# ON ESTIMATION OF THE CES PRODUCTION FUNCTION*

## BY J. KMENTA[1]

### 1. SINGLE EQUATION ESTIMATES

THE ORIGINAL SPECIFICATION of the constant-elasticity-of-substitution (CES) production function by Arrow, Chenery, Minhas, and Solow [1] was restricted to the case of constant returns to scale. With this restriction it is possible to estimate the elasticity of substitution from the marginal productivity condition by regressing the value of production per worker on wage rate (both variables measured in logarithms). If, however, the CES production function is generalized to allow for the possibility of non-constant returns to scale, this method of estimation is no longer feasible. The purpose of this paper is to consider estimation procedures applicable to the generalized version of the CES function under various circumstances.

An obvious starting point is to consider estimates obtained by fitting the production function to observations on output and inputs alone. These estimates are consistent if the input variables are non-stochastic or, if stochastic, independent of the disturbance in the production function. The CES function can be written in the form

$$(1.1) \qquad \log X_i = \log \gamma - \frac{\nu}{\rho} \log [\delta K_i^{-\rho} + (1 - \delta)L_i^{-\rho}] + u_i \ .$$

The subscript $i$ refers to the $i$-th firm, and $u_i$ is the stochastic error term assumed to be independently and normally distributed with zero mean and constant variance. This specification is analogous to that used in the context of the Cobb-Douglas function.

The parameters of (1.1) could be estimated by nonlinear least squares methods for which computer programs are now available. An alternative method, based on simple least squares estimation, is possible if we replace (1.1) by its approximation which is linear in $\rho$. This can be derived by using Taylor's formula for expansion around $\rho = 0$.[2] After disregarding the terms of third and higher orders, the expansion leads to[3]

$$(1.2) \qquad \begin{aligned} \log X_i &= \log \gamma + \nu\delta \log K_i + \nu(1 - \delta) \log L_i \\ &\quad - \frac{1}{2} \rho\nu\delta(1 - \delta)[\log K_i - \log L_i]^2 + u_i \ . \end{aligned}$$

The approximation to the CES function given by (1.2) can then be con-

---

* Manuscript received November 2, 1964, revised March 15, 1966.

[1] This research was supported by the Social Systems Research Institute at the University of Wisconsin.

[2] The results of Arrow, et al., described in [1], Table 2, column 2, suggest that values of $\rho$ tend to range from 0 to 0.4. We have chosen an expansion around 0 for its mathematical convenience.

[3] All logarithms are natural logarithms. If common logs are used, the last term on the right hand side of (1.2) before $u_i$ has to be multiplied by 2.302585.

veniently separated into two parts, one corresponding to the Cobb-Douglas form and one representing a "correction" due to the departure of $\rho$ from zero. The latter part, given by the term $- [\rho\nu\delta(1 - \delta)/2][\log K_i - \log L_i]^2$, will disappear if $\rho = 0$. The error of approximating the CES function by (1.2) depends on the extent to which $\rho$ departs from zero, on the ratio of the two inputs and on the values of the remaining parameters. Some specific calculations are given in Appendix A.1.

If we apply simple least squares regression to (1.2) we can easily calculate estimates of the parameters of the CES function and their standard errors.[4] More specifically, we can test for the Cobb-Douglas hypothesis by examining the significance of the coefficient attached to $[\log K - \log L]^2$. This test is invariant with respect to units of measurement of $K$ and $L$.

For illustrative purposes we have applied this method to the United States non-farm data for 1947 to 1960.[5] To avoid multicollinearity, we have followed the example of Arrow, $et$ $al.$ in [1] and took the value of $\delta$ as predetermined at 0.519. The resulting estimates are shown below with standard errors given in brackets:

$$\hat{\gamma} = 0.1112$$
$$\hat{\nu} = 1.1785 \quad (0.1487)$$
$$\hat{\rho} = 0.4884 \quad (0.4398) .$$

The implied value of the elasticity of substitution (0.6719) is not significantly different from unity, given the predetermined value of $\delta$. The results thus provide no evidence against the Cobb-Douglas model.

## 2. SIMULTANEOUS EQUATION ESTIMATES WITH UNIFORM PRICES

Let us now consider firms which operate under perfectly competitive conditions and obtain their inputs at fixed prices in the same market. Given that the appropriate production function is the CES function, the production model may be specified by the following relationships:

$$(2.1) \qquad \log X_i = \log \gamma - \frac{\nu}{\rho} \log [\delta K_i^{-\rho} + (1 - \delta)L_i^{-\rho}] + u_{0i} ,$$

$$(2.2) \qquad \left(\frac{\rho}{\nu} + 1\right) \log X_i - (\rho + 1) \log K_i = \log [\gamma\gamma^{\rho/\nu}(p\nu\delta)^{-1}]R_1 + u_{1i} ,$$

$$(2.3) \qquad \left(\frac{\rho}{\nu} + 1\right) \log X_i - (\rho + 1) \log L_i = \log [w\gamma^{\rho/\nu}(p\nu)^{-1}(1 - \delta)^{-1}]R_2 + u_{2i} ,$$

where $p =$ price of product, $w =$ wage rate, and $r =$ price of capital input. The model is formally equivalent to the traditional model of production analysis, except that the usual Cobb-Douglas production function has been replaced by the CES function. Equations (2.2) and (2.3) are the profit-max-

[4] For the formula of calculating (approximate) standard errors see Klein [7, (258)].
[5] The data were taken from Sato [10]. Following the suggestion of Griliches in [4], the input data have been adjusted for quality improvements at the annual rate of 1.5 per cent for capital and 1 per cent for labor, beginning with 1947.

imizing marginal productivity conditions, generalized by the inclusion of disturbances and of parametric restraints $R_1$ and $R_2$. The underlying assumption concerning the disturbance in the production function is that it is known to the firm (but not to the econometrician) *a priori*.[6]  $R_1$ and $R_2$ have been introduced to allow for the possibility of systematic deviation from profit maximization due to restrictions on firm behavior.[7]  If there are no restrictions, then $R_1 = R_2 = 1$.

The simultaneous equation model (2.1) through (2.3) is underidentified in the sense that the conditions for minimizing generalized residual variance (or, under normality, maximizing the likelihood function) do not contain enough information to solve for all the unknown parameters to be estimated. This difficulty has been also encountered in the context of the Cobb-Douglas production function and has been overcome by imposing restrictions on the variance-covariance matrix of the disturbances.[8]  We shall follow this and assume that the variance-covariance matrix of the disturbances is diagonal.

The estimation of the foregoing model could be handled by nonlinear full-information method with the help of high-speed computers.[9]  Alternatively, we may replace the CES function (2.1) by its approximation suggested in the previous section. By modifying the system in this way and by introducing a more compact terminology, we can rewrite (2.1) to (2.3) as

$$(2.4) \qquad x_{0i} - \nu\delta x_{1i} - \nu(1 - \delta)x_{2i} + \frac{1}{2}\rho\nu\delta(1 - \delta)(x_{1i} - x_{2i})^2 = k_0 + u_{0i} ,$$

$$(2.5) \qquad \left(\frac{\rho}{\nu} + 1\right) x_{0i} - (\rho + 1)x_{1i} = k_1 + u_{1i} ,$$

$$(2.6) \qquad \left(\frac{\rho}{\nu} + 1\right) x_{0i} - (\rho + 1)x_{2i} = k_2 + u_{2i} ,$$

where $x_0 = \log X$, $x_1 = \log K$, $x_2 = \log L$, and the constant terms $k_r (r = 0, 1, 2)$ are defined with reference to (2.1) through (2.3). The disturbances $u_r (r = 0, 1, 2)$ are assumed to be normally distributed with zero means and constant variances and to be "cross-sectionally" independent. With this modification the joint likelihood function becomes

$$(2.7) \qquad \begin{aligned} L = \frac{3n}{2} \log (2\pi) + n \log (\rho + 1)(1 - \nu) - \frac{n}{2} \sigma_{00}\sigma_{11}\sigma_{22} \\ - \frac{1}{2\sigma_{00}} u_0'u_0 - \frac{1}{2\sigma_{11}} u_1'u_1 - \frac{1}{2\sigma_{22}} u_2'u_2 . \end{aligned}$$

Here $u_r (r = 0, 1, 2)$ is an $n \times 1$ vector of the disturbances in the sample, and

[6] This specifications of a production model was originated by Marschak and Andrews [8]; it corresponds to Case A of Mundlak and Hoch in [9].

[7] These restrictions (if they exist) are envisaged to operate for "institutional" reasons, not because of departures from horizontality of output-demand or input-supply curves. For elaboration see Hoch [6, (35–56)] or Mundlak and Hoch [9, (815)].

[8] See, e.g. Hoch [5]. An alternative method, known as the "factor shares" method, exploits the assumption of unrestricted profit maximization. In the context of the CES production function, this method cannot be used.

[9] A program for this method is given by Eisenpress and Greenstadt and described in [3].

$\sigma_{rr} = \mathrm{E}(u'_r u_r)$. From (2.7) we can derive the maximum likelihood estimates of the parameters of the production function.

In practice, the straightforward determination of the maximum likelihood estimates turns out to be quite difficult because it requires solution of equations which are nonlinear in the unknowns. Fortunately the solution (i.e., the maximum likelihood estimates) can be obtained by using a modified indirect least squares method.[10] Let us first introduce new variables, defined as follows:

$$z_{1i} = Fx_{0i} - x_{1i} ,$$
$$z_{2i} = Fx_{0i} - x_{2i} ,$$
$$z_{3i} = (x_{1i} - x_{2i})^2 ,$$

where

$$F = \left( \frac{\rho}{\nu} + 1 \right) \Big/ (\rho + 1) .$$

Next, we form a new regression equation

(2.8) $$x_{0i} = a_0 + a_1 z_{1i} + a_2 z_{2i} + a_3 z_{2i} + a_3 z_{3i} + e_i .$$

Because of the definition of the $z$'s, the coefficients of (2.8) can be identified with those of the production function (2.4)

(2.9) $$\log \gamma = a_0/(1 - Fa_1 - Fa_2) ,$$

(2.10) $$\nu \delta = -a_1/(1 - Fa_1 - Fa_2) ,$$

(2.11) $$\nu(1 - \delta) = -a_2/(1 - Fa_1 - Fa_2) ,$$

(2.12) $$-\frac{1}{2}\rho\nu\delta(1-\delta) = a_3/(1 - Fa_1 - Fa_2) .$$

Further, the definition of $F$ together with (2.10) and (2.11) implies,

(2.13) $$a_3 = -\frac{1}{2}(F - 1)a_1 a_2 .$$

Finally, the disturbance of (2.4) is proportional to that of (2.8).

Equations (2.5) and (2.6) imply that the $z$'s depend only on $u_1$ and $u_2$. Since $u_1$ and $u_2$ are independent of $u_0$ by assumption, all the $z$'s are independent of $e_i$, and least squares estimates of (2.8), subject to the restriction of (2.13), will be consistent. Replacing the $a$'s in (2.9) through (2.12) by their consistent estimates will lead to consistent estimates of $\log \gamma$, $\nu$, $\delta$, and $\rho$.

The construction of the variables $z_1$ and $z_2$ requires *a priori* knowledge of $F$. This we do not have, but a consistent estimate of $F$, say $\hat{F}$, can be obtained by utilizing the assumption of independence of $u_1$ and $u_2$. By imposing this restriction on the sample we have from equations (2.5) and (2.6)

(2.14) $$\hat{F}^2 m_{00} - \hat{F}(m_{01} + m_{02}) + m_{12} = 0 ,$$

where $\hat{F} = (\hat{\rho}/\hat{\nu} + 1)/(\hat{\rho} + 1)$, and $m$'s are sample moments of the $x$'s. Equation

---

[10] For the proof of equivalence of the two methods, see Appendix A.2.

(2.14) has two roots; it can be shown that both will be real—at least asymptotically.[11] The question as to which of the two roots is appropriate for our purpose can be resolved with reference to equation (2.8). We wish to minimize the sum of squares of the residuals or, equivalently, to maximize the sum of squares due to regression (SSR). If $\hat{a}_1$ and $\hat{a}_2$ are both negative, the SSR would be maximized by choosing the smaller root of (2.14).

The foregoing production model applies only in a situation where the production function is completely known by the firms before the input decisions are made. If, however, the value of the production function disturbance is not known *a priori*, output—and therefore profit—is a stochastic variable. In this case one would find it more appropriate to postulate maximization of expected rather than actual profit. Given this, the estimation procedure suggested above would lead to inconsistent estimates. Consistent estimates can be obtained by using simple least squares estimates corrected for asymptotic bias.[12]

### 3. SIMULTANEOUS EQUATION ESTIMATES WITH NON-UNIFORM PRICES

In this section we shall consider estimation in the case in which prices of output and of inputs are no longer constant but are allowed to vary over the sample period. Such a situation is encountered when we have observations on a number of firms which operate in different markets, or on a single firm at different points of time. Since, under perfectly competitive conditions, individual firms have no influence on the market, prices can be regarded as exogenously determined. Using the specification given in the previous section, the model can be described as

$$(3.1) \qquad x_{0i} = \log \gamma - \nu\delta(x_{2i} - x_{1i}) + \nu x_{2i} - \frac{1}{2}\rho\nu\delta(1-\delta)(x_{2i} - x_{1i})^2 + u_{0i}\,,$$

$$(3.2) \qquad \left(\frac{\rho}{\nu} + 1\right)x_{0i} - (\rho + 1)x_{1i} = \log\left(\frac{r}{p}\right)_i + \log \gamma^{\rho/\nu}(\nu\delta)^{-1}R_1 + u_{1i}\,,$$

$$(3.3) \qquad \left(\frac{\rho}{\nu} + 1\right)x_{0i} - (\rho + 1)x_{2i} = \log\left(\frac{w}{p}\right)_i + \log \gamma^{\rho/\nu}\nu^{-1}(1-\delta)^{-1}R_2 + u_{2i}\,.$$

Our problem again is to estimate the parameters of the production function (3.1). Deducting (3.2) from (3.3) gives

$$(3.4) \qquad \begin{aligned}(x_{2i} - x_{1i}) &= \left(\frac{1}{\rho + 1}\right)\log\left(\frac{1-\delta}{\delta}\right)\frac{R_1}{R_2} - \left(\frac{1}{\rho + 1}\right)\log\left(\frac{w}{r}\right)_i \\ &\quad + \left(\frac{1}{\rho + 1}\right)(u_{2i} - u_{1i})\,.\end{aligned}$$

An application of the simple least squares method to (3.4), i.e., regressing

[11] For proof, see Appendix A.2.

[12] The suggestion to use the postulate of expected profit maximization in this case has been made by Arnold Zellner; the implications for estimation of production function parameters are discussed in detail in [11]. See also Case B2 in Mundlak and Hoch [9].

log $(L/K)$ on log $(w/r)$, will give the best linear unbiased estimate of $1/(\rho + 1)$. This, in turn, leads to a consistent estimate of $\rho$.[13]

The estimation of the remaining parameters can proceed as follows. Equation (3.4) is, in fact, the reduced form equation for $x_{2i} - x_{1i}$. One can obtain the reduced form equation for $x_{2i}$ by solving (3.1), (3.3) and (3.4) for $x_{2i}$. The result is, in general,

$$(3.5) \qquad x_{2i} = b_0 + b_1 \log \left(\frac{r}{p}\right)_i + b_2 \log \left(\frac{r}{w}\right)_i + b_3 \left[\log \left(\frac{r}{w}\right)_i\right]^2 + b_4 v_{1i}^2 + v_{2i} \,,$$

where $v_{1i} = (\rho + 1)^{-1}(u_{1i} - u_{2i})$, and $v_{2i}$ is a linear combination of all three disturbances. Since $v_{1i}^2$ is not observable, we can replace it by the squares of the residuals of (3.4), $\hat{v}_{1i}^2$. Asymptotically the two are identical. If $\hat{x}_{2i} - \hat{x}_{1i}$ represents the least squares fitted value of equation (3.4), which has the residual $\hat{v}_{1i}$, and $\hat{x}_{2i}$ represents the least squares fitted value of equation (3.5), which has the residual $\hat{v}_{2i}$, then we may write

$$(x_{2i} - x_{1i}) = (\hat{x}_{2i} - \hat{x}_{1i}) + \hat{v}_{1i} \,,$$
$$x_{2i} = \hat{x}_{2i} + \hat{v}_{2i} \,,$$
$$(x_{2i} - x_{1i})^2 = (\hat{x}_{2i} - \hat{x}_{1i})^2 + \hat{v}_{1i}^2 + 2\hat{v}_{1i}(\hat{x}_{2i} - \hat{x}_{1i}) \,.$$

This enables us to apply two-stage least squares methods to the production function (3.1) by replacing the values of the explanatory variables by their reduced form fitted values. The relationship to be estimated would then be

$$(3.6) \qquad x_{0i} = \log \gamma - \nu\delta(\hat{x}_{2i} - \hat{x}_{1i}) + \nu\hat{x}_{2i} - \frac{1}{2}\rho\nu\delta(1 - \delta)[\hat{x}_{2i} - \hat{x}_{1i})^2 + \hat{v}_{1i}^2] + v_{0i} \,.$$

An application of simple least squares method to (3.6) would give estimates of all four parameters as in the single equation case of Section 1. However, since a best linear unbiased estimate of $\rho$ can be obtained from (3.4), we are concerned only with estimating log $\gamma$, $\nu$ and $\delta$. One possibility is to assume *a priori* absence of parametric restraints on profit maximization, i.e., to let $R_1 = R_2 = 1$. Then we can obtain estimates of both $\rho$ and $\delta$ from (3.4), substitute them into (3.6), and obtain estimates of log $\gamma$ and $\nu$ by simple least squares.[14] Another possibility, not involving the assumption of $R_1 = R_2 = 1$, is to substitute $\hat{\rho}$ for $\rho$ into (3.6) and estimate the remaining parameters by restricted least squares method.

It should be noted that, unlike the case in Section 2, the method of estimation described above is not affected if the postulate of *actual* profit maximization is replaced by the postulate of *expected* profit maximization. The only effect of this would be a change in the constant term and in the definition of $v_{2i}$ of equation (3.5) .

## 4. CONCLUSION

In the preceding sections, we have discussed methods of estimating the parameters of the CES production function from various types of data. In

---

[13] The possibility of estimating $\rho$ in this way was also mentioned in the ACMS paper [1, (24)].

[14] A similar procedure was suggested in Dhrymes and Kurz [2].

contrast to the previous works on this subject, we have not assumed constant returns to scale *a priori*. This gain in generality involves a certain cost since we require data on capital input, which was not necessary in previously used estimates based on non-uniform prices (the case discussed in our Section 3) which assumed constant returns. However, if data on capital input are available, then the more general formulation of the CES function is clearly preferable. If such data are unobtainable, then the research worker who used the restrictive CES form has to face the possibility of a bias due to misspecification. Whether constant returns to scale are typical or not is, in the opinion of this author, largely an article of faith. With improved collection and measurement of capital, hopefully the dilemma may soon be resolved.

*Michigan State University, U.S.A.*

## APPENDIX

### A. 1. ERROR OF APPROXIMATING THE CES FUNCTION BY A "CORRECTED" COBB-DOUGLAS FORM

The error of approximating the CES function by (1.2) depends on the extent to which $\rho$ departs from zero, on the ratio of the two inputs, and on the values of the remaining parameters. More precisely, the (logarithmic) difference between the "approximate" and the "exact" value of the function is

$$\log X_{\text{approx.}} - \log X_{\text{exact}} = \nu(1 - \delta) \log(L/K) - \frac{1}{2} \rho\nu\delta(1 - \delta)[\log(L/K)]^2$$
$$+ (\nu/\rho) \log [\delta + (1 - \delta)(L/K)^{-\rho}] .$$

We have calculated the ratio of approximate value of the function $X_A$ to its exact value $X_B$ for various values of $\rho$ and $L/K$, taking $\nu = 0.9$ and $\delta = 4/9$. The results, in the form of ratios of approximate to exact values of $X$, are shown below.

| $\rho$ | L/K | | | | | |
|---|---|---|---|---|---|---|
| | 0.1 | 0.5 | 1.0 | 2.0 | 5.0 | 10.0 |
| −1.0 | 1.0636 | 0.9997 | 1.0000 | 1.0023 | 1.0403 | 1.1363 |
| −0.5 | 1.0033 | 0.9997 | 1.0000 | 1.0004 | 1.0076 | 1.0259 |
| −0.1 | 0.9995 | 0.9999 | 1.0000 | 0.9998 | 1.0000 | 1.0006 |
| 0 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |
| +0.1 | 0.9994 | 1.0002 | 1.0000 | 1.0000 | 1.0000 | 1.0005 |
| +0.2 | 0.9971 | 0.9999 | 1.0000 | 0.9999 | 1.0004 | 1.0013 |
| +0.5 | 0.9747 | 0.9996 | 1.0000 | 1.0002 | 1.0004 | 0.9967 |
| +1.0 | 0.8802 | 0.9972 | 1.0000 | 1.0003 | 0.9654 | 0.9402 |
| +10.0 | 0.0007 | 0.7335 | 1.0000 | 0.7710 | 0.0117 | 0.0081 |

## A. 2. RECONCILIATION OF MAXIMUM LIKELIHOOD AND INDIRECT LEAST SQUARES METHODS

In Section 2 we considered the problem of estimating production function parameters within the context of the Marschak-Andrews type of production model. Differentiating the likelihood function (2.7) and putting first derivatives equal to zero leads to the following equations:

(A.1)
$$\frac{\partial L}{\partial \nu} \equiv -n(1 - \nu)^{-1} - \sigma_{00}^{-1} \sum_{i=1}^{n} u_{0i}' \left[ -\delta x_{1i}' - (1 - \delta)x_{2i}' + \frac{1}{2} \rho\delta(1 - \delta)x_{3i}' \right]$$
$$- \sigma_{11}^{-1} \sum_{i=1}^{n} u_{1i}'[-\rho\nu^{-2}x_{0i}'] - \sigma_{22}^{-1} \sum_{i=1}^{n} u_{2i}' [-\rho\nu^{-2}x_{0i}'] = 0 \, ,$$

(A.2)
$$\frac{\partial L}{\partial \delta} \equiv -\sigma_{00}^{-1} \sum_{i=1}^{n} u_{0i}' \left[ -\nu x_{1i}' + \nu x_{2i}' + \frac{1}{2} \rho\nu(1 - 2\delta)x_{3i}' \right] = 0 \, ,$$

(A.3)
$$\frac{\partial L}{\partial \rho} \equiv n(\rho + 1)^{-1} - \sigma_{00}^{-1} \sum_{i=1}^{n} u_{0i}' \left[ \frac{1}{2} \nu\delta(1 - \delta)x_{3i}' \right]$$
$$- \sigma_{11}^{-1} \sum_{i=1}^{n} u_{1i}' [\nu^{-1}x_{0i}' - x_{1i}'] - \sigma_{22}^{-1} \sum_{i=1}^{n} u_{2i}' [\nu^{-1}x_{0i}' - x_{2i}'] = 0 \, ,$$

(A.4)
$$\frac{\partial L}{\partial \sigma_{00}} \equiv -\frac{n}{2} \sigma_{00}^{-1} + \frac{1}{2} \sigma_{00}^{-2} \sum_{i=1}^{n} u_{0i}'^2 = 0 \, ,$$

(A.5)
$$\frac{\partial L}{\partial \sigma_{11}} \equiv -\frac{n}{2} \sigma_{11}^{-1} + \frac{1}{2} \sigma_{11}^{-2} \sum_{i=1}^{n} u_{1i}'^2 = 0 \, ,$$

(A.6)
$$\frac{\partial L}{\partial \sigma_{22}} \equiv -\frac{n}{2} \sigma_{22}^{-1} + \frac{1}{2} \sigma_{22}^{-2} \sum_{i=1}^{n} u_{2i}'^2 = 0 \, ,$$

where primes indicate deviations from the respective sample means, and $x_3 = (x_2 - x_1)^2$ .

Equations (A.4) to (A.6) give the solution for $\sigma_{00}, \sigma_{11}$ and $\sigma_{22}$ in terms of the remaining parameters. Substituting these into (A.1) to (A.3) gives

(A.7)
$$\left( \frac{\rho}{\nu} + 1 \right)\left[ m_{00} - \nu\delta m_{01} - \nu(1 - \delta)m_{02} + \frac{1}{2} \rho\nu\delta(1 - \delta)m_{03} \right]$$
$$- (\rho + 1)\left[ m_{01} - \nu\delta m_{11} - \nu(1 - \delta)m_{12} + \frac{1}{2} \rho\nu\delta(1 - \delta)m_{13} \right] = 0 \, ,$$

(A.8)
$$\left( \frac{\rho}{\nu} + 1 \right)\left[ m_{00} - \nu\delta m_{01} - \nu(1 - \delta)m_{02} + \frac{1}{2} \rho\nu\delta(1 - \delta)m_{03} \right]$$
$$- (\rho + 1)\left[ m_{02} - \nu\delta m_{12} - \nu(1 - \delta)m_{22} + \frac{1}{2} \rho\nu\delta(1 - \delta)m_{23} \right] = 0 \, ,$$

(A.9)
$$\left( \frac{\rho}{\nu} + 1 \right)^2 m_{00} - \left( \frac{\rho}{\nu} + 1 \right)(\rho + 1)(m_{01} + m_{02}) + (\rho + 1)^2 m_{12} = 0 \, ,$$

where

$$m_{rs} = \frac{1}{n} \sum_{i=1}^{n} x_{ri}' x_{si}' \qquad (r, s = 0, 1, 2, 3) \, .$$

Now, ILS estimates are derived from the least squares estimates of

(A.10) $$z_{0i} = a_0 + a_1 z_{1i} + a_2 z_{2i} - \frac{1}{2} a_1 a_2 (\hat{F} - 1) z_{3i} + e_i ,$$

where

$$z_{0i} = x_{0i}, \quad z_{1i} = z_{0i} - \hat{F} x_{1i}, \quad z_{2i} = x_{0i} - \hat{F} x_{2i} , \quad z_{3i} = (x_{1i} - x_{2i})^2.$$

$\hat{F}$ is obtained from

(A.11) $$\hat{F}^2 m_{00} - \hat{F}(m_{01} + m_{02}) + m_{12} = 0 ,$$

i.e., $\hat{F} = (\hat{\rho}/\hat{v} + 1)(\hat{\rho} + 1)^{-1}$ .[15]

The "normal" equations for the least squares estimates of (A.10) are

(A.12)
$$M_{01} - \hat{a}_1 M_{11} - \hat{a}_2 M_{12} + \frac{1}{2} \hat{a}_1 \hat{a}_2 (\hat{F} - 1) M_{13}$$
$$= \hat{a}_2 [M_{03} - \hat{a}_1 M_{13} - \hat{a}_2 M_{23} + \frac{1}{2} \hat{a}_1 \hat{a}_2 (\hat{F} - 1) M_{33}] ,$$

(A.13)
$$M_{02} - \hat{a}_1 M_{12} - \hat{a}_2 M_{22} + \frac{1}{2} \hat{a}_1 \hat{a}_2 (\hat{F} - 1) M_{23}$$
$$= \hat{a}_1 [M_{03} - \hat{a}_1 M_{13} - \hat{a}_2 M_{23} + \frac{1}{2} \hat{a}_1 \hat{a}_2 (\hat{F} - 1) M_{33}] ,$$

where $M_{rs} = \frac{1}{n} \sum_{i=1}^{n} z'_{ri} z'_{si}$   $(r, s = 0, 1, 2, 3)$ .

Now if (a) the $M$'s are expresed in terms of the $m$'s and $\hat{F}$, (b) $\hat{F}$ is replaced by $(\hat{\rho}/\hat{v} + 1)(\hat{\rho} + 1)^{-1}$, (c) the $\hat{a}$'s are expressed in terms of $\hat{v}$, $\hat{\delta}$, and $\hat{\rho}$ as per relationships (2.10) through (2.13) of the text, it can easily be shown that equations (A.11), (A.12), and (A.13) are equivalent to (A.9), (A.7), and (A.8), respectively. Thus we can conclude that the ML and the ILS estimates are identical.

Let us, finally, prove that the roots of equation (A.11) are asymptotically real, i.e., that $(m_{01} + m_{02})^2 - 4 m_{00} m_{12} \geqq 0$ for $n \to \infty$. From the specification of the profit maximizing conditions (2.5) and (2.6) in the text) and the definition of $F$ we have

(A.14) $$x_{1i} = k_1/(\rho + 1) + F x_{0i} - u_{1i}/(\rho + 1) ,$$

(A.15) $$x_{2i} = k_2/(\rho + 1) + F x_{0i} - u_{2i}/(\rho + 1) .$$

That is,

$$m_{01} = F m_{00} - (\rho + 1)^{-1} \text{ cov } (x_0, u_1) ,$$

$$m_{02} = F m_{00} - (\rho + 1)^{-1} \text{ cov } (x_0, u_2) ,$$

$$m_{12} = F^2 m_{00} - F(\rho + 1)^{-1} \text{ cov } (x_0, u_1)$$
$$- F(\rho + 1)^{-1} \text{ cov } (x_0, u_2) + (\rho + 1)^{-2} \text{ cov } (u_1, u_2) .$$

---

[15] The question concerning which of the two roots is to be chosen is discussed in the text of Section 2.

Then,

$$(A.16) \quad (m_{01} + m_{02})^2 - 4m_{00}m_{12}$$
$$= (\rho + 1)^{-2}[\text{cov}\,(x_0, u_1) + \text{cov}\,(x_0, u_2)]^2 - 4(\rho + 1)^{-2}[\text{cov}\,(u_1, u_2)]m_{00}\,.$$

Since, by assumption, $\text{cov}\,(u_1, u_2) \to 0$ as $n \to \infty$ and the remaining term on the RHS is necessarily positive (or zero), the proposition is proved.

## REFERENCES

[1] ARROW, K. J., H. B. CHENERY, B. S. MINHAS, and R. M. Solow, "Capital-Labor Substitution and Economic Efficiency," *Review of Economics and Statistics*, XLIII (August, 1961), 225-50.

[2] DHRYMES, P. J., and M. KURZ, "Technology and Scale in Electricity Generation," *Econometrica*, XXXII (July, 1964), 287-315.

[3] EISENPRESS, H., and J. L. GREENSTADT, "Non-linear Full-Information Estimation," paper presented at the annual meeting of the Econometric Society in Chicago, Illinois, December 27-30, 1964 (mimeographed).

[4] GRILICHES, Z., "Production Functions, Technical Change, and All That," Netherlands School of Economics, Econometric Institute Report 6328, October 2, 1963 (mimeographed).

[5] HOCH, I., "Simultaneous Equation Bias in the Context to the Cobb-Douglas Production Function," *Econometrica*, XXVI (October, 1958), 566-78.

[6] ———, "Estimation of Production Function Parameters Combining Time-Series and Cross-Section Data," *Econometrica*, XXX (January, 1962), 34-53.

[7] KLEIN, L. R., *A Textbook of Econometrics*, (New York: Row, Peterson and Company, Inc., 1953).

[8] MARSCHAK, J., and W. J. ANDREWS, "Random Simultaneous Equations and the Theory of Production," *Econometrica*, XII (July-October, 1944), 143-205.

[9] MUNDLAK, Y., and I. HOCH, "Consequences of Alternative Specifications in Estimation of Cobb-Douglas Production Functions," *Econometrica*, XXXIII, (October, 1965), 814-28.

[10] SATO, R., "The Estimation of Biased Technical Progress and the Production Function," paper presented at the annual meeting of the Econometric Society in Chicago, Illinois, December 27-30, 1964 (mimeographed).

[11] ZELLNER, A., J. KMENTA, and J. DRÉZE., "Specification and Estimation of Cobb-Douglas Production Function Models," Social Systems Research Institute at the University of Wisconsin, Systems Formulation and Methodology Paper 6409 (forthcoming in *Econometrica*).