Cowles Foundation Paper 181

Econometrica, Vol. 30, No. 2 (April, 1962)

ON DEVISING UNBIASED ESTIMATORS FOR THE PARAMETERS OF THE COBB-DOUGLAS PRODUCTION FUNCTION[1]

BY PHOEBUS J. DHRYMES

1.

IN MANY empirical investigations the geometric mean of observed factor shares serves as an estimator of the factor exponents in a Cobb-Douglas[2] production function. This particular estimator is due to Klein [**7**, p. 194]. It is here shown that such an estimator is biased. Under certain conditions an alternative estimator is derived which is unbiased, sufficient, efficient, and consistent.

2.

The following description for the production process of a firm, sector, or economy is often employed:

$$Q(t) = A(t) \prod_{i=1}^{n} X_i^{\alpha_i}(t) \, e^{u(t)} \qquad (t = 1,2,\ldots,T) \tag{1}$$

where $Q(t)$, $X_i(t)$ are respectively the output and ith input at time t, $A(t)$ is some positive function of time, $\alpha_i > 0$, and $\sum_{i=1}^{n} \alpha_i = 1$, this latter preserving the condition imposed on (1) when $u(t) \equiv 0$, i.e., when it is nonstochastic, due to the requirements of competitive theory. By virtue of the latter assumption, we may also write:

$$P_i(t) = P_Q(t) \frac{\partial Q^*(t)}{\partial X_i(t)} e^{v_i(t)} (i = 1,2,\ldots,n\,;\, t = 1,2,\ldots,T) \tag{2}$$

where $P_i(t)$, $P_Q(t)$ are, respectively, the prices of the ith factor and of output at time t and $Q^*(t) = Q(t)e^{-u(t)}$, i.e., it is the nonstochastic part of (1). A variant of this formulation is given, e.g., by Wolfson [**13**]. What (2) implies

[1] This is a revised version of part of my Thesis submitted at the Massachusetts Institute of Technology. During its preparation I had the benefit of helpful advice by Professors E. Kuh and R. M. Solow. I have also benefited from several suggestions by the referee of this paper. To all I wish to express my gratitude. Acknowledgment is also due to the Ford Foundation which partially financed this investigation through a Doctoral Dissertation Fellowship and to the Cowles Foundation at which the present version was completed during my visit there in the summer of 1961.

[2] The name "Cobb-Douglas" traditionally given to such functions is somewhat inappropriate. Wicksell derives these functions from elementary economic considerations in the course of his analysis of Akerman's problem. (Wicksell, K, *Lectures on Political Economy*, Vol. II, Routledge and Kegan Paul, Ltd., London, 1951, pp. 285–286.)

is that decisions on factor employment are made on the basis of the anticipated output $Q^*(t)$ and not on the basis of the output actually materializing.

The term $u(t)$ in (1) and $v_i(t)$ in (2) are assumed to be serially uncorrelated and to be distributed (the latter marginally) normally with mean zero and variance σ^2 and $\sigma^2_{v_1}$, respectively. We may also assume that $u(t)$ and $v_i(t)$ are independent.

Manipulating (1) and (2) it is then an easy matter to show that

$$\frac{P_i(t)X_i(t)}{P_Q(t)Q(t)} = \alpha_i(t) = \alpha_i e^{w_i(t)} (i=1,2,\ldots,n\,;\, t=1,2,\ldots,T) \tag{3}$$

where $w_i(t) = u(t) - v_i(t)$. Under the assumption made above, the $w_i(t)$ are (marginally) normal with mean zero and variance $\sigma_i^2 = \sigma^2 + \sigma^2_{v_i}$; they are also serially uncorrelated.

Now, if the $\alpha_i(t)$ are interpreted as observed factor shares, then the factor exponents in (1) can be estimated by first transforming into logarithms. Thus:

$$\log \alpha_i(t) = \log \alpha_i + w_i(t) \quad (i=1,2,\ldots,n\,;\, t=1,2,\ldots,T). \tag{4}$$

From (4), minimizing the second moment of estimated disturbances gives

$$\log \hat{\alpha}_i = \frac{1}{T} \sum_{i=1}^{T} \log \alpha_i(t) \qquad (i = 1, 2, \ldots, n) \tag{5}$$

where T is the sample size.

On the hypotheses maintained, the estimators (5) are unbiased, efficient, and consistent. This is due to the Markov theorem. Because of the (marginal) normality of the $w_i(t)$, the estimators (5) are also sufficient, since they are maximum likelihood estimators.

Now it is obvious that equations (5) imply a certain set of estimators for the α_i, *viz.*,

$$\hat{\alpha}_i = \left[\prod_{t=1}^{T} \alpha_i(t)\right]^{1/T} \qquad (i = 1, 2, \ldots, n). \tag{6}$$

Thus a rather convenient estimator is deduced which requires only readily available information, obviating or at least ameliorating the problems involved in obtaining accurate information concerning, in particular, capital inputs.

It is also occasionally claimed in the literature, e.g., Hoch [**4**, p. 567] and Walters [**11**, p. 26] that the estimators (6) are, on the hypotheses maintained except for normality, "best linear unbiased." This "result" is attributed to Klein [**7**, p. 193]. The ascription, however, is not correct in that Klein nowhere explicitly claims this.

The claim concerning the estimators (6) results presumably from the properties that the Markov theorem under the classical assumptions ascribes to (5) as estimators of $\log \alpha_i$.

Now, although the estimators (6) are transforms of (5) under a homeomorphism (i.e., a continuous one-one correspondence whose inverse is also (one-one and) continuous), yet not all properties ascribed to (5) can be ascribed to (6) as well.

It is a simple matter to show that the estimators (6) are sufficient essentially because of the Fisher-Neyman criterion for sufficient statistics [**5**, p. 101] and the homeomorphism involved in obtaining (6) from (5).

Consistency could also be shown to be preserved using the following heuristic argument which can be made rigorous:

$$\hat{\alpha}_i = e^{\log \hat{\alpha}_i} = f(\log \hat{\alpha}_i) \ . \tag{7}$$

Then we have

$$\operatorname*{plim}_{T\to\infty} \hat{\alpha}_i = f(\operatorname*{plim}_{T\to\infty} \log \hat{\alpha}_i) = f(\log \alpha_i) = \alpha_i \ .$$

Unbiasedness, however, is not necessarily preserved under a homeomorphism; in fact, since $f(\log \hat{\alpha}_i)$ obeys $f''(\widehat{\log \alpha_i}) > 0$ and is everywhere analytic, we have by using the finite form of Taylor's theorem (with remainder after two terms) about $\log \alpha_i$:

$$E[f(\log \hat{\alpha}_i)] \geq f[E(\log \hat{\alpha}_i)] = f(\log \alpha_i) = \alpha_i \quad (i = 1, 2, \ldots, n) \ . \tag{8}$$

In (8) the equality holds if and only if $\operatorname{var}(\widehat{\log \alpha_i}) = 0$.

This suggests that the estimators (6) are asymptotically unbiased. By virtue of the asymptotic unbiasedness and sufficiency of the estimators (6), it follows that they are asymptotically efficient. The term "efficient" here means the property of being unbiased and of minimum variance within the class of unbiased estimators of the parameter in question.

The asymptotic efficiency of (6) is essentially due to the completeness[3] of the density function of the mean of a sample derived from a normal parent.

In the present case we may show asymptotic unbiasedness directly by utilizing the fact that $E(\alpha_i) = \varphi_i(1)$ where $\varphi_i(1)$ is the moment generating function of $\widehat{\log \alpha_i}$ with parameter 1:

$$E(\hat{\alpha}_i) = \alpha_i e^{\frac{1}{2T}\sigma_i^2} > \alpha_i \qquad (i = 1, 2, \ldots, n) \ . \tag{9}$$

The result in (9) points out another difficulty with this particular formulation, *viz.*,

$$E\left[\sum_{i=1}^{n} \hat{\alpha}_i\right] > \sum_{i=1}^{n} \alpha_i = 1 \ . \tag{10}$$

[3] For an elementary development of the notion of completeness, see Hogg and Craig [**5**, p. 106]. For a full development see Lehmann and Scheffé [**8**, p. 305 ff.].

We may wish to preserve the requirement $\Sigma_{i=1}^{n} \alpha_i = 1$ so that the model would remain a true generalization of the nonstochastic model of production and distribution. This may be accomplished by treating one of the factors, say the nth, asymmetrically by requiring its share to accrue residually.

This has the following consequences: In (2) we do not have n conditions, but rather $n-1$, so that optimization occurs only with respect to $n-1$ factors; in (5) and consequently in (6) we have only $n-1$ independent estimators, so that we always put by definition:

$$\hat{\alpha}_n \equiv 1 - \sum_{i=1}^{n-1} \hat{\alpha}_i . \tag{11}$$

What follows then should be interpreted in the light of (11). From (9) it is easily seen that the inverse of the relative bias of (6) is given by:

$$e^{-\frac{1}{2T}\sigma_i^2} = \sum_{k=0}^{\infty} \left(-\frac{\sigma_i^2}{2T}\right)^k / k! \qquad (i = 1, 2, \ldots, n-1) . \tag{12}$$

Bearing this in mind we have:

LEMMA 1: *There exist unbiased estimators for the α_i, $i = 1, 2, \ldots, n-1$.*

Proof. We proceed constructively: Since $\hat{\log \alpha_i}$ and S_i^2 (the sample mean and variance of the logarithms of $\alpha_i(t)$) are distributed independently, and since

$$E(S_i^{2k}) = \frac{2^k \sigma_i^{2k} \Gamma\left(\frac{T-1}{2} + k\right)}{T^k \Gamma\left(\frac{T-1}{2}\right)} , \tag{12a}$$

it will suffice to determine a function $f(S_i^2)$ such that $E[f(S_i^2)] = e^{-\frac{1}{2T}\sigma_i^2}$. Writing formally

$$f(S_i^2) = \sum_{k=0}^{\infty} a_k S_i^{2k} \qquad (i = 1, 2, \ldots, n-1) , \tag{13}$$

then taking the expectation of (13) with respect to the density function of S_i^2 and equating coefficients of like powers of σ_i^2 in (12) we find

$$f(S_i^2) = \sum_{k=0}^{\infty} (-\tfrac{1}{4})^k \frac{B\left(\frac{T-1}{2}, k\right)}{\Gamma(k)\, k!} S_i^{2k} \quad (i=1,2,\ldots,n-1) . \tag{13a}$$

It follows[4] that:

$$\bar{\alpha}_i = \hat{\alpha}_i f(S_i^2) \qquad (i = 1, 2, \ldots, n-1) \tag{14}$$

are unbiased estimators of the α_i, *q.e.d.*

LEMMA 2: *The estimators $\bar{\alpha}_i$ in* (14) *are sufficient.*

Proof. This follows from the fact that the $\bar{\alpha}_i$ are simple functions of the (jointly) sufficient statistics $\log \hat{\alpha}_i$ and S_i^2 alone, i.e., not involving unknown parameters.

LEMMA 3: *The estimators $\bar{\alpha}_i$ in* (14) *are efficient.*

Proof. The joint density function of $\log \hat{\alpha}_i$ and S_i^2 is complete;[5] due to a

[4] Expectation of (13) is justified since it is uniformly convergent for $S_i^2 < \infty$; further $f(S_i^2)$ is an analytic function of S_i^2, hence independent of $\hat{\alpha}_i$, so that $E[\hat{\alpha}_i f(S_i^2)] = E(\hat{\alpha}_i) E[f(S_i^2)] = \alpha_i$.

[5] It will be noted that the properties ascribed to the estimators $\bar{\alpha}_i$ are deduced from the characteristics of the marginal distribution of the $w_i(t)$, $i = 1, 2, \ldots, n-1$, and not from those of the joint distribution of the $w_i(t)$. This is in a sense the logical consequence of the fact that the problem as it naturally arises involves the marginal distribution of the $v_i(t)$, $i = 1, 2, \ldots, n-1$, exhibited in equation (2). It is not, however, difficult to derive the joint distribution of the $w_i(t)$. This is seen as follows:

Let the joint distribution of the $v_i(t)$, $i = 1, 2, \ldots, n-1$, be multivariate normal with mean vector 0 and covariance matrix Σ. We still retain the assumption that the $v_i(t)$ are serially uncorrelated but we shall not insist on the independence of the $v_i(t)$ for different indices i. Then consider the transformation:

$$w_i^*(t) = u(t) - v_i(t) = w_i(t) \qquad (i = 1, 2, \ldots, n-1),$$

$$w_n^*(t) = u(t),$$

or for simplicity

$$W^* = A V^*,$$

where W^* is a vector whose components are the $w_i^*(t)$, $i = 1, 2, \ldots, n$, and V^* is the vector whose first component is $u(t)$ and whose remaining $n-1$ components are the $v_i(t)$, $i = 1, 2, \ldots, n-1$, A being the matrix of the transformation. Since $u(t)$ is independent of the $v_i(t)$, it follows that V^* has the multivariate normal distribution with mean vector 0 and covariance matrix Σ^*, where

$$\Sigma^* = \begin{bmatrix} & & 0 \\ & \Sigma & \vdots \\ & & 0 \\ 0 & \cdots\, 0 & \sigma^2 \end{bmatrix}.$$

Hence, since A is nonsingular, W^* has the (proper) multivariate normal distribution with mean vector 0 and covariance matrix $A\ \Sigma^*\ A'$. But then it is obvious, by "integrating out" $w_n^*(t)$, that the $w_i(t)$, $i = 1, 2, \ldots, n-1$, have the joint multivariate normal distribution with mean vector zero and covariance matrix obtained by deleting the last row and column from $A\ \Sigma^* A'$. It then, of course, follows that the $w_i(t)$ have result by Rao [**9**, p. 86] if an efficient estimator exists it must be an explicit function of the (jointly) sufficient statistics $\log\hat{\alpha}_i$ and S_i^2 alone. Thus, if another continuous[6] estimator

$$\varrho_i = \varrho_i(\log\hat{\alpha}_i, S_i^2) \qquad (i = 1, 2, \ldots, n-1) \tag{15}$$

exists which is also unbiased, then

$$E[\bar{\alpha}_i - \varrho_i] = 0 . \tag{15a}$$

Due to the completeness of the joint density function of $\log\hat{\alpha}_i$ and S_i^2 it follows that $\bar{\alpha}_i \equiv \varrho_i$, *q.e.d.*

LEMMA 4: *The estimators $\bar{\alpha}_i$ in* (14) *are consistent.*

Proof. Had efficiency been defined by the Cramer-Rao equation, [**2**, p. 481] the proof would have been immediate. Since we have not employed this definition, it seems simpler to give a direct argument from the variance of the $\bar{\alpha}_i$. Thus:

$$\operatorname{Var}(\bar{\alpha}_i) = e^{2\log\alpha_i}\left[e^{\frac{2\sigma_i^2}{T}} V_i(f) - 1\right] \qquad (i = 1, 2, \ldots, n-1) \tag{16}$$

where

$$V_i(f) = E[f(S_i^2)]^2 \qquad (i = 1, 2, \ldots, n-1) . \tag{16a}$$

It can be shown that

$$\lim_{T\to\infty} V_i(f) = 1 . \tag{16b}$$

Hence (16b) together with Lemma 1 and Chebyshev's inequality imply consistency, *q.e.d.*

We may conclude this section by formally stating the theorem:

THEOREM: *Under the hypotheses* (1) *and* (2) *as modified by* (11) *the estimators* (6) *are biased*; *the estimators given in* (14) *are unbiased, sufficient, efficient, and consistent.*[7]

---

individually the univariate normal (marginal) distribution as claimed earlier. One could if one wished characterize the properties of the $\bar{a}_i$ on the basis of their joint distribution, or derive appropriate estimators by first transforming the $w_i(t)$ so that they become independent. This, however, would represent a departure from the natural way in which the problem is posed.

[6] Continuity is assumed only in order to disallow the possibility that $\bar{a}_i \neq \varrho_i$ on a set of measure zero.

[7] In the interest of completeness one should remark that these attributes are derived from the properties of the marginal distribution of the $v_i(t)$ — and hence of the $w_i(t)$. Of course if $u(t) \equiv 0$ and the $v_i(t)$ are independent, then this qualification would be unnecessary.

3. CONCLUSION

We give below an indication of the size of the bias in estimating the α_i from share-of-capital data in various sectors of the American Economy for the postwar period (1945–1958). The entries in Table I are computed values of $f(S_i^2)$ truncated after 8 terms.

TABLE I

Sector	truncated $f(S_i^2)$
Manufacturing	.996
Transportation	.976
Services	.999

As is apparent from (9) the relative bias will be small if the variance of $\log \hat{\alpha}_i(t)$ is also small.

We may be concerned with unbiasedness, however, if we wish to use estimators of the α_i in obtaining an estimator of a productivity parameter as follows:

It will be noted that the term $A(t)$ in (1) has not appeared in subsequent developments. If a specific functional from were to be employed, say, $A(t) = Ae^{\lambda t}$, then using the estimators $\bar{\alpha}_i$ we easily find:

$$(17) \qquad Q(t) \prod_{i=1}^{n} X_i^{-\bar{\alpha}_i}(t) = A e^{\lambda t} \prod_{i=1}^{n} X_i^{-(\bar{\alpha}_i - \alpha_i)}(t) e^{u(t)} \quad (t = 1, 2, \ldots, T) .$$

The right member of (17) involves the random term

$$\prod_{i=1}^{n} X_i^{-(\bar{\alpha}_i - \alpha_i)}(t) e^{u(t)} .$$

Hence we could derive an estimator for λ and study its properties. In this context it would be desirable to have unbiased estimators for the α_i. An application of this will be made in a subsequent paper.

This procedure gives a statistical counterpart to the method suggested by Solow [**10**, p. 312] and provides a rather simple means of testing a statistical hypothesis on the "rate of change" of productivity as between two economies, firms, or sectors whose productive processes are characterized by a relation such as (1).

Harvard University

REFERENCES

[1] AITCHISON, J., AND J. A. C. BROWN: *The Lognormal Distribution*, Cambridge: University Press, 1957.

[2] CRAMER, H.: *Mathematical Methods of Statistics*, Princeton: Princeton University Press, 1946.

[3] FINNEY, D. J.: "On the Distribution of a Variate Whose Logarithm is Normally Distributed," *Journal of the Royal Statistical Society*, Supplement I, Vol. VII, No. 2, 1941, p. 155.

[4] HOCH, I.: "Simultaneous Equation Bias in the Context of the Cobb-Douglas Function," *Econometrica*, Vol. 29, 1958, p. 566.

[5] HOGG, R. V., AND A. T. CRAIG: *Introduction to Mathematical Statistics*, New York: The Macmillan Co., 1959.

[6] ———: "Sufficient Statistics in Elementary Distribution Theory," *Sankhya*, Vol. 17, 1956–57, p. 209.

[7] KLEIN, L. R.: *A Textbook of Econometrics*, Evanston: Row, Peterson and Co., 1953.

[8] LEHMANN, E. L., AND H. SCHEFFÉ: "Completeness, Similar Regions and Unbiased Estimation," Part I, *Sankhya*, Vol. 10, 1950, p. 305.

[9] RAO, C. RADHAKRISHNA: "Information and Accuracy Attainable in the Estimation of Statistical Parameters," *Bulletin of the Calcutta Mathematical Society*, Vol. 37, 1945, p. 81.

[10] SOLOW, R. M.: "Technical Change and the Aggregate Production Function," *Review of Economics and Statistics*, Vol. 39, 1957, p. 312.

[11] WALTERS, A. A.: *Some Notes on Simultaneous Equations and the Cobb-Douglas Function*, Series A, No. 17, University of Birmingham, Faculty of Commerce and Social Science, June, 1960.

[12] ———: *A Survey of Statistical Production and Cost Functions*, Series A, No. 20, University of Birmingham, Faculty of Commerce and Social Science, June, 1960.

[13] WOLFSON, R. J.: "An Econometric Investigation of Regional Differentials in American Agricultural Wages," *Econometrica*, 1959, pp. 225–257.