasticity of substitution, whereas Fair's GIV estimates are also consistent but inefficient. Under the alternative hypothesis, the FGLS estimates are inconsistent, while the GIV ones remain consistent. Using a Hausman (1978) specification test, I tested for the null hypothesis of exogeneity of the regressors and for all six equations the null was not rejected at significance levels well above 5%.

Figure 1: Nonstationarity of the Series



trends. This suggests that the correlations captured in the regressions above might all be of the so-called spurious type (cf. Granger and Newbold, 1974, and Phillips, 1986). The very high R-squares and the low Durbin-Watson statistics obtained in the OLS estimation certainly point out towards that direction. I next turn to a formal investigation of this possibility.

Table 3 reports a summary of the unit root tests I performed on each of the series. The first row presents the results of a simple Dickey-Fuller test of a unit root in the series against the alternative hypothesis of trend-stationarity. It is clear from Table 3 that for none of the six series does the test reject the hypothesis of a unit root. The next two rows extend this simple test to allow for serial correlation by adding higher-order autoregressive terms to the test. I performed this so-called Augmented Dickey-Fuller test with one and two lags, and the null hypothesis of a unit root was again not rejected for any of the series. Finally, I also implemented a Phillips-Perron test at truncation lags 2, 3 and 4 reaching again the same conclusion. In the bottom panel of Table 3, I report the results of the same tests performed on each of the six series expressed in first differences. In this case, the results indicate a rejection of the null hypothesis of the series being integrated of order two.

I therefore conclude that all six series are nonstationary and integrated of



**Table 3. Unit Root Tests**

	$\log\left(\frac{Y}{K}\right)$	$\log(r)$	$\log\left(\frac{Y}{L}\right)$	$\log(w)$	$\log\left(\frac{K}{L}\right)$	$\log\left(\frac{w}{r}\right)$	5% Critical Value
ADF 0	-1.082	-1.759	-1.089	-1.764	-1.083	-1.760	-3.499
ADF 1	-1.304	-1.823	-1.342	-1.882	-1.311	-1.836	-3.501
ADF 2	-1.249	-1.815	-1.305	-1.867	-1.259	-1.827	-3.504
PP 2	-1.333	-1.872	-1.341	-1.894	-1.334	-1.877	-3.499
PP 3	-1.379	-1.860	-1.385	-1.886	-1.380	-1.866	-3.499
PP 4	-1.405	-1.911	-1.404	-1.931	-1.404	-1.915	-3.499

	$\Delta \log\left(\frac{Y}{K}\right)$	$\Delta \log(r)$	$\Delta \log\left(\frac{Y}{L}\right)$	$\Delta \log(w)$	$\Delta \log\left(\frac{K}{L}\right)$	$\Delta \log\left(\frac{w}{r}\right)$	5% Critical Value
ADF 1	-4.338	-4.782	-4.250	-4.745	-4.322	-4.773	-3.503
PP 3	-5.652	-6.828	-5.553	-6.666	-5.634	-6.793	-3.501

**Table 4. Cointegration Tests**

<b>A. Residual-Based Augmented Dickey-Fuller Tests</b>							
	Residuals of eq. (1)	Residuals of eq. (2)	Residuals of eq. (3)	Residuals of eq. (4)	Residuals of eq. (5)	Residuals of eq. (6)	5% Critical Value
ADF 0	-3.008	-2.953	-2.995	-3.026	-2.971	-3.013	-2.920
ADF 1	-3.382	-3.378	-3.380	-3.371	-3.368	-3.369	-2.921
ADF 2	-2.778	-2.779	-2.778	-2.866	-2.868	-2.866	-2.923

<b>B. Johansen-Juselius Cointegration Tests</b>									
Test	Num. of lags	Max-Lambda				Trace			
		$r = 0$ vs $r = 1$		$r \leq 1$ vs $r = 2$		$r = 0$ vs $r = 1$		$r \leq 1$ vs $r = 2$	
		1	2	1	2	1	2	1	2
$\log(Y/K)$ & $\log(r)$		74.58	7.83	0.02	0.56	74.59	8.39	0.02	0.56
$\log(Y/L)$ & $\log(w)$		67.11	8.11	0.01	0.40	67.12	8.51	0.01	0.40
$\log(K/L)$ & $\log(w/r)$		72.91	7.85	0.01	0.52	72.93	8.37	0.01	0.52
95 % Critical Values		15.67		9.24		19.96		9.24	

order one, which implies that the OLS, FGLS and GIV estimates computed above are all potentially subject to a spurious regression bias. In fact, as shown by Phillips (1986), in this situation, OLS estimates will not be consistent unless a linear combination of the dependent and independent variables is stationary, that is, only if the two variables entering each regression are cointegrated.<sup>27</sup>

Table 4 presents the results from two cointegration tests. The top panel considers Engle and Granger's (1987) residual-based Augmented Dickey-Fuller test, which hinges on testing the stationarity of the residuals from the OLS regressions (1) through (6). As pointed out by Engel and Granger (1987), the critical values of standard unit root tests are not appropriate when applied to the OLS residuals because they lead to too many rejections of the null hypothesis of no cointegration. MacKinnon (1991) has linked the appropriate critical values to the sample size and to a set of parameters that only vary with the specification of the cointegration equation, the number of variables and the significance level. These critical values are reported in the last column. As is apparent from the Table, the results are not conclusive. When the estimation includes one lagged first difference of the residuals, the null hypothesis of nonstationarity of the residuals is rejected for all six specifications, implying that the series in the estimation seem to be cointegrated. Nevertheless, when a second lagged first difference is added, the results are overturned and the tests suggest instead that the series in the estimation may in fact not be cointegrated.

The bottom panel of Table 4 implements the maximum likelihood cointegration test suggested by Johansen and Juselius (1990), which tests the null hypothesis of the existence of  $r$  cointegrating vectors against the alternative of the existence of  $r + 1$  cointegrating vectors. Implementing the test requires specifying a particular model for the cointegration equation as well as choosing the number of lags of the first difference of the variables to be included in the estimation. In light of equations (1) through (6), I choose a model with a constant and no trend and compute the statistics with one and two lagged first differences of the data. The results in the bottom panel of Table 4, indicate that, when the estimation includes one lag, the null hypothesis of no cointegration is clearly rejected for all of specifications. As with the Engle and Granger (1987) tests, however, when the number of lags is increased to two, the null hypothesis of zero cointegrating vectors cannot be rejected for any of the specifications.

Overall, these mixed results of the cointegration tests indicate that the OLS estimates in Table 2 should be interpreted with caution because of a potential

---

<sup>27</sup>As discussed below, this is not necessarily true for our FGLS and GIV estimates.

spurious regression bias. As discussed by Hamilton (1994, p. 562), a natural cure for spurious regressions would be to difference the data before estimating the equations. The disadvantage of this approach is that important long-run information would be lost and, in particular, the interpretation of our estimates of  $\sigma_i$  would become much less transparent. Interestingly, the existence of a unit root in the OLS residuals implies that the FGLS and GIV estimates in Table 2 are asymptotically equivalent to the estimates that would be obtained with the differenced data (cf. Blough, 1992, and the discussion in Hamilton, 1994, p. 562). This implies that the estimates in columns VI and VII of Table 2 are consistent, but it also indicates that these estimates might be neglecting important long-run information in the data.

## 5 Estimates under Biased Technological Change

The results in the previous section provide some evidence in favor of a Cobb-Douglas specification of the U.S. aggregate production function. In particular, the GIV estimates on the preferred data configuration lead to no rejections of the null hypothesis of a unit elasticity of substitution between capital and labor. Nevertheless, an explicit treatment of the nonstationary nature of the data suggested that the results should be treated with caution.

In this section, I will cast further doubts on the validity of the estimates obtained under the assumption of Hicks-neutral technological change. I will first show that in the presence of biased technical change, Berndt's (1976) estimation equations are misspecified in a critical way, in fact biasing the estimates towards finding the results that support a Cobb-Douglas view of the U.S. economy. I will next offer some solutions to this misspecification problem and will examine how the estimates are affected by these modifications.

### The Source of the Bias

Consider again the Arrow et al. (1961) CES production function now expanded to allow for non-neutral technological change

$$Y_t = \left[ \delta (A_t^K K_t)^{\frac{\sigma-1}{\sigma}} + (1 - \delta) (A_t^L L_t)^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}}, \quad (8)$$

where  $A_t^K$  is an index of capital-augmenting efficiency and  $A_t^L$  is an index of labor-augmenting efficiency. It is straightforward to show that, given equation (8), there is no total-factor productivity index  $A_t$  such that  $F_t = Y_t/A_t$  is only a function of the capital-labor ratio, and not of  $A_t^K$ ,  $A_t^L$ , or their ratio. In other words, under biased technological change, it becomes impossible to

remove factor efficiency from the aggregate input index  $F_t$ . This implies that, under these circumstances, the estimations equations (1), (2), (4), and (5), which include  $F_t$  and its associated price, are all misspecified in the sense that they suffer from an omitted-variable bias. Furthermore, taking the first order conditions with respect to capital and labor, one can obtain the following expression from which equations (3) and (6) are derived:

$$\log\left(\frac{W_t}{R_t}\right) = \log\left(\frac{1-\delta}{\delta}\right) + \left(\frac{1}{\sigma}\right)\log\left(\frac{K_t}{L_t}\right) + \left(\frac{1-\sigma}{\sigma}\right)\log\left(\frac{A_t^K}{A_t^L}\right) \quad (9)$$

Equation (9) clearly illustrates that as long as  $A_t^K \neq A_t^L$  (i.e., as long as technological change is non-neutral) equations (3) and (6) also suffer from an omitted-variable bias. To get a better understanding on the direction of the bias, it is useful to subtract  $\log(K_t/L_t)$  from equation (9):

$$\log\left(\frac{W_t L_t}{R_t K_t}\right) = \log\left(\frac{1-\delta}{\delta}\right) + \left(\frac{1-\sigma}{\sigma}\right)\log\left(\frac{A_t^K K_t}{A_t^L L_t}\right). \quad (10)$$

The left-hand side of equation (10) is simply the logarithm of the labor share in total output divided by the capital share. As discussed in the introduction, this variable has been remarkably stable over the period 1948-1998. On the other hand, the capital-labor ratio  $K_t/L_t$  on the right-hand side of (10) has steadily increased during the same period. Consequently, whenever the bias in technological change is ignored, i.e., whenever the ratio  $A_t^K/A_t^L$  is not included in the regression, the estimate of  $(1-\sigma)/\sigma$  will necessarily be close to zero, implying that the estimate of  $\sigma$  will necessarily be close to one. At first glance this does not seem surprising: if a Cobb-Douglas function describes well aggregate production in the U.S. private sector, the labor share should be approximately constant. The problem is that in the presence of biased technological change the argument does not run both ways. As it is clear from equation (10), if  $K_t/L_t$  and  $A_t^L/A_t^K$  grow at the same rate, then steady factor shares can be consistent with *any* well-behaved production function, and certainly with aggregate production functions with non-unit elasticities of substitution.

A notable example is a version of the neoclassical growth model with labor-augmenting technological change (cf., Barro and Sala-i-Martin, 1995, Chapter 2). With rather weak conditions on the aggregate production function, the model delivers a balance growth path in which the capital-labor ratio grows at the same rate as the index of labor-augmenting efficiency. Another example is provided by Acemoglu (2003), who develops a model in which the incen-

tives to innovate depend on the size of the bill paid to each factor. In his model an increase in the labor share encourages labor-augmenting technological change, which in turn increases the capital-labor ratio and the wage-rental ratio. The fact that capital and labor are assumed to be gross complements ( $\sigma < 1$ ) ensures that the increase in  $K/L$  is higher than the increase in  $w/r$ , thereby bringing the labor share back to its steady-state equilibrium value. Acemoglu's (2003) model nicely illustrates the misspecification inherent in restricting technological change to be neutral in the regressions above. By not controlling for the bias in technological change, one is implicitly ascribing the full variation in the capital-labor to factor substitution, when in fact part of the variation is explained by technological change.

### Model Specification and Additional Data

As discussed above, in the presence of biased technological change, it is impossible to construct an index of aggregate input  $F_t$  that is independent of the efficiency indices  $A_t^K$  and  $A_t^L$ . Consequently, in order to consistently estimate the elasticity of substitution, it becomes necessary to devise a method to control for these indices, which in turn requires the imposition of some type of structure on the form of technological change (cf., Diamond, McFadden and Rodriguez, 1978). I follow the bulk of the literature in assuming that  $A_t^K$  and  $A_t^L$  grow at constant rates  $\lambda_K$  and  $\lambda_L$ .<sup>28</sup> Under this assumption, the production function (7) becomes

$$Y_t = \left[ \delta (A_0^K e^{\lambda_K \cdot t} K_t)^{\frac{\sigma-1}{\sigma}} + (1-\delta) (A_0^L e^{\lambda_L \cdot t} L_t)^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}}.$$

The first-order conditions for profit maximization can then be manipulated to obtain the following six specifications, analogous to equations (1) through (6) in section 2:<sup>29</sup>

$$\log(Y_t/K_t) = \alpha'_1 + \sigma \log(R_t/P_t^Y) + (1-\sigma)\lambda_K \cdot t + \varepsilon_{1,t} \quad (1')$$

$$\log(Y_t/L_t) = \alpha'_2 + \sigma \log(W_t/P_t^Y) + (1-\sigma)\lambda_L \cdot t + \varepsilon_{2,t} \quad (2')$$

$$\log(K_t/L_t) = \alpha'_3 + \sigma \log(W_t/R_t) + (1-\sigma)(\lambda_L - \lambda_K) \cdot t + \varepsilon_{3,t} \quad (3')$$

$$\log(R_t/P_t^Y) = \alpha'_4 + (1/\sigma) \log(Y_t/K_t) - [(1-\sigma)/\sigma] \lambda_K \cdot t + \varepsilon_{4,t} \quad (4')$$

$$\log(W_t/P_t^Y) = \alpha'_5 + (1/\sigma) \log(Y_t/L_t) - [(1-\sigma)/\sigma] \lambda_L \cdot t + \varepsilon_{5,t} \quad (5')$$

<sup>28</sup>Below, I briefly discuss an alternative specification with technological change being partly stochastic.

<sup>29</sup>Lucas (1969), David and van de Klundert (1965), and Kalt (1978) estimate different subsets of these specifications.

$$\log(W_t/R_t) = \alpha'_6 + (1/\sigma) \log(K_t/L_t) - [(1 - \sigma)/\sigma] (\lambda_L - \lambda_K) \cdot t + \varepsilon_{6,t}. \quad (6')$$

There are two important differences between the specifications in (1') through (6') and those in (1) through (6). First, because of the impossibility of constructing an aggregate input index  $F_t$ , this index is replaced by real output  $Y_t$ . Second, all six specifications now include a time trend. As is clear from the equations, the exclusion of the time trend in the regressions above, would in general lead to inconsistent estimates of the elasticity of substitution. The only exception is the Cobb-Douglas case ( $\sigma = 1$ ), in which the bias would effectively be zero.<sup>30</sup>

Estimation of equations (1') through (6') requires data on real output  $Y_t$  and its associated price  $P_t^Y$ . A natural candidate to proxy for  $Y_t$  is real GDP in the U.S. private sector. As argued by Berndt (1976), using value added to measure  $Y_t$  is in general problematic. As he points out, most studies consider only a subset of capital inputs, namely equipment and structures, thus ignoring land, inventories, and working capital. Because in this paper I have also focused on capital equipment and structures, the use of value added as a proxy for  $Y_t$  is arguably inappropriate. Nevertheless, in the presence of biased technological change, Berndt's (1976) approach of constructing the aggregate input index is infeasible, because the index  $F_t$  will necessarily depend on  $A_t^K$  and  $A_t^L$ , which are unobservable. Furthermore, as discussed above, ignoring the adjustment for  $A_t^K$  and  $A_t^L$  is not an option, since this leads to estimates of the elasticity that are necessarily biased towards one. The infeasibility of using Berndt's (1976) aggregate input approach inclines me to use series on value added. In particular, I use GDP in the U.S. private sector to proxy for  $Y_t$ , and the corresponding GDP deflator to proxy for  $P_t^Y$ .

### Estimation Results

Column I in Table 5 presents OLS estimates of equations (1') through (6'). In order to assess the effect of controlling for biased technological change, these estimates should be compared with those in column V of Table 2. As it is clear from the results, the point estimates of the elasticity drop to values that are, in general, well below one and that range from 0.551 to 0.948. Furthermore, the standard errors of the estimates are quite low, implying that the null hypothesis of a unit elasticity is rejected at the 1% significance level for five of the six specifications, while it is rejected at the 10% level for the remaining equation (see the t-stats in Table 5). Interestingly, the estimates are also consistent with the empirical regularity discussed in Berndt (1976), by which the estimates of

<sup>30</sup>Notice also that the constant terms are not only a function of  $\delta$ , but depend also on  $\sigma$ ,  $A_0^K$  and  $A_0^L$  (see Klump and De La Grandville, 2000, for more on this).

the elasticity based on the marginal product of labor equations tend to be higher than the estimates based on the marginal product of capital equations. As in the regressions in section 4, the Durbin-Watson statistics indicate the existence of serial correlation in the residuals.<sup>31</sup> I performed again Ljung-Box tests and the null hypothesis of the residuals following an AR(1) process was again not rejected.

Column II in Table 5 presents FGLS estimates that apply the Prais-Winsten procedure. The results are qualitatively similar to the OLS ones, with the lowest estimates becoming even lower and the highest estimate reaching a value slightly higher than one. Finally, in column III, I present estimates based on Fair's (1970) GIV technique. Instrumentation seems to help a great deal in bringing the estimates from different specifications closer together. In particular, the GIV estimates range from 0.681 to 0.891, and even the highest point estimate,  $\sigma_5$ , is significantly lower than one at the 5% significance level. Overall, of the eighteen estimates in columns I, II, and III, seventeen are below one, with fifteen of these estimates being significantly below one at the 5% level. Furthermore, the *typical* elasticity lies in the range 0.6 to 0.9.

The high R-squares and low Durbin-Watson statistics obtained under OLS suggest that, as in section 4, the results might suffer from a spurious regression bias. Section 4 uncovered significant nonstationarities in the data used in the regressions with Hicks-neutral technological change. I repeated the unit root tests for the series involved in the estimation of equations (1') through (6') and found very similar results. The null hypothesis of the series being integrated of order one was not rejected in any of the six cases, while the null hypothesis of their first difference being integrated of order one was clearly rejected in all cases. These results call for a cointegration test to assess whether or not the regressions above are spurious. The top panel of Table 6 presents the results of Engle and Granger's (1987) residual-based test. The residuals are obtained from a model that includes a time trend, and the specification of the test includes one, two or three lags of the first difference of the residuals. The statistics reported in Table 6 should be compared with the critical values computed à la MacKinnon (1991), which are also adjusted to take into account the time trend in the equations. As in section 4, the results are not conclusive. When the estimation includes one lag, the tests generally reject the null hypothesis of no cointegration, but the results are not robust to adding a second lagged first difference of the residuals. Similarly, in the bottom panel of Table 6 I report the Johansen and Juselius (1990) test, which equally de-

---

<sup>31</sup>These statistics are not reported to save space. Even the highest Durbin-Watson statistic was lower than one.

Table 5. Estimates Allowing for Biased Technological Change

		OLS	FGLS	GIV	Saikkonen	With Lags
		I	II	III	IV	V
<b>1</b>	$\sigma_1$	<b>0.551</b>	<b>0.407</b>	<b>0.681</b>	<b>0.641</b>	<b>0.408</b>
	t-stat for $H_0: \sigma_1 = 1$	-9.55	-14.2	-3.43	-2.79	-15.07
	$R^2$	0.972	0.996	0.905	0.977	0.993
<b>2</b>	$\sigma_2$	<b>0.908</b>	<b>0.918</b>	<b>0.880</b>	<b>0.872</b>	<b>1.004</b>
	t-stat for $H_0: \sigma_1 = 1$	-3.36	-2.02	-3.00	-2.94	0.04
	$R^2$	0.998	0.996	0.992	0.999	0.999
<b>3</b>	$\sigma_3$	<b>0.596</b>	<b>0.378</b>	<b>0.685</b>	<b>0.687</b>	<b>0.313</b>
	t-stat for $H_0: \sigma_1 = 1$	-8.33	-11.7	-3.06	-2.27	-15.74
	$R^2$	0.992	0.995	0.981	0.994	0.998
<b>4</b>	$\sigma_4$	<b>0.743</b>	<b>0.637</b>	<b>0.770</b>	<b>0.807</b>	<b>0.578</b>
	t-stat for $H_0: \sigma_1 = 1$	-4.05	-5.65	-2.08	-0.70	-7.58
	$R^2$	0.933	0.965	0.623	0.940	0.969
<b>5</b>	$\sigma_5$	<b>0.948</b>	<b>1.013</b>	<b>0.891</b>	<b>0.892</b>	<b>1.521</b>
	t-stat for $H_0: \sigma_1 = 1$	-1.84	0.28	-2.70	-1.95	3.20
	$R^2$	0.998	0.996	0.994	0.999	0.999
<b>6</b>	$\sigma_6$	<b>0.785</b>	<b>0.750</b>	<b>0.772</b>	<b>0.797</b>	<b>0.587</b>
	t-stat for $H_0: \sigma_1 = 1$	-3.35	-2.59	-1.93	-0.77	-5.03
	$R^2$	0.983	0.977	0.896	0.984	0.991
No. of Obs.		51	51	50	48	50

Table 6. Cointegration Tests for Model with a Time Trend

A. Residual-Based Augmented Dickey-Fuller Tests								
	Residuals of eq. (1)	Residuals of eq. (2)	Residuals of eq. (3)	Residuals of eq. (4)	Residuals of eq. (5)	Residuals of eq. (6)	5% Critical Value	
ADF 0	-2.407	-4.783	-2.860	-2.990	-4.509	-3.262	-3.501	
ADF 1	-3.071	-4.165	-3.219	-3.533	-3.948	-3.719	-3.502	
ADF 2	-2.641	-3.086	-2.315	-2.965	-3.027	-3.006	-3.504	
B. Johansen-Juselius Cointegration Tests								
Test	Max-Lambda				Trace			
	$r = 0$ vs $r = 1$		$r \leq 1$ vs $r = 2$		$r = 0$ vs $r = 1$		$r \leq 1$ vs $r = 2$	
	1	2	1	2	1	2	1	2
log( $Y/K$ ) & log( $r$ )	26.83	12.97	0.00	5.41	26.83	18.38	0.00	5.41
log( $Y/L$ ) & log( $w$ )	40.72	17.28	0.12	0.10	40.84	17.38	0.12	0.10
log( $K/L$ ) & log( $w/r$ )	27.12	9.76	0.05	1.47	27.17	11.23	0.05	1.47
95 % Critical Values	18.96		12.25		25.32		12.25	



livers different conclusions depending on the number of lags specified in the estimation. Both types of tests suggest, however, that the null hypothesis of no cointegration is more easily rejected in the case of the marginal product of labor equations (2') and (5').<sup>32</sup>

The mixed results of the cointegration tests complicate the evaluation of the consistency of the OLS estimates presented in Table 5. On the one hand, if we are willing to reject the null hypothesis of no cointegration (i.e., if we believe the tests should be specified with one lag), OLS estimates are not only consistent but also superconsistent, in the sense that they converge to their true value at a higher speed than in the absence of nonstationarity in the series (cf., Phillips and Durlauf, 1986, and Stock, 1987). Furthermore, in that case, OLS estimates are consistent even in the presence of autocorrelation in the disturbances and endogeneity of the regressors. Phillips and Durlauf (1986) showed, however, that OLS estimates of cointegrated relationships have nonstandard asymptotic distributions, thus invalidating standard inference techniques. Furthermore, small-sample biases, which are likely to be important in our regressions, suggest the need to use alternative superior estimates. Saikkonen (1991) suggests a simple modification to the OLS procedure, which delivers both consistent and efficient estimates that have asymptotically standard distributions.<sup>33</sup>

On the other hand, if we interpret the results in Table 6 as indicating that the null hypothesis of no cointegration cannot be rejected (i.e., if we believe the tests should be specified with two lags), then OLS estimates suffer from a spurious regression bias, and are therefore inconsistent. As argued before, a natural cure for this bias is to difference the data before estimating the equations, the disadvantage being that, by doing so, valuable long-run information may be lost. An alternative approach is to include lagged values of both the dependent and independent variables in the regression. This procedure leads to consistent estimates of the elasticity and to *t*-tests of the hypothesis  $\sigma_i = 1$  that are asymptotically  $N(0, 1)$ .<sup>34</sup>

Rather than taking a strong stance on whether the variables in the regressions are in fact cointegrated or not, I next report the results of applying Saikkonen's (1991) procedure for estimating cointegrating vectors, as well as

---

<sup>32</sup>The statistics in the Engle and Granger (1987) are always higher for these two equations. Furthermore, even with two lagged first differences in the estimation, the max-lambda statistic in the Johansen and Juselius test is very close to its critical value for this pair of variables.

<sup>33</sup>See King, Plosser, Stock and Watson (1991) for an application of a similar technique.

<sup>34</sup>Conversely, *F* tests of hypotheses that a set of estimates are jointly significant have nonstandard limiting distribution (c.f., Hamilton, pp. 562).

standard OLS estimates that include lagged values of both variables to correct for spurious regression bias in the absence of cointegration.<sup>35</sup>

Column IV of Table 5 presents the results of the implementation of Saikkonen's (1991) procedure for  $l = 1$  and  $p = 1$ . The details of the estimation procedure are relegated to Appendix B. For each of the six regressions, the table reports the estimate of the elasticity of substitution and the modified t-statistic, which should be compared with the associated critical value from a standard normal distribution.<sup>36</sup> The results indicate that the inclusion of the leads and lags does not have much of an effect on the point estimates of the elasticity, which are very similar to those obtained under the GIV procedure in column III of Table 5. The modified t-statistics are somewhat lower than the standard ones, but they still lead to a rejection of the null hypothesis of a unit elasticity (at reasonable significance levels) in four of the six cases.

Finally, in column V of Table 5, I report OLS estimates extended to include lagged values of both the dependent variable  $y_t$  and the independent variable  $x_t$ , as well as a time trend. This has a larger impact on the estimates, suggesting that spurious regression biases might be important. The estimates of the elasticity obtained from equations (1'), (3'), (4'), and (6') are substantially lower than the ones obtained in the first four columns of Table 5, and provide evidence that the elasticity of substitution might well be lower than 0.5. These values are consistent with the findings of David and van de Klundert (1965), Eisner and Nadiri, and Lucas (1969). On the other hand, the estimates of the elasticity obtained from equations (2') and (5') indicate that the elasticity is much higher, and might even be larger than one. Remember, however, that the approach of adding lags of both variables in the model is only appropriate under the null hypothesis of no cointegration of the variables. The fact that this hypothesis was relatively easier to reject for the variables in equations (2') and (5') suggests that these high values of the elasticity should not be taken at face value.

### Estimates of the Bias in Technological Change

As is apparent from equations (1') through (6'), with the our estimates of the elasticity of substitution at hand, we can also obtain estimates of the parameters  $\lambda_K$  and  $\lambda_L$ , that is, estimates of the growth rate of capital- and labor-augmenting technological change. This was already recognized by David

<sup>35</sup>The Saikkonen (1991) procedure also adds lags to the estimated equation (see Appendix B). Although the introduction of these lags is here justified on statistical grounds, these lags could also be rationalized appealing to adjustment costs in capital and labor (see Lucas, 1969).

<sup>36</sup>This modified t-statistic corresponds to  $t \cdot \sqrt{s^2}/\lambda$  in Appendix B.

and van de Klundert (1965) who ran an equation analogous to (2') for the U.S. private sector in the period 1899-1960 and found  $\lambda_L = 0.019$ , or a growth rate of labor efficiency of 1.9 percent a year. They also ran equations analogous to (3'), both with and without lags, thus obtaining values for  $\lambda_L - \lambda_K$  of 0.72% and 0.86% per year, respectively. As shown in Table 6, using my own estimates of  $\sigma_3$  from equation (2') under Saikonnen's procedure (i.e., column IV), yields an estimate of  $\lambda_L$  of 1.85% per year, a figure remarkably close to David and van de Klundert's. Furthermore, when computing  $\lambda_L$  with the estimates of the reverse equation (5'), I obtain a value of 1.90% which matches their figure to the second decimal. On the other hand, as is clear from Table 6, my findings suggest that the bias in technological change,  $\lambda_L - \lambda_K$ , is much larger than the one they estimated. In particular, I find that labor-augmenting efficiency grew about 3% faster than capital-augmenting efficiency. In fact, my estimates suggest that capital efficiency shows a downward trend. To understand this result, remember that  $A_t^L$  and  $A_t^K$  are actually indices of *unmeasured* quality or *unmeasured* efficiency. A possible interpretation of the findings on Table 7 is that the Krusell et al. (2000) price of capital deflator does a better job of incorporating quality improvements than does the Jorgenson and Ho (2000) quality-adjusted labor input index.

**Table 7. Estimates of the Bias in Technological Change**

	Estimate of Elasticity	Coefficient on time trend	Implied coefficient of technological change
Eq. (1')	0.641	-0.005	$\lambda_K = -1.34\%$
Eq. (2')	0.872	0.002	$\lambda_L = 1.85\%$
Eq. (3')	0.687	0.010	$\lambda_L - \lambda_K = 3.08\%$
Eq. (4')	0.807	0.004	$\lambda_K = -1.56\%$
Eq. (5')	0.892	-0.002	$\lambda_L = 1.90\%$
Eq. (6')	0.797	-0.008	$\lambda_L - \lambda_K = 3.15\%$

As discussed above, the imposition of a particular structure on the form of technological change is dictated by the need to identify the elasticity of substitution. Given this constraint, the choice of constant exponential growth rates of factor efficiency seems a natural one. Alternatively, one could consider a specification that allowed for a stochastic component in technological efficiency. In particular, consider the case in which  $A_t^K = A_0^K e^{\lambda_K \cdot t + \xi_t^K}$  and

$A_t^L = A_0^L e^{\lambda_L \cdot t + \xi_t^L}$ . It is straightforward to check that the stochastic components  $\xi_t^K$  and  $\xi_t^L$  would enter as disturbance terms in the first-order conditions that form the basis of the empirical specifications. In section 2, these error terms had been justified appealing to optimization errors on the part of firms. Technological shocks provide a second plausible explanation for these terms. Notice, however, that omitted efficiency shocks are much more likely to be correlated with factor demands than optimization errors are. The reason is that although  $\xi_t^K$  and  $\xi_t^L$  are unobserved by the econometrician, they may be in the information set of firms, which will take them into account in choosing input demands. This suggests that, to the extent that firms respond quickly to productivity shocks, my estimates of the elasticity might be biased. Remember, however, that the generalized instrumental variable (GIV) estimates in Table 5 (column III) are remarkably close to those obtained under Saikkonen's (1991) method. To the extent that the instruments used in the GIV estimation (U.S. population, wages in the government sector, and real capital stock owned by the government) are unresponsive to technology shocks but are correlated with the variables in equations (1') through (6'), the results in Table 5 indicate that ruling out stochastic components in  $A_t^K$  and  $A_t^L$  does not have a sizeable effect on the estimates of the elasticity.<sup>37</sup>

## 6 Conclusion

This paper has argued that a Cobb-Douglas specification of the U.S. aggregate production function may be misleading. My estimates suggest that, controlling for biased technological change, the elasticity of substitution between capital and labor is likely to be considerably below one, and may even be lower than 0.5. This contrasts with the results of Berndt (1976), who reported estimates of the elasticity insignificantly different from one under the assumption of Hicks-neutral technological change. I have shown, however, that ignoring the bias in technological change puts the data in a straightjacket that naturally leads to an acceptance of the null hypothesis of a unit elasticity of substitution between capital and labor. I illustrated this source of bias by showing that, in my sample, ignoring biased technological change also leads to estimates of the elasticity insignificantly different from one.

---

<sup>37</sup>I have also experimented with an alternative specification of  $A_t^K$  and  $A_t^L$  that allows for different growth rates of factor-bias in different subperiods. This amounts to including dummy variables for different subperiods, e.g., 1948-1960, 1961-1972, 1973-1984, 1985-1998. This leads to slightly lower estimates of the elasticity for most specifications and estimation techniques, but the results are very similar to those reported in Table 5.

The results are relevant for the debates on the sources of economic growth, as well as for the debate on the effects of tax behavior on investment. Hsieh (2000) shows that when the elasticity of substitution between capital and labor is lower than one, standard growth accounting exercises tend to understate the role of productivity growth as a determinant of economic growth. Similarly, as pointed out by Eisner and Nadiri (1968), low values of the elasticity of substitution imply effects of tax policy on investment behavior that are significantly lower than the ones advocated by Hall and Jorgenson (1967), who considered only the Cobb-Douglas case.

Although the analysis has been conducted with only U.S. data, the findings of this paper lend support to a recent literature that has pushed the view that certain cross-country stylized patterns are not reconcilable with aggregate output being represented by an aggregate production function featuring a unit elasticity of substitution (e.g., Acemoglu, 2002, Caselli and Coleman, 2003, and Jones, 2003). Consistent with my findings, Gollin (2002) shows that even after adjusting labor income to include self-employment income, employee compensation as a share of GDP differs substantially across countries. Furthermore, other studies have documented large movements of the labor share in OECD countries and have related these movements to the capital-labor ratio (e.g., Blanchard, 1997 and Bentolila and Saint-Paul, 2003). My estimates suggest that even for a country, the United States, with a relatively stable labor share, the evidence seems to reject a Cob-Douglas specification of the aggregate production function.

## Colophon

*Acknowledgments:* I am grateful to Daron Acemoglu, Lucia Breierova, Robert Chirinko, Jerry Hausman, Rainer Klump, two anonymous referees and the Editor (Chad Jones) for very helpful comments, and to David Romer for encouraging me to pursue this project. Financial support from the Bank of Spain is gratefully acknowledged. All remaining errors are my own.

*Email:* pantras@fas.harvard.edu.

*Correspondence:* Department of Economics, Harvard University, Littauer 230, Cambridge, MA 02138.

## Appendix A

This Appendix discusses Fair's (1970) GIV procedure. The model I consider can be summarized by the following two equations:

$$\begin{aligned} y_t &= \alpha + \beta x_t + \varepsilon_t ; \quad t = 1, \dots, T \\ \varepsilon_t &= \rho \varepsilon_{t-1} + u_t ; \quad t = 1, \dots, T, \end{aligned}$$

where  $E[\varepsilon_t | x_t] \neq 0$ , but there exists a vector of instrumental variables  $Z$  that satisfies  $\text{plim} \frac{1}{T} Z_t' x_t \neq 0$  and  $\text{plim} \frac{1}{T} Z_t' \varepsilon_t = 0$ . Fair (1970) showed that this type of model could be estimated consistently and efficiently by the following three-step procedure:<sup>38</sup>

1. First, project  $x_t$  on the set of instruments  $Z$  and compute the fitted values  $\hat{x}_t = Z_t(Z_t'Z_t)^{-1}Z_t'x_t$ . Estimate the equation  $y_t = \alpha + \beta\hat{x}_t + v_t$ , thus obtaining consistent estimates  $\tilde{\alpha}$  and  $\tilde{\beta}$ .

2. Letting  $\hat{\varepsilon}_t = y_t - \tilde{\alpha} - \tilde{\beta}x_t$ , obtain an estimate  $\hat{\rho}$  in a regression of  $\hat{\varepsilon}_t$  on  $\hat{\varepsilon}_{t-1}$ .

3. Estimate the equation  $y_t - \hat{\rho}y_{t-1} = (\hat{x}_t - \hat{\rho}x_{t-1})\beta + u_t$  and obtain  $\hat{\beta}$ , the final estimate.

The only additional requirement needed to achieve consistency is to include the lagged values of both  $Y$  and  $X$  in the set of instruments. Fair also showed that an asymptotically consistent estimate of the variance of  $\hat{\beta}$  is given by

$$\text{Est.Var}[\hat{\beta}] = \frac{\frac{1}{T} \sum_{t=2}^T \left( y_t - \hat{\rho}y_{t-1} - (x_t - \hat{\rho}x_{t-1})\hat{\beta} \right)^2}{\sum_{t=2}^T (w_t - \bar{w})^2}$$

<sup>38</sup>Fair's (1970) procedure is in fact much more general than the simplified example presented here.

where  $w_t = \hat{x}_t - \hat{\rho}x_{t-1}$ .

## Appendix B

This Appendix discusses Saikkonen's (1991) two-step procedure. If the initial model is

$$y_t = \alpha + \beta x_t + \lambda t + \varepsilon_t,$$

Saikkonen's (1991) two-step procedure first involves adding leads and lags of the first difference of the independent variable to this equation, and then modifying the standard error of the estimate of  $\beta$  making use of the estimated residuals, whose stochastic process is approximated by an AR(p) process. More specifically, the procedure can be summarized as follows:

1. Run  $y_t = \alpha + \beta x_t + \lambda t + \sum_{s=-l}^l \gamma_l \Delta x_{t-l} + u_t$ , compute  $\hat{u}_t$  and construct  $s^2 = \sum_t \hat{u}_t^2 / (T - K)$ , where  $K$  is the number of regressors including the constant.
2. Run  $\hat{u}_t = \phi_1 \hat{u}_{t-1} + \phi_2 \hat{u}_{t-2} + \dots + \phi_p \hat{u}_{t-p} + \eta_t$  and construct

$$\lambda = \frac{1}{T - p} \left( \frac{\sum_{t=p+1}^T \hat{\eta}_t^2}{1 - \hat{\phi}_1 - \dots - \hat{\phi}_p} \right).$$

Saikkonen (1991) showed that the estimate of  $\beta$  obtained in step 1 is both consistent and asymptotically efficient. More importantly, if  $t$  is the t-statistic corresponding to a test on the estimate of  $\beta$ , Saikkonen (1991) proved that:<sup>39</sup>

$$t \cdot \frac{\sqrt{s^2}}{\lambda} \xrightarrow{d} N(0, 1).$$

---

<sup>39</sup>Stock and Watson (1993) also proved that this procedure yields estimates that are asymptotically equivalent to those obtained with Johansen's (1988, 1991) Gaussian Maximum Likelihood technique.

## References

- Acemoglu, Daron K. (2002), "Directed Technical Change," *Review of Economic Studies*, 69, pp. 781-810.
- Acemoglu, Daron K. (2003), "Labor- and Capital-Augmenting Technical Change," *Journal of the European Economic Association*, 1, pp. 1-37.
- Antràs, Pol (2003), "Is the U.S. Aggregate Production Function Cobb-Douglas? New Estimates of the Elasticity of Substitution," Harvard University, mimeo.
- Arrow, Kenneth J., Hollis B. Chenery, Bagicha Minhas, and Robert M. Solow (1961), "Capital-Labor Substitution and Economic Efficiency," *Review of Economics and Statistics*, 43:5, pp. 225-254.
- Barro, Robert, and Xavier Sala-i-Martin (1995), *Economic Growth*, Cambridge, MA: MIT Press.
- Basu, Susanto (1996), "Procyclical Productivity: Increasing Returns or Cyclical Utilization?," *Quarterly Journal of Economics*, 111:3, pp. 719-751.
- Bentolila, Samuel, and Gilles Saint-Paul (2003), "Explaining Movements in the Labor Share", *Contributions to Macroeconomics*, 3:1, Article 9, <http://www.bepress.com/bejm/contributions/vol3/iss1/art9>.
- Berndt, Ernst (1976), "Reconciling Alternative Estimates of the Elasticity of Substitution," *Review of Economics and Statistics*, 58:1, pp. 59-68.
- \_\_\_\_\_ (1991), *The Practice of Econometrics: Classic and Contemporary*. Reading, Mass.: Addison-Wesley.
- \_\_\_\_\_ and Laurits R. Christensen (1973), "The Translog Function and the Substitution of Equipment, Structures and Labor in U.S. Manufacturing 1929-68," *Journal of Econometrics*, March, pp. 81-114.
- Blanchard, Olivier (1997), "The Medium Run," *Brookings Papers on Economic Activity*, pp. 89-158.
- Blough, Stephen R. (1992), "Spurious Regressions, with AR(1) Correction and Unit Root Pretest," Johns Hopkins University, Mimeo.



- Bound, John, David A. Jaeger and Regina M. Baker (1995), "Problems with Instrumental Variables Estimation When the Correlation Between the Instruments and the Endogeneous Explanatory Variable is Weak," *Journal of the American Statistical Association*, 90:430, pp. 443-450.
- Burnside, Craig, Martin Eichenbaum, and Sergio Rebelo (1995), "Capital Utilization and Returns to Scale," *NBER Macroeconomics Annual 1995*, pp. 67-110.
- Chirinko, Robert S. (2002), "Corporate Taxation, Capital Formation, and the Substitution Elasticity between Capital and Labor," *National Tax Journal*, 55:2, pp. 339-355.
- Coen, R. M. (1969), "Tax Policy and Investment Behavior: A Comment," *American Economic Review*, 59:3, pp. 370-379.
- David, Paul A. and Th. van de Klundert (1965), "Biased Efficiency Growth and Capital-Labor Substitution in the U.S., 1899-1960," *American Economic Review*, 55, pp. 357-393.
- Dhrymes, P. J., and P. Zarembka (1970), "Elasticities of Substitution for for Two-Digit Manufacturing Industries: A Correction," *Review of Economics and Statistics*, February, pp. 115-117.
- Diamond, Peter, Daniel McFadden, and Miguel Rodriguez (1978), "Measurement of the Elasticity of Factor Substitution and Bias of Technical Change," in *Production Economics: A Dual Approach to Theory and Applications (Volume 2)*, edited by Melvyn Fuss and Daniel McFadden, Amsterdam: North-Holland, pp. 125-147.
- Diewert, W. Erwin (1971), "An Application of the Shepard Duality Theorem: A Generalized Linear Production Function," *Journal of Political Economy*, 79:3, pp. 482-50.
- Eisner, R., and M. I. Nadiri (1968), "Investment Behavior and the Neoclassical Theory," *Review of Economics and Statistics*, 50:3, pp. 369-382.
- Engle, Robert F. and Clive W. J. Granger (1987), "Co-Integration and Error Correction: Representation, Estimation, and Testing," *Econometrica*, 55:2, pp. 251-276.
- Gollin, Douglas (2002), "Getting Income Shares Right," *Journal of Political Economy*, 110:2, pp. 458-474.

- Gordon, R. J. (1990), *The Measurement of Durable Goods Prices*, National Bureau of Economic Research Monograph Series, Chicago: University of Chicago Series.
- Granger, Clive W. J. and Paul Newbold (1974), "Spurious regressions in econometrics," *Journal of Econometrics*, 2, pp. 111-120.
- Greene, William H. (1999), *Econometric Analysis*, fourth edition, New York: MacMillan.
- Fair, Ray C. (1970), "The Estimation of Simultaneous Equation Models with Lagged Endogenous Variables and First Order Serially Correlated Errors," *Econometrica*, 38:3, pp. 507-516.
- Fraumeni, Barbara M. (1997), "The Measurement of Depreciation in the NIPA's," *Survey of Current Business*, July, Bureau of Economic Activity.
- Hall, Robert E. and Dale W. Jorgenson (1967), "Tax Policy and Investment Behaviour", *American Economic Review*, 57, pp. 391- 414.
- Hamilton, James (1994), *Time Series Analysis*, Princeton, NJ: Princeton University Press.
- Hausman, Jerry A. (1978), "Specification Tests in Econometrics," *Econometrica*, 46:6, pp. 1251-1271.
- \_\_\_\_\_ (1983), "Specification and Estimation of Simultaneous Models," in Z. Griliches and M. Intriligator (eds.), *Handbook of Econometrics*, Vol. 1, 1983, pp. 391-448.
- Herman, Shelby W. (2000), "Fixed Assets and Consumer Durable Goods—Estimates for 1925-98 and New NIPA Table: Changes in Net Stock of Produced Assets," *Survey of Current Business*, April, Bureau of Economic Activity.
- Hsieh, Chang-Tai (2000), "Measuring Biased Technological Change," Princeton University, mimeo.
- Johansen, Soren (1988), "Statistical Analysis of Cointegration Vectors," *Journal of Economic Dynamics and Control*, 12, pp. 231-54.
- \_\_\_\_\_ (1991), "Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models," *Econometrica*, 59:6, pp. 1551-1580.

- \_\_\_\_\_ and Katarina Juselius (1990), "Maximum Likelihood Estimation and Inference on Cointegration - With Applications to the Demand for Money," *Oxford Bulletin of Economics and Statistics*, 52, pp. 169-210.
- Jones, Charles I. (2003), "Growth, Capital Shares, and a New Perspective on Production Functions," University of California at Berkeley, mimeo.
- Jorgenson, Dale W. and Mun S. Ho (2000), "The Quality of the U.S. Workforce 1948-95," Harvard University, mimeo with data update through 1999.
- \_\_\_\_\_ and Kevin J. Stiroh (2000), "Raising the Speed Limit: U.S. Economic Growth in the Information Age," *Brookings Papers on Economic Activity*, 1, pp. 125-211.
- Judd, Kenneth L. (1987), "The Welfare Cost of Factor Taxation in a Perfect-Foresight Model," *Journal of Political Economy*, 95:4, pp. 675-709.
- Kalt, Joseph P. (1978), "Technological Change and Factor Substitution in the United States: 1929-1967," *International Economic Review*, 19, pp.761-775.
- King, Robert G., Charles I. Plosser, James H. Stock and Mark W. Watson (1991), "Stochastic Trends and Economic Fluctuations," *American Economic Review*, 81:4, pp. 819-840.
- Klump, Rainer and Olivier De La Grandville (2000), "Economic Growth and the Elasticity of Substitution: Two Theorems and Some Suggestions," *American Economic Review*, 90:1, pp. 282-291.
- Krueger, Alan B. (1999), "Measuring Labor's Share," *American Economic Review*, 89:2, pp. 45-51.
- Krusell, Per, Lee Ohanian, Jose Victor Rios-Rull and Gianluca Violante (2000), "Capital-Skill Complementarity and Inequality: A Macroeconomic Analysis", *Econometrica*, 68, pp. 1029-1054.
- Lucas, Robert E. (1969), "Labor-Capital Substitution in U.S. Manufacturing," in *The Taxation of Income from Capital*, edited by Arnold Harberger and Martin J. Bailey, Washington, DC: The Brookings Institution, pp. 223-274.

- MacKinnon, J. (1991), "Critical Values for Co-Integration Tests," in R.F. Engle and C.W.J. Granger (eds.) *Long-Run Economic Relationships*, Oxford University Press, pp. 267-76.
- Maddala, G. S. (1965), "Differential Industry Effects and Differential Factor Effects of Technological Change," Memo 36, Research Center in Economic Growth, Stanford University.
- Mankiw, Gregory N., David Romer, David N. Weil (1992), "A Contribution to the Empirics of Economic Growth," *Quarterly Journal of Economics*, 107:2, pp. 407-437.
- Nerlove, Marc (1963), "Returns to Scale in Electricity Supply," in Carl Christ, ed., *Measurement in Economics: Studies in Mathematical Economics and Econometrics in Memory of Yehuda Grunfeld*, Stanford, CA: Stanford University Press, pp. 167-198.
- \_\_\_\_\_ (1967), "Recent Empirical Studies of the CES and Related Production Functions." In M. Brown, editor, *The Theory and Empirical Analysis of Production*, New York: Columbia University Press for Studies in Income and Wealth Series 32, pp. 55-122.
- Phillips, Peter C. B. (1986), "Understanding Spurious Regressions in Econometrics," *Journal of Econometrics*, 33, pp. 311-40.
- \_\_\_\_\_ and Stephen Durlauf (1986), "Multiple Time Series Regression with Integrated Processes," *Review of Economic Studies*, 53:4, pp. 473-495.
- Saikkonen, Pentti (1991), "Asymptotically Efficient Estimation of Cointegration Regressions," *Econometric Theory*, 7, pp. 1-21.
- Stock, James H. (1987), "Asymptotic Properties of Least Squares Estimators of Cointegrating Vectors," *Econometrica*, 55:5, pp. 1035-1056.
- Trostel, Philip A. (1993), "The Effect of Taxation on Human Capital," *Journal of Political Economy*, 101:2, pp. 327-350.