THE ESTIMATION OF COBB-DOUGLAS TYPE FUNCTIONS WITH MULTIPLICATIVE AND ADDITIVE ERRORS: A FURTHER ANALYSIS

H. H. Kelejian

Econometric Research Program
Research Paper No. 27
March 1971

The research described in this paper was supported in part by NSF Grant GS 2799.

Econometric Research Program
PRINCETON UNIVERSITY
207 Dickinson Hall
Princeton, New Jersey

The Estimation of Cobb-Douglas Type Functions With Multiplicative and Additive Errors: A Further Analysis*

H. H. Kelejian

1. Introduction

In a recent article in this journal Goldfeld and Quandt (4) suggested a maximum likelihood approach to the estimation of a Cobb-Douglas type model when the model includes both multiplicative and additive disturbance terms. As expected, an analytical expression for the solution to the maximization problem did not exist. Indeed, because of the complexity of the likelihood function, their maximization algorithm had to be used in conjunction with a numerical integration technique.

The purpose of this paper is to generalize and simplify the work by Goldfeld and Quandt. Specifically, an estimation technique is suggested which does not require the specification of the disturbance terms beyond their means and variances, which does not require the compounding of a maximization algorithm with a numerical integral technique, but yet leads to asymptotically efficient estimates of the parameters of the regression function. In addition, the procedure readily lends itself to interpretation. For instance, it will become evident that if the distribution of the multiplicative disturbance term is not known, the scale parameter of the model (unlike the other parameters) will not be identified.

---

*This research was supported in part, by the Ford Foundation. I would like to thank R. R. Cornwall for introducing me to this problem. I would also like to thank R. Fair and R. Quandt for helpful comments. These people, of course, are not responsible for any shortcomings of this paper.

2. Model Specification and Estimation

Consider the model suggested by Goldfeld and Quandt,

$$(1)\quad y_t = \alpha_0 x_{1t}^{\alpha_1} \cdots x_{kt}^{\alpha_k} e^{u_t} + v_t,\ t = 1, \ldots n ,$$

where $x_{1t} \ldots x_{kt}$ are the t^{th} observations on the k independent variables,[1] and u_t and v_t are the t^{th} values of the disturbance terms. We assume that u_t and v_t are independent of each other, are independent of $x_{1t} \ldots x_{kt}$, and $Eu_t = Ev_t = 0$, and $Eu_t^2 = \sigma_u^2$, and $Ev_t^2 = \sigma_v^2$. We also assume that u_t and v_t are independent of u_s and v_s for all $s \neq t$. We do not, however, assume normality on any other particular distribution.

Let A be the mean of e^{u_t}. Then e^{u_t} may be expressed as

$$(2)\quad e^{u_t} = A + \phi_t ,$$

where $E\phi_t = 0$. Substituting (2) into (1) we obtain

$$(3)\quad y_t = Bx_{1t}^{\alpha_1} \cdots x_{kt}^{\alpha_k} + w_t ,$$

where $B = \alpha_0 A$, and $w_t = \alpha_0 x_{1t}^{\alpha_1} \cdots x_{kt}^{\alpha_k}\phi_t + v_t$. Since ϕ_t and v_t are independent of $x_t = (x_{1t} \ldots x_{kt})$, it follows that $E[w_t|x_t] = 0$. Therefore if we ignore the heteroscedasticity of w_t and apply nonlinear least squares to (3), the resulting estimates of B and $\alpha_1, \ldots \alpha_k$, although not efficient, will be consistent.[2] The remainder of the procedure may now be evident. These estimates will be used to obtain a consistent estimate of the variance of w_t, then (3) will be transformed to rid w_t of its heteroscedasticity, and finally, nonlinear least

[1] Goldfeld and Quandt assume that the independent variables are nonstochastic. This assumption, however, is unnecessary for the results of this paper.

[2] The proof of this would be almost identical to the one described in Aigner and Goldberger (1, p. 715). Actually, in a different context, such a proof is outlined below.

squares will be applied to the transformed model.

The conditional variance of w_t is easily shown to be

(4) $E[w_t^2|x_t] = (\alpha_0^2\sigma_\phi^2)\, x_{1t}^{2\alpha_1} \ldots x_{kt}^{2\alpha_k} + \sigma_v^2 = cz_t + \sigma_v^2$

where σ_ϕ^2 is the variance of ϕ_t, $c = \alpha_0^2\sigma_\phi^2$, and $z_t = x_{1t}^{2\alpha_1} \ldots x_{kt}^{2\alpha_k}$. Now, the consistent estimates of B and $\alpha_1, \ldots \alpha_k$ described above enable us to obtain a consistent estimate of w_t from (3), namely

(5) $\hat{w}_t = y_t - \hat{B}x_{1t}^{\hat{\alpha}_1} \ldots x_{kt}^{\hat{\alpha}_k}$.

The consistency of $\hat{w}_t$ implies that

(6) $\hat{w}_t^2 = w_t^2 + \Delta_t$

where the probability limit of Δ_t with respect to n is zero. Given (6) and (4) it is clear that a consistent estimate of the conditional variance of w_t can be easily obtained by ordinary least squares via the linear regression[3]

(7) $\hat{w}_t^2 = \sigma_v^2 + c\hat{z}_t + \varepsilon_t$

where $\hat{z}_t = x_{1t}^{2\hat{\alpha}_1} \ldots x_{kt}^{2\hat{\alpha}_k}$, and ε_t is a disturbance term. Specifically, if the estimates of σ_v^2 and c are $\hat{\sigma}_v^2$ and $\hat{c}$, then our consistent estimate of the conditional variance of w_t is

(8) $\hat{\sigma}_{wt}^2 = \hat{\sigma}_v^2 + \hat{c}\hat{z}_t$.

Finally, nonlinear least squares can now be applied to our basic model (3) after it has been transformed by dividing it though by $\hat{\sigma}w_t$. The resulting estimates of B and $\alpha_1, \ldots \alpha_k$ are obviously consistent, and, since $\hat{\sigma}w_t$ is a consistent estimate of σw_t, these estimates should

[3]To see this note that (4) implies that $w_t^2 = cz_t + \sigma_v^2 + r_t$, where $E[r_t|x_t] = 0$. Thus $\hat{w}_t^2 = cz_t + \sigma_v^2 + (r_t + \Delta_t)$. Since $\hat{z}_t$ converges to z_t the results from (7) follow. The reader should note that limits of expectations are not being taken. That convergence in probability need not imply convergence in moments is nicely demonstrated in Dhrymes (2, pp. 88-89).

also be asymptotically efficient.[4]

To see this let y, f, and w be the n x 1 vectors whose t^{th} elements are y_t, $Bx_{1t}^{\alpha_1} \ldots x_{kt}^{\alpha_k}$, and w_t. Then the proposed estimates can be obtained by minimizing S where

$$(9) \quad S = (y - f)' \hat{\Sigma}^{-1}(y - f)$$

where $\hat{\Sigma}^{-1}$ is the n x n diagonal matrix whose t^{th} diagonal element is $\hat{\sigma}^2_{wt}$. Let P denote the (k+1) x 1 vector of parameters: $P' = (B\ \alpha_1 \ldots \alpha_k)$. Then minimizing S with respect to P yields

$$(10) \quad \hat{f}_p' \hat{\Sigma}^{-1} (\hat{f} - y) = 0$$

where $\hat{f}_p$ is the n x k + 1 matrix $\frac{\partial f}{\partial P}$ evaluated at our estimate $\hat{P}$, and $\hat{f}$ is the n x 1 vector, f, evaluated at $\hat{P}$. Linearizing (10) about the true parameters we have

$$(11) \quad f_p' \Sigma^{-1} (f + f_p(\hat{P} - P) - y) = 0$$

or

$$(12) \quad \hat{P} - P = (f_p' \Sigma^{-1} f_p)^{-1} f_p' \Sigma^{-1} w ,$$

where Σ is the n x n diagonal matrix whose t^{th} diagonal element is σ^2_{wt}. It follows from (12) that $\hat{P}$ is consistent. Further, under general conditions, it can be shown[5] that $\sqrt{T}\ (\hat{P} - P)$ convergences in distribution to a multivariate normal with mean vector zero and

---

[4]The estimate of the scale parameter α_0 can be obtained from the final estimate of $B = \alpha_0 A$, and our estimate of $c = \alpha_0^2 \sigma_\phi^2$. It is interesting to note that to do this (solve for α_0) the density of u_t must be such that σ_ϕ^2 and A depend upon a single common parameter. Furthermore, to know how σ_ϕ^2 and A relate to each other in terms of that parameter the density of u_t must be specified.

[5]See Dhrymes (2, p. 108). For a concise discussion of some of the problems involved in discussing sequences of random variables see Chapter 3 of Dhrymes (2).

conditional covariance matrix

$$(13) \quad V_{\hat{p}} = T(f_p' \, \Sigma^{-1} f_p)^{-1} .$$

A consistent estimate of $V_{\hat{p}}$ is obviously

$$(14) \quad \hat{V}_{\hat{p}} = T(\hat{f}_p' \, \hat{\Sigma}^{-1} \hat{f}_p)^{-1} .$$

Since $V_{\hat{p}}$ is the asymptotic covariance matrix of the Aitken estimator, it follows that $\hat{P}$ is asymptotically efficient.[6]

---

[6] One could, of course, iterate on this technique. That is, $\hat{P}$ could be used to obtain another estimate of Σ which would lead to another estimate of P, etc. The asymptotic properties, however, of the resulting estimate of P would not be different from those of $\hat{P}$ described above.

BIBLIOGRAPHY

1. Aigner, D. J. and A. S. Goldberger, "Estimation of Pareto's Law from Grouped Observations," Journal of the American Statistical Association, Vol. 65 (September 1970), 712-723.

2. Dhrymes, P. J. (1970), Econometrica, (New York: Harper and Row, 1970).

3. Goldberger, A. S. (1964), Econometric Theory, (New York: Wiley and Sons, 1964).

4. Goldfeld, S. M. and R. E. Quandt (1970), "The Estimation of Cobb-Douglas Type Functions with Multiplicative and Additive Errors," International Economic Review, Vol. 11 (June 1970).

5. Kelejian, H. H. (1971), "Two Stage Least Squares and Nonlinear Systems," Forthcoming, Journal of the American Statistical Association, (June 1971.)