

Journal of Econometrics 5 (1977) 155–165. © North-Holland Publishing Company

ON BAYESIAN AND NON-BAYESIAN ESTIMATION OF A TWO-LEVEL CES PRODUCTION FUNCTION FOR THE DUTCH MANUFACTURING SECTOR*

Rins HARKEMA and Sybrand SCHIM VAN DER LOEFF

Econometric Institute, Erasmus University, Rotterdam, The Netherlands

Received May 1975, revised version received February 1976

In this paper maximum-likelihood estimates of the parameters of the two-level CES function, obtained by direct estimation of this function, are given. In addition the authors propose to show how a Bayesian analysis may help to find a solution to the difficulties related with, but not specific to, this particular estimation problem. It is shown that numerical integration of the posterior distribution may give an indication as to which parameter has to be pinpointed and at which value when multi-collinearity precludes unconditional maximization of the likelihood. It is suspected that this approach has a wider field of application.

1. Introduction

The two-level constant-elasticity-of-substitution function has been introduced by Sato (1967) almost a decade ago. Since then most of its applications are to be found in the literature on consumer behavior. Sato gives estimation results for the two-level CES production function, that are based on the first-order conditions for a cost minimum under the assumption of perfectly competitive markets of factor supplies. The absence of direct estimation results¹ [direct in the sense of Hodges (1969)], may well be explained by the usual difficulties which are encountered when estimating an intrinsically non-linear relationship, such as the one at hand. Moreover this particular function is typical in that it reduces to the ordinary CES production function for a certain combination of the parameters, which is liable to present additional difficulties, in particular when the available time series are afflicted with a considerable degree of multi-collinearity and not devoid of measurement errors. Apart from its own interest it is therefore interesting to study the estimation of the two-level CES production function as an example of some of the difficulties that show up in many a present-day application of econometric methods.

*Paper presented at the Third World Congress of the Econometric Society at Toronto, August 20–26, 1975.

¹A notable exception is a recent paper by Mizon (1974).

The present paper provides maximum-likelihood estimates of the parameters of the two-level CES production function, obtained by direct estimation of this function. In addition it is shown how a Bayesian analysis of the problem may increase the chances to find a solution to the difficulties mentioned in the preceding paragraph. When faced with the problem of multicollinearity, a common procedure is to pinpoint one or more parameters at a predetermined value. It is demonstrated how numerical integration of the posterior distribution may give an indication as to which parameter has to be pinpointed and at which value. It should be remarked at the outset, that the purpose of this paper is not to give a full-fledged Bayesian analysis, but rather to demonstrate how a Bayesian analysis may be helpful in solving estimation problems, that otherwise are to be solved by trial and error methods, which in many cases will prove to be too costly and too time-consuming.

The plan of the paper is as follows. Section 2 outlines the model and its underlying assumptions. Section 3 deals with the nature of the data used in the estimation procedure and with the effects on the parameters of the transformation to index numbers. Section 4, finally, discusses the Bayesian and maximum-likelihood approaches to the estimation of the model and presents the results.

2. The model

In this paper we suppose that production takes place according to the following two-level CES production function:

$$V_t = \gamma e^{gt} \{ \theta_1 K_{1t}^{-\rho_1} + [\theta_2 K_{2t}^{-\rho_2} + \theta_3 L_t^{-\rho_2}]^{\rho_1/\rho_2} \}^{-1/\rho_1} e^{\varepsilon_t}, \quad t = 0, 1, \dots, T, \quad (1)$$

where V_t denotes value added in constant prices, and K_{1t} , K_{2t} and L_t are measures for the inputs of services of structures, equipment and labor, respectively. γ is usually called the efficiency parameter, g is the rate of Hicks-neutral growth, and ρ_1 and ρ_2 are so-called substitution parameters. Defining

$$\sigma_i = (1 + \rho_i)^{-1}, \quad i = 1, 2, \quad (2)$$

it is well-known that σ_1 may be used to measure the substitution possibilities between structures and the mixture of labor and equipment, while σ_2 measures the possibilities to substitute equipment for labor.

The θ_i ($i = 1, 2, 3$) are parameters that, in fact, adjust the dimensions. The disturbance term ε_t indicates that production is subject to random influences like weather conditions, unexpected breakdowns of machines, etc. This random nature leads us to suppose that the entrepreneur will aim at maximizing the mathematical expectation of profits, i.e.,

$$E[\Pi_t] = E[p_t V_t] - E[w_{1t}] K_{1t} - E[w_{2t}] K_{2t} - E[w_{3t}] L_t, \quad (3)$$

where p_t and w_{it} ($i = 1, 2, 3$) denote the prices of output and factor services, respectively. Demand for output and supply of factor services are supposed to be given by the following relations:

$$\begin{aligned} V_t &= \zeta_{0t} p_t^{\eta_0} e^{u_{0t}}, \\ K_{it} &= \zeta_{it} w_{it}^{\eta_i} e^{u_{it}}, \quad i = 1, 2, \\ L_t &= \zeta_{3t} w_{3t}^{\eta_3} e^{u_{3t}}, \end{aligned} \tag{4}$$

where the η_i are the respective price elasticities and the ζ_{it} comprise the non-price effects. The disturbance terms u_{it} ($i = 0, 1, 2, 3$) represent the random elements in demand and supply; it will be assumed that the disturbances u_{it} ($i = 0, 1, 2, 3$) as well as the disturbance ε_t are normally distributed. In order to determine the expected profit maximum we first eliminate the price variables p_t and w_{it} ($i = 1, 2, 3$) from (3) by substituting the relationships (4). Differentiation of the resulting expression with respect to K_{1t} , K_{2t} and L_t , and setting the first-order conditions equal to zero yields the optimal levels K_{1t}^* , K_{2t}^* and L_t^* . The observed quantities of K_{1t} , K_{2t} and L_t are assumed to be given in terms of these optimal levels as follows:

$$\begin{aligned} K_{it} &= K_{it}^* e^{v_{it}}, \quad i = 1, 2, \\ \text{and} \\ L_t &= L_t^* e^{v_{3t}}, \end{aligned} \tag{5}$$

where the v_{it} ($i = 1, 2, 3$) denote the failure to adjust the input levels to the optimal levels. The model is completed with four definitional equations linking quantities of output and factor services demanded with quantities supplied.

As a straightforward generalization of the result obtained by Hodges (1969) it can be shown that if the disturbance terms v_{it} ($i = 1, 2, 3$) are statistically independent of ε_t in (1), K_{it} ($i = 1, 2$) and L_t are independent of ε_t as well. Consequently consistent estimates of the parameters of the two-level CES production function can be obtained by direct estimation of (1), provided that the assumptions of this section hold good.

3. Data and index number transforms

Apart from the time series for structures the data that have been used in the estimation procedure are the same as in Schim van der Loeff and Harkema (1976), to which the interested reader is referred for a detailed discussion, as well as the data themselves. For structures a somewhat crude series has been constructed on the basis of figures on gross investment in structures in the Dutch manufacturing sector, I_t , and the relationship

$$K_{1t} = (1 - \delta) \cdot K_{1,t-1} + I_t, \tag{6}$$

where δ is the rate of depreciation. In accordance with the findings of Jorgenson and Griliches (1967), the value of δ has been set equal to 0.0513. Given an initial stock of structures or, equivalently, an initial growth rate of the stock of structures an index of inputs of services of structures can be constructed, under the, admittedly crude, assumption that the services are proportional to stock. Since this assumption was thought untenable for equipment and labor, an index of inputs of services of equipment was constructed on the basis of a series of total energy consumption of the Dutch manufacturing sector; in addition the series of labor volume was corrected for effects of sex, age, and schooling. For output a series of gross value added against factor prices has been used.

Since some of the data are only available in index numbers, the production function has been rewritten in terms of index numbers. Dividing (1) through by the production function in the base period $t = 0$ yields

$$(V_t/V_0) = e^{gt} \{ \delta_1 (K_{1t}/K_{10})^{-\rho_1} + (1 - \delta_1) [\delta_2 (K_{2t}/K_{20})^{-\rho_2} + (1 - \delta_2) (L_t/L_0)^{-\rho_2}]^{\rho_1/\rho_2} \}^{-1/\rho_1} e^{v_t}, \tag{7}$$

where $v_t = \varepsilon_t - \varepsilon_0$, and δ_1 and δ_2 are defined as

$$\begin{aligned} \delta_1 &= \theta_1 K_{10}^{-\rho_1} \{ \theta_1 K_{10}^{-\rho_1} + [\theta_2 K_{20}^{-\rho_2} + \theta_3 L_0^{-\rho_2}]^{\rho_1/\rho_2} \}^{-1}, \\ \delta_2 &= \theta_2 K_{20}^{-\rho_2} \{ \theta_2 K_{20}^{-\rho_2} + \theta_3 L_0^{-\rho_2} \}^{-1}. \end{aligned} \tag{8}$$

Now δ_1 and δ_2 have a rather clear-cut interpretation under assumptions that are slightly different from those in the preceding section. Suppose, for the time being, that the supply schedules for each of the production factors are completely known, i.e., the disturbance terms u_{it} ($i = 1, 2, 3$) in (4) are identically zero. On dividing the first-order condition for a profit maximum with respect to K_{1t} , through by the sum of the first-order conditions with respect to K_{2t} and L_t , one obtains

$$\begin{aligned} \xi_1 w_{1t} (\xi_2 w_{2t} + \xi_3 w_{3t})^{-1} &= \theta_1 K_{1t}^{-\rho_1 - 1} \{ [\theta_2 K_{2t}^{-\rho_2} + \theta_3 L_t^{-\rho_2}]^{\rho_1/\rho_2 - 1} \\ &\quad \times [\theta_2 K_{2t}^{-\rho_2 - 1} + \theta_3 L_t^{-\rho_2 - 1}] \}^{-1}, \end{aligned} \tag{9}$$

where $\xi_i = 1 + \eta_i^{-1}$ ($i = 1, 2, 3$). Likewise, dividing the profit-maximizing condition with respect to K_{2t} by the profit-maximizing condition with respect to L_t , yields

$$\xi_2 w_{2t} (\xi_3 w_{3t})^{-1} = \theta_2 K_{2t}^{-\rho_2 - 1} (\theta_3 L_t^{-\rho_2 - 1})^{-1}. \tag{10}$$

As (9) and (10) also apply in the base period, some simple algebraic operations are sufficient to show that (8), (9), and (10) give rise to the following equalities:

$$\begin{aligned} \delta_1 &= \xi_1 w_{10} K_{10} (\xi_1 w_{10} K_{10} + \xi_2 w_{20} K_{20} + \xi_3 w_{30} L_0)^{-1}, \\ \delta_2 &= \xi_2 w_{20} K_{20} (\xi_2 w_{20} K_{20} + \xi_3 w_{30} L_0)^{-1}. \end{aligned} \tag{11}$$

So δ_1 and δ_2 are the base-period shares of the cost of structures in total costs and of the cost of equipment in the costs of the inner 'nest', respectively, where it has to be understood that each cost component is weighted by one and the same function of the associated price elasticity of supply, viz., $1 + \eta_i^{-1}$ ($i = 1, 2, 3$). This knowledge would make it abundant to estimate the parameters δ_1 and δ_2 , if the elasticities were known exactly. Unfortunately knowledge about these elasticities is very scarce or inexact.

4. Estimation²

Under the assumption that ε_t ($t = 0, 1, \dots, T$) in (1) is normally distributed with zero mean and covariance matrix $\sigma^2 I$, it holds good that v_t ($t = 1, \dots, T$) in (7) is also normally distributed with zero mean and covariance matrix $\sigma^2 A = \sigma^2(I + \iota \iota')$, where ι denotes the $T \times 1$ vector with all elements equal to unity. Therefore, maximum-likelihood estimates of the parameters $g, \delta_1, \delta_2, \rho_1$ and ρ_2 can be obtained by minimizing the quadratic form $v' A^{-1} v$, where v is the $T \times 1$ vector composed of elements v_t defined as [see eq. (7)]

$$v_t = \ln(V_t/V_0) - g t + \rho_1^{-1} \ln\{\delta_1(K_{1t}/K_{10})^{-\rho_1} + (1 - \delta_1)[\delta_2(K_{2t}/K_{20})^{-\rho_2} + (1 - \delta_2)(L_t/L_0)^{-\rho_2}]^{\rho_1/\rho_2}\}. \quad (12)$$

However, as indicated before, the maximum-likelihood estimates turn out to be unacceptable for most parameters. Some experimentation with pinpointed values of δ_1 and δ_2 in the neighbourhood of the shares, computed according to (11) under the assumption of perfectly competitive markets of factor supplies, did not lead to satisfactory results. Because any further information about the price elasticities was lacking, another approach was attempted.

During the last decade much research has been done and much information has become available about elasticities of substitution between capital and labor. Therefore an obvious way out seems to consist of incorporating this information into the estimation procedure by passing to a Bayesian approach. A similar analysis has been carried out by Sankar (1970) for the case of the ordinary CES production function. Following his line of approach, the joint prior distribution for δ_1, δ_2, g , and h (the precision of the disturbance ε_t), given the parameters σ_1 and σ_2 in (2) is specified to be of the following (diffuse) form:

$$p(\delta_1, \delta_2, g, h | \sigma_1, \sigma_2) \propto h^{-1}, \\ 0 < \delta_1 < \delta_2 < 1, \quad -\infty < g < \infty, \quad 0 < h < \infty. \quad (13)$$

As regards the parameters σ_1 and σ_2 , it can be shown³ that σ_2 represents the

²The authors are indebted to Mr. A.S. Louter of the Econometric Institute for his assistance in preparing the required computer programs.

³See Sato (1967, p. 203).

(direct) elasticity of substitution between labor and equipment. From numerous empirical investigations of the ordinary CES production function, substantive information has become available about this parameter. Parameter σ_1 denotes the elasticity of substitution between inputs of services of structures and the 'mixture' of labor and equipment services. It may seem troublesome to interpret an elasticity of substitution between structures and this mixture (a generalized harmonic mean) of inputs of labor and equipment services. Sato, however, also demonstrates that the Allen partial elasticity of substitution between structures on the one hand, and equipment and labor on the other, is indeed given by σ_1 . Clearly, the value of σ_1 indicates *grasso modo* the possibilities of substituting inputs of services of structures for inputs of equipment and for labor services, whichever definition of elasticity of substitution is employed. The prior distributions for σ_1 and σ_2 will be specified to be statistically independent and of the following form:

$$p(\sigma_1) \propto e^{-\sigma_1/0.29} \quad \text{and} \quad p(\sigma_2) \propto \sigma_2^8 e^{-\sigma_2/0.11},$$

$$0 < \sigma_1, \sigma_2 < \infty. \quad (14)$$

These are gamma distributions with one and nine degrees of freedom, respectively. The prior distribution for σ_1 supposedly reflects that the substitution possibilities between structures and the other factors are very scarce. It has its mode at zero and the scale parameter has been chosen such that its median equals 0.20. The parameters of the prior distribution for σ_2 have been chosen so as to make this distribution almost symmetric with median and mean close to one in order to reflect the substantive empirical evidence about the elasticity of substitution between labor and equipment.

If we rewrite (12) as $v_t = y_t - gt$, where y_t in terms of σ_1 and σ_2 is given by

$$y_t = \ln(V_t/V_0) - \sigma_1(\sigma_1 - 1)^{-1} \ln\{\delta_1(K_{1t}/K_{10})^{(\sigma_1 - 1)/\sigma_1} \\ + (1 - \delta_1)[\delta_2(K_{2t}/K_{20})^{(\sigma_2 - 1)/\sigma_2} \\ + (1 - \delta_2)(L_t/L_0)^{(\sigma_2 - 1)/\sigma_2}]^{\sigma_1 - 1} \sigma_2 / (\sigma_2 - 1) \sigma_1\}, \quad (15)$$

the joint posterior distribution of δ_1 , δ_2 , σ_1 , σ_2 , g and h can be written as

$$p(\delta_1, \delta_2, \sigma_1, \sigma_2, g, h | \text{data}) \propto h^{T/2 - 1} \exp\{-\frac{1}{2}h(y - gt)'A^{-1}(y - gt)\} \\ \times e^{-\sigma_1/0.29} \cdot \sigma_2^8 e^{-\sigma_2/0.11}, \quad (16)$$

where y and t are the $T \times 1$ vectors composed of the elements y_t and t_t , respectively. Integration with respect to h yields the joint posterior distribution of δ_1 , δ_2 , σ_1 , σ_2 and g , viz.,

$$p(\delta_1, \delta_2, \sigma_1, \sigma_2, g | \text{data}) \propto [(y - gt)'A^{-1}(y - gt)]^{-T/2} \\ \times e^{-\sigma_1/0.29} \cdot \sigma_2^8 e^{-\sigma_2/0.11}. \quad (17)$$

Rewriting the term within square brackets in (17) as

$$(y-gt)'A^{-1}(y-gt) = (t'A^{-1}t)(g-\hat{g})^2 + (y-\hat{g}t)'A^{-1}(y-\hat{g}t), \quad (18)$$

where $\hat{g} = (t'A^{-1}t)^{-1}t'A^{-1}y$, the Student- t distributed variable g can be integrated out, to yield the joint posterior distribution of $\delta_1, \delta_2, \sigma_1$ and σ_2 , viz.,

$$p(\delta_1, \delta_2, \sigma_1, \sigma_2 | \text{data}) \propto [(y-\hat{g}t)'A^{-1}(y-\hat{g}t)]^{-(T-1)/2} \times e^{-\sigma_1/0.29} \cdot \sigma_2^8 e^{-\sigma_2/0.11}. \quad (19)$$

The marginal posterior density functions for $\delta_1, \delta_2, \sigma_1$ and σ_2 can be obtained by resorting to numerical integration techniques.

In appendix A we have plotted the marginal posterior distributions of $\delta_1, \delta_2, \sigma_1$ and σ_2 resulting from the joint posterior distribution (19) (unbroken line). In order to ascertain to what extent the particular form of the prior distribution (14) influences the shape of these marginal density functions, the marginal posterior distributions of $\delta_1, \delta_2, \sigma_1$ and σ_2 , resulting from numerically integrating

$$p'(\delta_1, \delta_2, \sigma_1, \sigma_2 | \text{data}) \propto [(y-\hat{g}t)'A^{-1}(y-\hat{g}t)]^{-(T-1)/2} \sigma_1^{-1} \sigma_2^{-1}, \quad (20)$$

have been drawn in (broken line). The joint posterior distribution (20) emerges when Bayes theorem is used to combine the likelihood function with so-called diffuse prior distributions for σ_1 and σ_2 . As can be seen, the use of diffuse prior distributions for σ_1 and σ_2 rather than the, admittedly not very informative prior distributions (14), does not change the shapes of the marginal posterior density functions to a great extent. It is striking that all marginal posterior distributions have distinct modes at values which are not implausible from an economist's viewpoint, while the maximum-likelihood procedure invariably produces at least some estimates which are unreasonable. An explanation for this fact may be found by looking at the contours of the likelihood function in appendix B. It clearly shows how for δ_1 -values in the neighborhood of zero – the ordinary CES case admitting ρ_1 to take on any real value – the contours are stretched out. The contours have been drawn for two sets of values of g, ρ_2 and δ_2 for one of which – the first diagram – a conditional maximum for ρ_1 and δ_1 is found as indicated. For the other set the 'maximum' is to be found on the $\delta_1 = 0$ axis. The conclusion seems warranted, that by the process of numerical integration the high but narrow ridge along the $\delta_1 = 0$ axis is given a small weight and the mass of the distribution is found in the lower area more to the left in the figure.

The modes of the respective marginal posterior distributions may be expected to give an indication as to which parameter should preferably be pinpointed in order to obtain conditional maximum-likelihood estimates of the parameters of the two-level CES production function. As the mode of the marginal posterior

distribution of δ_2 deviates substantially from the theoretical value of this parameter computed according to (11) under the assumption of perfectly competitive markets for factor supplies, viz., 0.24, it was decided to experiment with values of δ_2 pinpointed around 0.55, viz., the mode of the posterior distribution of δ_2 . At values below 0.64 for δ_2 the parameter δ_1 attained its lower bound, i.e., zero. The conditional maximum-likelihood estimates for the parameters g , ρ_2 , δ_1 and ρ_1 , given that δ_2 takes on the value 0.64, are given in table 1. The figures between parentheses denote the asymptotic standard errors, which have been calculated via direct estimation of the Hessian matrix of the log-likelihood. For both the method of optimization – the complex method – and the method of computation of these standard errors, the interested reader is referred to Schim van der Loeff and Harkema (1976). As can be seen the standard errors

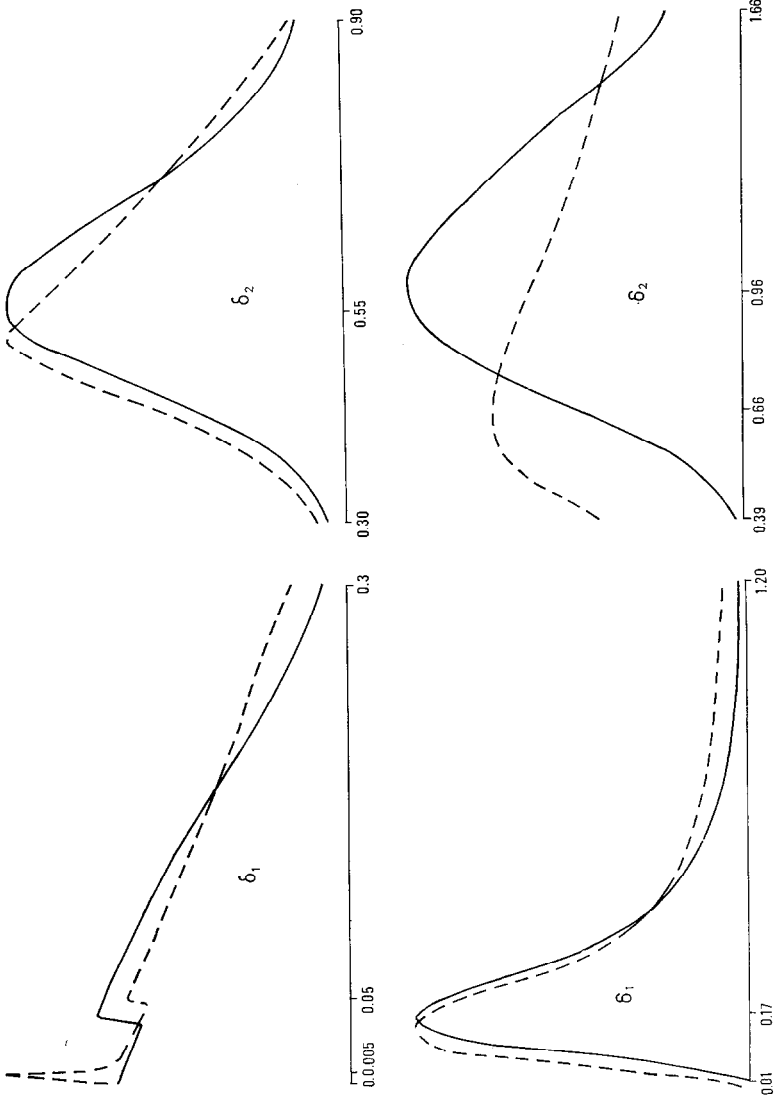
Table 1
Conditional estimates of the parameters
of the two-level CES production function.

g	δ_2	ρ_2	δ_1	ρ_1
0.021	0.64	0.027	0.017	12.27
(0.0018)	(-)	(0.72)	(0.037)	(8.05)

are quite large, which points to multicollinearity as could be expected in view of the short (15 years) time series available. In fact the correlation coefficient between the estimates of δ_1 and ρ_1 , as calculated from the estimated covariance matrix, is -0.92 .

The preceding analysis demonstrates the usefulness of a Bayesian approach, where the structure of the model – in this case intrinsically non-linear and with a definite secondary optimum at $\delta_1 = 0$, where ρ_1 may take on any value – and the lack of sufficiently long time series, provide for a situation where unconstrained maximum-likelihood estimates prove unsatisfactory. It tries to make plausible the point that, apart from its own relevance, a Bayesian approach may also be a useful tool for obtaining (conditional) maximum-likelihood estimates, because of the process of numerical integration, which indicates where the mass of the likelihood function is concentrated.

Appendix A



17/76

Fig. A. Marginal posterior densities of the parameters of the two-level CES production function.

Appendix B

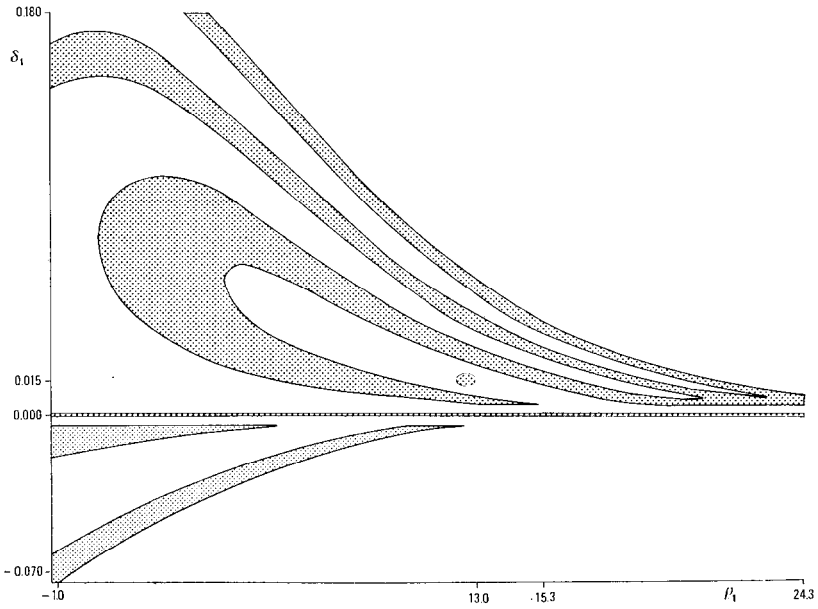


Fig. B-1. Contour lines of the concentrated likelihood function for $g = 0.021$, $\delta_2 = 0.64$ and $\rho_2 = 0.031$.

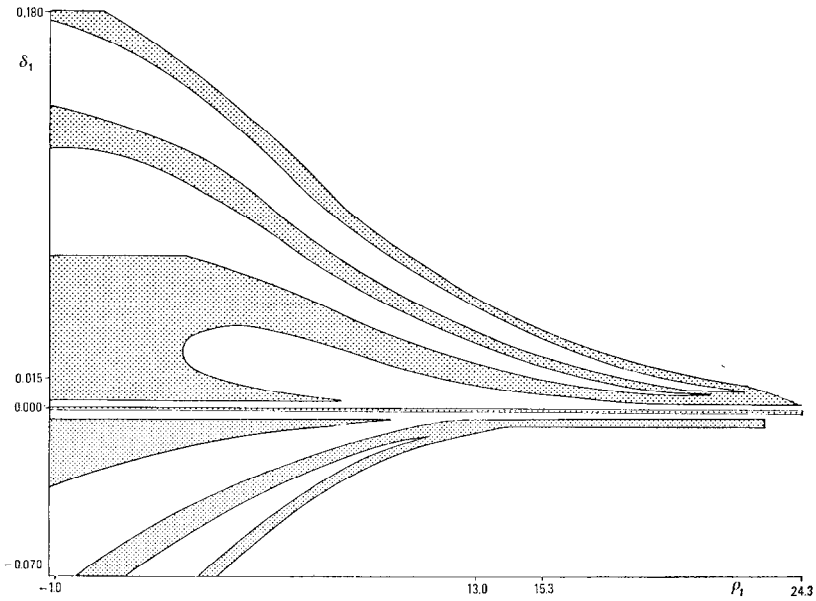


Fig. B-2. Contour lines of the concentrated likelihood function for $g = 0.021$, $\delta_2 = 0.63$ and $\rho_2 = 0.25$.

References

- Hodges, D.J., 1969, A note on estimation of Cobb–Douglas and CES production function models, *Econometrica* 37, no. 4, 721–725.
- Jorgensen, D.W. and Z. Griliches, 1967, The explanation of productivity change, *Review of Economic Studies* 34, 249–283.
- Mizon, G.E., 1974, Factor substitution and returns to scale in a cross section of U.K. industries: An exercise in nonlinear inference, mimeo. (London School of Economics and Political Science, London).
- Sankar, U., 1970, Elasticities of substitution and returns to scale in Indian manufacturing industries, *International Economic Review* 11, no. 3, 399–411.
- Sato, K., 1967, A two-level constant-elasticity-of-substitution production function, *Review of Economic Studies* 34, 201–218.
- Schim van der Loeff, S. and R. Harkema, 1976, Three models of firm behavior; Theory and estimation, with an application to the Dutch manufacturing sector, *Review of Economics and Statistics* 58, no. 1, 13–21.