Report 7407/E

THREE MODELS OF FIRM BEHAVIOUR; THEORY AND ESTIMATION,
WITH AN APPLICATION TO THE DUTCH MANUFACTURING SECTOR.

by

S. Schim van der Loeff and R. Harkema

June 1974
THREE MODELS OF FIRM BEHAVIOUR; THEORY AND ESTIMATION, WITH AN APPLICATION TO THE DUTCH MANUFACTURING SECTOR.

by

S. Schim van der Loeff and R. Harkema

Contents

1. Introduction 1
2. Models of Firm Behaviour 2
   2.1. Deterministic Profit Maximization 2
   2.2. Expected Profit Maximization 4
   2.3. Cost Minimization with Exogenous Output 6
3. The Data and Some Specific Estimation Problems 8
   3.1. The Data 8
   3.2. Consequences of Aggregation and the Use of Index Numbers 9
   3.3. Methods of Estimation 11
4. Estimation 14
   4.1. Deterministic Profit Maximization 14
   4.2. Expected Profit Maximization 16
   4.3. The Cost Minimization Model 16
References 19
Appendix 21

1. INTRODUCTION AND SUMMARY

In the last decade much research has been done in the field of estimating the parameters of production functions. Quite recently stress has been laid on obtaining estimates that are logically consistent within the context of a completely specified theoretical model of firm behaviour. As a first attempt Coen and Hickman (1970) have applied the dynamic adjustment model, suggested by Dhrymes (1967), to the estimation of the parameters of the Cobb-Douglas (CD) function. In this paper the dynamic adjustment model has been used to estimate the parameters of the
Constant Elasticity of Substitution (CES) function. In addition to this theoretical model, two more models of firm behaviour are brought forward and empirically tested, the assumptions underlying these models being deterministic and expected profit maximization, respectively. With regard to the last-mentioned model, a generalization of Hodges' (1969) result is given, to include monopsony situations on the factor markets as well as a monopoly situation in the output market.

In the field of estimation, a method is presented which obtains the asymptotic variance-covariance matrix of the various maximum likelihood estimates by means of a direct estimate of the Hessian matrix of the log-likelihood function. This and the application of a direct search method to compute the maximum of the likelihood function, makes it unnecessary to calculate the untractable first and second order derivatives of a non-linear likelihood function, such as the one at hand and provides a rather easy method to solve similar small-order non-linear estimation problems.

The order of discussion will be as follows. First, the theoretical models will be derived in Section 2. Then, Section 3 will be devoted to the data and a few problems stemming from aggregation and the use of index numbers. Finally, in Section 4 the results of the estimation will be presented.

2. MODELS OF FIRM BEHAVIOUR

2.1. Deterministic Profit Maximization

The model presented in this subsection is the most conventional one of the models considered in this section. It assumes that the entrepreneur has full knowledge of his production function, i.e., he knows, given the factor inputs, exactly what output he will obtain. He can thus maximize profits under the constraint of the CES production function

\[ V_t = y_t e^{\gamma t} \left[ \delta K_t^{-\rho} + (1 - \delta) L_t^{-\rho} \right]^{-\nu/\rho} \quad (t = 1, \ldots, T) \]

1 The authors are indebted to Prof. T. Kloek, who suggested this approach.
where $V_t$ denotes value added in constant prices, $K_t$ and $L_t$ are measures for capital and labor inputs and $\gamma$, $g$, $\delta$, $\rho$, and $\nu$ are parameters, the meaning of which is supposed to be sufficiently known to the reader.

In addition it is assumed in this subsection that the entrepreneur has full knowledge of the price schedules of the output market and the factor markets. He thus knows exactly beforehand the price he will obtain or have to pay for whatever quantity of output or factors he turns out to decide on. The market for output is supposedly characterized as follows

\begin{equation}
V_t = \xi_{1t}^n p_t \quad (t = 1, \ldots, T)
\end{equation}

The factor $\xi_{1t}$ represents the total effect on output demand of variables other than the price variable $p_t$, like e.g. total disposable income. These variables are supposed to be exogenous, while in addition it is assumed that $\xi_{1t}$ can be adequately approximated by

\begin{equation}
\xi_{1t} = c e^{ht} \quad (t = 1, \ldots, T)
\end{equation}

Lest the analysis become too complicated it is assumed that there is perfect competition on the factor markets, i.e., that prices per unit of capital and labor service, $r_t$ and $w_t$, respectively, are exogenously determined. The profit maximizing conditions become, after substitution of (2.1.3) in (2.1.2)

\begin{equation}
r_t = c^{-\frac{1}{\eta_1}} v_{0\gamma}^{-p/v} \xi_1^{-\frac{1}{\eta_1}+p/v} \eta_1 (1+p) \exp\{-h\eta_1 + g \nu^{-1} t\}
\end{equation}

\begin{equation}
w_t = c^{-\frac{1}{\eta_1}} v_{(1-\delta)\gamma}^{-p/v} \xi_1^{-\frac{1}{\eta_1}+p/v} \eta_1 (1+p) \exp\{-h\eta_1 + g \nu^{-1} t\}
\end{equation}

where $\xi_1 = 1 + \eta_1^{-1}$.

In the above model the entrepreneurial services are supposed to consist out of two components, namely, services pertaining to technical and economic expertise, respectively. The entrepreneur's technical expertise, his expertise as an engineer, shows up in his obtaining the most output out of a given set of inputs. His economic expertise becomes apparent in his adequately reacting to prices, when choosing the quantities of inputs. As in reality the entrepreneurial economic
and technical abilities will not be perfect, random terms $u_{0t}$, $u_{1t}$, and $u_{2t}$ are added to the production constraint (2.1.1) and the profit maximizing conditions (2.1.4). Taking the natural logarithm of all three equations, the following mixed structural-reduced form model obtains

\[
\begin{align*}
\ln V_t + \nu \ln K_t + (1 - \delta) \ln L_t - \gamma &= u_{0t} \\
(\xi_1 + \rho \nu^{-1}) \ln V_t - (1 + \rho) \ln K_t &= u_{1t} \\
(\xi_1 + \rho \nu^{-1}) \ln V_t - (1 + \rho) \ln L_t &= u_{2t}
\end{align*}
\]

Proceeding in the line of thought of Hoch (1958), a simultaneous model has thus been specified. This is to say, that in the reduced-form equations for $\ln K_t$ and $\ln L_t$, the term $u_{0t}$ will appear. With respect to the disturbances $u_{0t}$ and $u_{it}$ ($i = 1, 2$), which are supposed to be due to lack of engineering and economic expertise, respectively, it is assumed that $u_{0t}$ is distributed independently of both $u_{it}$ ($i = 1, 2$). This leads to the following specification of the covariance matrix $\phi$ of the disturbances in the simultaneous model (2.1.5)

\[
\phi = \begin{bmatrix}
\sigma_0^2 & 0 & 0 \\
0 & \sigma_1^2 & \omega \\
0 & \omega & \sigma_2^2
\end{bmatrix}
\]

2.2. Expected Profit Maximization

In line with Zellner, Kmenta and Drèze (1966), we assume in this subsection that the entrepreneur has full knowledge of his production function up to a random disturbance $u_{0t}$, representing such factors as weather, unpredictable variations in machine or worker performance etc., in any case factors that are explicitly unknown to the entrepreneur.
This leads to expected profit maximization, or the maximization of the mathematical expectation of profits $\Pi_t$, in time period $t$ ($t = 1, \ldots, T$)

\[(2.2.1) \quad E[\Pi_t] = E[p_t V_t] - E[r_t] K_t - E[w_t] L_t \]

The production constraint in this subsection is

\[(2.2.2) \quad V_t = Y e^{\delta t} [\delta K_t^{-\rho} + (1 - \delta) L_t^{-\rho}]^{-\nu/\rho} e^{u_0 t} \]

The output and factor markets are characterized as follows

\[(2.2.3) \quad K_t = \xi_1 t e^{v_1 t} \quad \text{or} \quad p_t = \xi_1 t e^{v_1 t} \quad \text{or} \quad p_t = \xi_2 t e^{v_2 t} / 2 \]

Assuming that the $u_{it}$ ($i = 0, 1, 2, 3$) are normally and independently distributed with zero mean and variance $\sigma_i^2$ ($i = 0, 1, 2, 3$), (2.2.1) becomes, taking expectations

\[(2.2.4) \quad E[\Pi_t] = \xi_1 t e^{v_1 t} \exp\{\delta \sigma_0^2 \xi_1^2 + i \sigma_1^2 / n_1 - \xi_1 u_0 t\} \]

where $\xi_i = 1 + n_i^{-1} (i = 1, 2, 3)$. Differentiating with respect to $K_t$ and $L_t$, the following profit maximizing conditions are obtained

\[(2.2.5) \quad c_t = e^{c_0 t - \xi_1 t + \nu} \cdot e^{-(\xi_1 + \rho \nu) u_0 t} = 1 \]

\[c_t c_0 t e^{x_0 t} = \exp\{(\xi_1 + \rho \nu) u_0 t\} = 1 \]

where $c_t = e^{c_0 t - \xi_1 t + \nu} \cdot e^{-(\xi_1 + \rho \nu) u_0 t}$ (i = 1, 2, 3) and

\[c_0 t = e^{c_0 t - \xi_1 t + \nu} \cdot e^{-(\xi_1 + \rho \nu) u_0 t} \]

Substituting the production constraint (2.2.2) in the two equations
In (2.2.5), it is seen that the terms with \( u_{\omega t} \) cancel out and consequently \( K_t \) and \( L_t \) are independent of \( u_{\omega t} \).

The assumptions of the expected profit maximization model thus make for a situation in which the simultaneous equation bias does not occur, when the equation (2.2.2) is estimated by itself. It may be expected, however, that these estimates will be less efficient since the extra information embodied in the factor demand equations is neglected. Nevertheless the obvious gain in simplicity makes it worthwhile to compare this model with the other two, substantially more complicated, models in this section.

2.3. Cost Minimization with Exogenous Output

As said in the introduction, Dhrymes (1967) has developed a model of dynamic adjustment to the optimal (in the sense of cost minimizing) levels of factor inputs. This model is similar to the deterministic profit maximization model in the sense that both models take into account that the parameters of the production function also show up in the factor demand equations. As far as the authors are aware, the actual estimation has first been undertaken by Coen and Hickman (1970), with the CD production model. In the cost minimization model it is assumed that entrepreneurs form expectations \( \hat{V}_t (t = 1, ..., T) \) about their output demand in time period \( t \). As, in this paper, entrepreneurs are supposed to produce according to a CES production function schedule it thus holds good that

\[
(2.3.1) \quad \hat{V}_t = \gamma e^{\hat{\gamma} t} \left[ \delta K_t^{-\varepsilon} + (1 - \delta)L_t^{-\varepsilon} \right]^{-\frac{1}{\varepsilon}}
\]

where \( \hat{V}_t \) is given and known.

If the factor markets are typified by perfect competition, factor prices \( r_t \) and \( w_t \) are exogenously determined. The entrepreneurs may then proceed to determine optimal factor employment - given their expectations about factor prices - by minimizing costs. Some algebraic operations yield the following optimal levels of factor employment.

\[
(2.3.2) \quad \ln K_t^* = \rho^{-1} \ln \delta - \nu^{-1} \ln \gamma + \nu^{-1} \ln \hat{V}_t + \rho^{-1} \ln x_{1t} - g\nu^{-1} t
\]

\[
\ln L_t^* = \rho^{-1} \ln(1 - \delta) - \nu^{-1} \ln \gamma + \nu^{-1} \ln \hat{V}_t + \rho^{-1} \ln x_{2t} - g\nu^{-1} t
\]

The derivations are given in Dhrymes (1967).
where $x_{1t}$ and $x_{2t}$ are defined by the equations

$$(2.3.3) \quad x_{1t} = \left(1 + (w_{1t} r_{t}^{-1})^{1-\sigma}\delta(1-\delta)^{-1}\right)^{-\sigma} \text{ and } x_{2t} = x_{1t}(x_{1t} - 1)^{-1}$$

$\sigma$ denoting the elasticity of substitution, $\sigma = (1 + \rho)^{-1}$. Following Dhrymes (1967) the following model of dynamic adjustment shall be postulated

$$(2.3.4) \quad \ln K_{t} = \alpha_{1} \ln K_{t-1} + (1 - \alpha_{1}) \ln K_{t-1} + \omega_{1t}$$

$$\ln L_{t} = \alpha_{2} \ln L_{t-1} + (1 - \alpha_{2}) \ln L_{t-1} + \omega_{2t}$$

where $\alpha_{i}$ $(i = 1, 2)$ denote parameters between zero and one and the $\omega_{it}$ represent random terms. Using (2.3.2) and (2.3.4) the following set of estimating equations is obtained

$$(2.3.5) \quad \ln K_{t} = \alpha_{1} \rho^{-1} \ln \delta - \alpha_{1} \zeta^{-1} \ln \gamma + (1 - \alpha_{1}) \ln K_{t-1} +$$

$$+ \alpha_{1} \rho^{-1} \ln x_{1t} + \alpha_{1} \zeta^{-1} \ln V_{t-1} - \alpha_{1} \gamma^{-1} t + \omega_{1t}$$

$$\ln L_{t} = \alpha_{2} \rho^{-1} \ln(1-\delta) - \alpha_{2} \zeta^{-1} \ln \gamma + (1 - \alpha_{2}) \ln L_{t-1} +$$

$$+ \alpha_{2} \rho^{-1} \ln x_{2t} + \alpha_{2} \zeta^{-1} \ln V_{t-1} - \alpha_{2} \gamma^{-1} t + \omega_{2t}$$

where for simplicity it is assumed that $\hat{V}_{t} = V_{t-1}$ although undoubtedly better specifications of $\hat{V}_{t}$ could be thought of. Finally, it shall be assumed that $\omega_{t} = (\omega_{1t} \omega_{2t})'$ has a binormal distribution with zero mean and covariance matrix $\Sigma$.

Under the assumption that the factor prices are correctly anticipated and given exogenously, (2.3.5) defines a (non-linear) system of reduced-form equations, that can be estimated making use of a non-linear estimation procedure.
3. THE DATA AND SOME SPECIFIC ESTIMATION PROBLEMS

3.1. The Data

The data in this section all apply to the Dutch manufacturing sector. A measuring approach is adopted such that growth in total output is largely explained by growth in total inputs. The principles that should underlie such an approach are excellently set forward in Jorgenson and Griliches (1967). As to the output data it has therefore seemed appropriate to use a Laspeyres index of gross value added against factor prices.

The authors have deemed it inappropriate to use the capital-services-proportional-to-capital-stock approach, the major argument being that this approach does not take into account the degree of utilization of capital stock. It has therefore been decided not to tackle the difficult problem of constructing a time series of capital stock for the Dutch manufacturing sector. Rather an approach first proposed by Hilhorst (1961), was used to obtain a series that was expected to be closer to the "services" concept namely a series on energy consumption of the Dutch manufacturing sector. Much can be and has been said about the appropriateness of using such a series but the disadvantages are outweighed, in the opinion of the authors, by the advantage of having a series that may be expected to be closer to the services concept than a series on capital stock, even when a good one would exist.

The CBS (the Dutch Central Bureau of Statistics) publishes a series of labor volume, where no weighting according to the relative share of each laborservice in the total wage bill has been applied (which is a desideratum in the Jorgenson and Griliches approach). To construct an index of labor services it was thought that corrections should be made with respect to age, sex, and schooling. For these corrections a publication by Kooyman and Merkies (1971) was used. They present correction factors to allow for productivity differences between labor units of different age, sex and schooling. The CBS series allows for "part-time" work, vacations, and sickness- and accidentleave. So applying the Kooyman and Merkies correction to the CBS series, an index of labor services is obtained in which all manyears worked are made comparable by converting them into manyears worked by a forty-year-old male with primary school education.

For the user cost of capital services r the following very simple

---

3 For more details the interested reader is referred to an unpublished working paper by Schim van der Loeff (1974), that can be obtained on request from the Econometric Institute.
formula has been used

\[(3.1.1) \quad r_t = p_t \left[ i_t + \delta \right] \quad (t = 1, \ldots, T)\]

where \( p_t \) represents the price of investment goods, \( i_t \) the long-term interest rate both at time \( t \) and \( \delta \) is the rate of replacement. For \( \delta \) a weighted average has been used of the replacement rates given by Jorgenson and Griliches (1967) for structures and equipment. As for the wage rate, use has been made of the constructed index of labor services and a series on the total wage bill, to obtain an index of remuneration for labor services.

3.2. Consequences of Aggregation and the Use of Index Numbers

As is well known the aggregates for output and inputs of capital and labor services that are appropriate in the CES production function model are generalized means of the form

\[(3.2.1) \quad V_t = \left[ \prod_{i=1}^{I} \left( V_{it}^{-\rho} \right)^{1-\rho} \right]^{1-\rho} \]
\[K_t = \left[ \prod_{i=1}^{I} \left( K_{it}^{-\rho} \right)^{1-\rho} \right]^{1-\rho} \]
\[L_t = \left[ \prod_{i=1}^{I} \left( L_{it}^{-\rho} \right)^{1-\rho} \right]^{1-\rho} \]

where \( V_{it} \), \( K_{it} \), and \( L_{it} \) denote output and inputs of capital and labor services of firm or group of firms \( i \) (\( i = 1, \ldots, I \)) in time period \( t \). Under fairly general conditions it can be shown that the relative error made in using the published arithmetic means instead of the generalized means (3.2.1), is proportional to the second order sample moments of the natural logarithms of \( V_{it} \), \( K_{it} \), and \( L_{it} \) respectively.\(^5\) This is to say that if for \( V_t \), \( K_t \), and \( L_t \) the published arithmetic averages are used, the production function should be written as

\[^4\] This formula for the user cost of capital resembles that given by Jorgenson and Griliches (1967), and is used by a.o. Barten (1972).

\[^5\] See Schim van der Loeff and Harkema (1973). It can be shown that if \( V_{it} \), \( K_{it} \) and \( L_{it} \) are lognormally distributed the relationships (3.2.3) hold exactly, while in any case (3.2.3) is valid up to terms in third and higher order moments.
\begin{align}
(3.2.2) \quad c_0 V_t &= \gamma e^{gt} [\delta (c_1 K_t)^{-\rho} + (1 - \delta)(c_2 L_t)^{-\rho}]^{-\nu/\rho} \\
\text{where the terms } c_i (i = 0, 1, 2) \text{ are defined by} \\
(3.2.3) \quad c_0 &= \exp(i(-1 - \nu - \sigma)I^{-1} \sum_{i=1}^I (\ln V_{it} - \ln G_V)^2) \\
&= \exp(i(-1 - \nu)I^{-1} \sum_{i=1}^I (\ln K_{it} - \ln G_K)^2) \\
&= \exp(i(-1 - \nu)I^{-1} \sum_{i=1}^I (\ln L_{it} - \ln G_L)^2)
\end{align}

where \( G_V, G_K, \) and \( G_L \) denote the geometric averages of the \( V_{it}, K_{it}, \) and \( L_{it} \) respectively. It should be noted that in not adding time subscripts to the \( c_i \) \( (i = 0, 1, 2) \) in (3.2.3) it has been implicitly assumed that the second order moments stay the same over time. A justification for this assumption may be found in Cramer (1971, p.181).

In that case the parameters \( \rho, \nu \) and \( \sigma \) are left unaffected by the aggregation errors, while the parameters \( \gamma \) and \( \delta \) become

\begin{align}
(3.2.4) \quad \gamma' &= c_0^{-1} \gamma [\delta c_1^{-\rho} + (1 - \delta)c_2^{-\rho}]^{-\nu/\rho} \\
\delta' &= c_1^{-\rho} \delta [\delta c_1^{-\rho} + (1 - \delta)c_2^{-\rho}]^{-1}
\end{align}

Another complication arises because the data are specified in the form of index numbers. Dividing (3.2.2) through by the relation in the base year \( t = 0 \), one obtains

\begin{align}
(3.2.5) \quad \left( \frac{V_t}{V_0} \right) &= e^{gt} [\delta''(K_t/K_0)^{-\rho} + (1 - \delta'')(L_t/L_0)^{-\rho}]^{-\nu/\rho} \\
\text{where} \\
(3.2.6) \quad \delta'' &= \delta (c_1 K_0)^{-\rho} [\delta (c_1 K_0)^{-\rho} + (1 - \delta)(c_2 L_0)^{-\rho}]^{-1}
\end{align}

As can be seen the use of index numbers prevents the parameter \( \gamma \) from being estimated while the parameter \( \delta \) obtains a much less clearcut interpretation than the name "distribution parameter" leads to believe. A further complication of the use of index numbers may be illustrated.
as follows. Consider the following relation between two quantities $y_t$ and $x_t$.

\[ (3.2.7) \quad \ln y_t = \alpha_0 + \alpha_1 \ln x_t + u_t \quad (t = 0, 1, \ldots, T) \]

where $u_t$ is a random term with zero mean and variance $\sigma^2$, which is supposed to be serially uncorrelated. If $y_t$ and $x_t$ are only available in index numbers, subtracting the relation in the base period yields

\[ (3.2.8) \quad \ln(y_t/y_0) = \alpha_1 \ln(x_t/x_0) + u_t - u_0 \quad (t = 1, \ldots, T) \]

The covariance matrix of the transformed disturbances $v_t = u_t - u_0$ becomes up to the scalar multiple $\sigma^2$

\[ (3.2.9) \quad A = [I + 11'] \]

where $1$ is the column vector consisting of ones only ($1 = (1 \ 1 \ \ldots \ 1)'$). The inverse, determinant and two matrices $V$ such that $VV' = A$, are

\[ (3.2.10) \quad \det(A) = T + 1; \ A^{-1} = [I - (T + 1)^{-1}11']; \]

\[ V = [I - (T+1) + \sqrt{T+1} \cdot (T+1)11'] \]

3.3. Methods of Estimation

To estimate the parameters of the three models of the preceding section, use has been made of the (full-information) maximum likelihood method. To maximize a non-linear likelihood function several methods exist which involve computing first and second order derivatives of the function. In the case of the production function models discussed in this paper, this leads to rather intricate and untractable formulae. Therefore a direct search method has been used namely the "Complex" method a description of which can be found in Box (1965) or in the aforementioned working paper. This method provides a relatively simple algorithm to solve very involved non-linear maximum-likelihood problems. With most non-linear methods it has in common that no information is obtained about the standard deviations of the estimates. In order to
fill this gap, the following paragraph describes a method which estimates the Hessian matrix of the log-likelihood function without it being necessary to compute the second-order derivatives.

Denoting the parameter vector of the likelihood function by $\beta$ and the vector of maximum likelihood estimates of $\beta$ by $b$, we may expand the log-likelihood function in a Taylor series around $\beta = b$. This yields

\begin{equation}
\ln \mathcal{L}(\beta) = \ln \mathcal{L}(b) + (\beta - b)' \left[ \frac{\partial \ln \mathcal{L}(\beta)}{\partial \beta} \right]_{\beta = b} + \frac{1}{2} (\beta - b)' \left[ \frac{\partial^2 \ln \mathcal{L}(\beta)}{\partial \beta \partial \beta'} \right]_{\beta = b} (\beta - b) + R_3
\end{equation}

where $R_3$ denotes the remainder term.

As $b$ is the vector of maximum likelihood estimates, the first-order derivatives of $\ln \mathcal{L}(\beta)$ with respect to $\beta$ in $\beta = b$ will be zero. In addition, taking $\beta$ in a sufficiently small neighborhood around $b$, it may be assumed that the remainder term of (3.3.1) is small. Then, (3.3.1) may be written as

\begin{equation}
\ln \mathcal{L}(\beta) - \ln \mathcal{L}(b) \approx \frac{1}{2} (\beta - b)' \left[ \frac{\partial^2 \ln \mathcal{L}(\beta)}{\partial \beta \partial \beta'} \right]_{\beta = b} (\beta - b)
\end{equation}

where it should be noted that the right-hand side of (3.3.2) is linear in the elements of the Hessian matrix. If the vectors $\beta$ and $b$ contain $N$ elements, we may arrange the $N(N+1)$ distinct elements of the Hessian matrix in a vector $a$. Similarly, after attaching a factor one half to the squares $(\beta_i - b_i)^2$ $(i = 1, ..., N)$, the $N(N+1)$ products $(\beta_i - b_i)(\beta_j - b_j)$ $(i, j = 1, ..., N; i \neq j)$ may be arranged in a vector $z$. Therefore (3.3.2) may be rewritten in the form of the well-known linear model as follows

\begin{equation}
\ln \mathcal{L}(\beta) - \ln \mathcal{L}(b) = z' a + v
\end{equation}

In order to estimate the vector $a$ we proceed as follows. According to an $N$-dimensional uniform distribution we draw vectors $\beta^r$ $(r = 1, ..., R)$ from a sufficiently small neighborhood of the vector $b$ and compute the

\footnote{This method has been suggested by Prof. T. Kloek.}
differences \( \ln \ell(\beta^r) - \ln \ell(b) \) \((r = 1, \ldots, R)\). The vectors \( \beta^r \) are used to construct vectors \( z \) as above and these are arranged in the \((R \times N(N + 1))\)-matrix \( Z \) and the differences \( \ln \ell(\beta^r) - \ln \ell(b) \) in the \( R \)-dimensional vector \( y \). Regressing \( y \) on \( Z \) yields our estimates of the \( N(N + 1) \) distinct elements of the Hessian matrix, arranged in the vector \( a \).

As is well known\(^7\), the asymptotic distribution of the maximum likelihood estimators \( b \) is multivariate normal with mean \( \beta \) and variance-covariance matrix \( V \), where

\[
V = - \left\{ E \left[ \frac{\partial^2 \ln \ell(\beta)}{\partial \beta \partial \beta'} \right] \right\}^{-1}
\]

A consistent estimate of \( V \) is given by the matrix \( \hat{V} \), where

\[
\hat{V} = - \left\{ \left[ \frac{\partial^2 \ln \ell(\beta)}{\partial \beta \partial \beta'} \right]_{\beta=b} \right\}^{-1}
\]

Evidently, a consistent estimate of the asymptotic variance-covariance matrix of the maximum likelihood estimates may be obtained by reverting the sign of the estimates of the elements of the Hessian matrix and inverting the matrix thus obtained. In practice, the actual calculations have been performed on the so called concentrated likelihood function. A justification for this procedure, which simplifies matters considerably may be found in Koopmans, Rubin and Leipnik (1950, pp. 148-153).

In the preceding paragraphs a relatively easy way of obtaining maximum likelihood estimates and their asymptotic covariance matrix has been described. It is obvious that it is not clear a priori how large the neighborhood from which the \( \beta^r \) are drawn, should be. Certainly it should not be too large, lest the Taylor approximation be invalid. The policy has been to select a neighborhood within which the estimate of \( a \) in (3.3.3) proved not to be too sensitive to changes in the boundaries chosen. This usually proved to be well within a \( \sigma \)-boundary (within one unit of the estimated standard deviation). The interpretation of the standard deviations given with the estimates in the next section should be done with some caution. All that can be said for now is that further theoretical research will have to be done in order to find the optimal neighborhood from which should be sampled.

\(^7\) The argument that follows is based on Theil (1971), Ch. 8.
4. ESTIMATION

4.1. Deterministic Profit Maximization

In this subsection the estimation results of the model (2.1.5) will be presented. In index numbers (2.1.5) becomes

\[
\ln(V_t/V_0) + \nu^{-1}\ln[\delta''(K_t/K_0)^{-\rho} + (1 - \delta'')(L_t/L_0)^{-\rho}] - g_t = v_{0t}
\]

\[
(1 + \nu^{-1} + \eta^{-1}_1)\ln(V_t/V_0) - (1+\rho)\ln(K_t/K_0) - (\rho\nu^{-1} + \eta^{-1}_1)t - \ln(r_t/r_0) = v_{1t}
\]

\[
(1 + \nu^{-1} + \eta^{-1}_1)\ln(V_t/V_0) - (1+\rho)\ln(L_t/L_0) - (\rho\nu^{-1} + \eta^{-1}_1)t - \ln(w_t/w_0) = v_{2t}
\]

where $\delta''$ is defined as in (3.2.6) and $v_{it} = u_{it} - u_{i0}$ ($i = 0, 1, 2$).

Combining the arguments that led to the specification of (2.1.6) and (3.2.9), the variance-covariance matrix of the transformed disturbances of the system (4.1.1) becomes

\[
\Omega = A \otimes \phi
\]

where $\otimes$ denotes the Kronecker product.

The full-information maximum likelihood method calls for maximizing

\[
f(y_1, \ldots, y_T) = (2\pi)^{-T/2}|A|^{-3/2}|\phi|^{-T/2}\exp\{-\frac{1}{2}\epsilon'\Omega^{-1}\epsilon\} \prod_{t=1}^{T} |J_t|
\]

where $\epsilon = (\epsilon_1', \epsilon_2', \ldots, \epsilon_T')'$ if $\epsilon_t = (v_{0t}', v_{1t}', v_{2t}')'$,

$y_t = [\ln(V_t/V_0)', \ln(K_t/K_0)', \ln(L_t/L_0)']'$, and $|J_t|$ is the Jacobian of the transformation from $\epsilon_t$ to $y_t$. Application of the method of stepwise maximization yields the following function to be minimized with respect to the parameters $\rho, \nu, \delta'', g, \eta_1$, and $h$

\[
f^*(\rho, \nu, \delta'', g, \eta_1, h) = s_0^2(s_1^2s_2^2 - r^2)[(1 + \rho)(1 - \nu - \nu \eta^{-1})]^{-2}
\]

the expression between square brackets being the Jacobian $|J| = |J_t|$ ($t = 1, \ldots, T$), and $s_i^2$ ($i = 0, 1, 2$) and $r$ being equal to
After minimizing (4.1.4) with respect to all six parameters some estimates turned out to be unacceptable (e.g. a negative rate of disembodied technical progress). As this is probably due to multicollinearity (see Table 4.2 below), it was decided to pinpoint $g$ at a predetermined value of 1.6%. The resulting conditional estimates are presented in Table 4.1.

**TABLE 4.1. CONDITIONAL ESTIMATES OF THE DETERMINISTIC PROFIT MAXIMIZATION MODEL**

<table>
<thead>
<tr>
<th>$\hat{\rho}$</th>
<th>$\hat{\gamma}$</th>
<th>$\hat{\sigma}^2$</th>
<th>$\hat{g}$</th>
<th>$\hat{\eta}_1$</th>
<th>$\hat{h}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.07</td>
<td>1.44</td>
<td>0.47</td>
<td>0.016</td>
<td>-0.15</td>
<td>0.069</td>
</tr>
<tr>
<td>(0.05)</td>
<td>(0.04)</td>
<td>(0.02)</td>
<td>(0.13)</td>
<td>(0.005)</td>
<td></td>
</tr>
</tbody>
</table>

The figures between brackets denote conditional standard deviations that have been computed according to the estimation method described in Section 3. To be more specific, the standard deviations shown with the present as well as with the following two models have been estimated by drawing 1500 times according to a uniform distribution from intervals centered around the parameter estimates and with length equal to one third of the eventually estimated values of the standard deviations. In most cases the t-ratios are rather high which may have been caused by the conditional character of the standard deviations. The degree of correlation between $\bar{g}$ and $\bar{\gamma}$ is illustrated in Table 4.2 where $g$ has been pinpointed at several values.

**TABLE 4.2. EFFECTS ON $\bar{\gamma}$ OF CHANGING $\bar{g}$**

<table>
<thead>
<tr>
<th>$\bar{g}$</th>
<th>$\bar{\gamma}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.2 %</td>
<td>1.55</td>
</tr>
<tr>
<td>1.6 %</td>
<td>1.44</td>
</tr>
<tr>
<td>2.2 %</td>
<td>1.25</td>
</tr>
<tr>
<td>2.6 %</td>
<td>1.12</td>
</tr>
<tr>
<td>3.0 %</td>
<td>1.00</td>
</tr>
</tbody>
</table>
4.2. Expected Profit Maximization

It has been shown that consistent estimates for the parameters of the production function can be obtained by direct estimation of the model (2.2.2). The maximum likelihood estimates of this model are presented below. Application of the stepwise maximization principle leads to the following function to be minimized with respect to the parameters \( \rho, \nu, \delta'', \) and \( g \)

\[
(4.2.1) \quad f^{*}(\rho, \nu, \delta'', g) = s_0^2
\]

where \( s_0^2 \) is defined as in (4.1.5). Minimization of this function with the "Complex"-method yields the parameter estimates presented in Table 4.3.

**TABLE 4.3. ESTIMATES OF THE PARAMETERS OF THE EXPECTED PROFIT MAXIMIZATION MODEL**

<table>
<thead>
<tr>
<th>( \hat{\rho} )</th>
<th>( \hat{\nu} )</th>
<th>( \hat{\delta}' )</th>
<th>( \hat{g} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.05</td>
<td>1.42</td>
<td>0.47</td>
<td>0.016</td>
</tr>
<tr>
<td>(0.20)</td>
<td>(0.10)</td>
<td>(0.04)</td>
<td>(0.003)</td>
</tr>
</tbody>
</table>

Comparing the estimates in the foregoing subsection with those in Table 4.3, it can be seen that not using the extra information embodied in the factor demand equations yields estimates that are less precise. This is especially noticeable for the estimate of \( \rho \), the substitution parameter. The evidence is not conclusive, however, because in the preceding subsection the value of \( g \) was pinpointed. The degree of correlation between \( \hat{g} \) and \( \hat{\nu} \) as calculated from the estimated variance-covariance matrix amounts to -0.4 in the present model.

4.3. The Cost-Minimization Model

Subtracting from (2.3.5) the equations in the base period yields the estimating equations in index-numbers
\begin{align*}
(4.3.1) \quad & \ln(K_t/K_0) = (1 - \alpha_1) \ln(K_{t-1}/K_{-1}) + \alpha_1 \rho^{-1} \ln x_{1t}^i + \\
& + \alpha_1 \nu^{-1} \ln(V_{t-1}/V_{-1}) - \alpha_1 \eta^{-1} \ln x_{1t}^i + v_{1t} - \psi_{10} \\
\ln(L_t/L_0) = (1 - \alpha_2) \ln(L_{t-1}/L_{-1}) + \alpha_2 \rho^{-1} \ln x_{2t}^i + \\
& + \alpha_2 \nu^{-1} \ln(V_{t-1}/V_{-1}) - \alpha_2 \eta^{-1} \ln x_{2t}^i + v_{2t} - \psi_{20}
\end{align*}

where the \( x_{it}^i \) (\( i = 1,2 \)) are defined by

\begin{align*}
(4.3.2) \quad & x_{1t}^i = \epsilon(1 + (1 - \epsilon) \epsilon^{-1})[\ln(w_t/w_0)(r_t/r_0)^{-1}]^{1-\sigma} \\
& x_{2t}^i = (1 - \epsilon)x_{1t}^i(x_{1t}^i - \epsilon)^{-1}
\end{align*}

\[ \epsilon = (r_0/w_0)^{1-\sigma}[(1 - \delta')/\delta']^{-\sigma} + (1 + (r_0/w_0)^{1-\sigma}[(1 - \delta')/\delta']^{-\sigma}]^{-1} \]

It was not assumed a priori that the disturbances were serially uncorrelated. A first order Markov scheme was tried in the estimation but produced unacceptable estimates. In the end the disturbances in \((4.3.1)\) have been assumed to be normally distributed with variance-covariance matrix

\[(4.3.3) \quad E[w w'] = A \otimes \Sigma\]

where \( w = (w_{11} - w_{10}, w_{21} - w_{20}, \ldots, w_{1T} - w_{10}, w_{2T} - w_{20})' \). Application of the concentrated likelihood method leads to the minimization of the function

\[(4.3.4) \quad f^*(\alpha_1, \alpha_2, \rho, \nu, \epsilon, g) = s_1^2 s_2^2 - r^2 \]

where \( s_i^2 (i = 1,2) \) and \( r \) are similarly defined as in \((4.1.5)\). The minimization of \((4.3.4)\) with respect to all six parameters again yielded estimates with a wrong sign or boundary solutions. It was decided to pinpoint the value of \( \alpha_1 \) at 1. The argument behind this choice is the following. Because of the way the capital input data are obtained, they may be expected to take into account the degree of

---

\( ^8 \) A problem that was sidestepped, is that in the year \( t = -1 \) one of the explanatory variables is correlated with the disturbance term.
utilization of the capital stock. In a period of expansion the manufacturing industries are supposed to foresee the need for capital services sufficiently to adjust the utilization of their capital stock within one year. Other values for \( a_1 \) were tried, but did not produce great differences in the estimated values of the other parameters. Furthermore a value of 3% was chosen for \( g \), since a value of 1.6% for \( g \) yielded an estimate for \( \rho \) so close to zero, that it would have been impossible to avoid discontinuities in drawing from an area sufficiently large around \( \hat{\theta} \). The estimates are presented in Table 4.4.

**TABLE 4.4. CONDITIONAL ESTIMATES OF THE COST-MINIMIZATION MODEL**

<table>
<thead>
<tr>
<th>( \hat{a}_1 )</th>
<th>( \hat{a}_2 )</th>
<th>( \hat{\beta} )</th>
<th>( \hat{\vartheta} )</th>
<th>( \hat{\varepsilon} )</th>
<th>( \hat{g} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.00</td>
<td>0.16</td>
<td>-0.05</td>
<td>1.34</td>
<td>0.39</td>
<td>0.03</td>
</tr>
<tr>
<td>(0.08)</td>
<td>(0.08)</td>
<td>(0.28)</td>
<td>(0.11)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Most surprising is the low value of the estimate for \( a_2 \). However, pinpointing \( a_2 \) at values between one and one-half did not have a substantial effect on the estimates of the other parameters. The high value of the standard deviation of \( \vartheta \) makes this the only model where the returns to scale are not significantly different from unity. Again the evidence is troubled by multicollinearity. The high correlation between the estimates presents a problem also for the estimation of the variance-covariance matrix, since the estimated Hessian is almost singular. The correlation coefficient between the estimates of \( \varepsilon \) and \( \vartheta \) is by way of example -0.9.
REFERENCES


**APPENDIX * **

**TIME SERIES FOR THE DUTCH MANUFACTURING SECTOR 1949-1972**

<table>
<thead>
<tr>
<th>Year</th>
<th>A.1</th>
<th>A.2</th>
<th>A.3</th>
<th>A.4</th>
<th>A.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1949</td>
<td>42.</td>
<td>50.0</td>
<td>77.6</td>
<td>59.0</td>
<td>31.2</td>
</tr>
<tr>
<td>1950</td>
<td>47.</td>
<td>54.5</td>
<td>80.6</td>
<td>62.6</td>
<td>34.1</td>
</tr>
<tr>
<td>1951</td>
<td>49.</td>
<td>57.5</td>
<td>82.2</td>
<td>70.4</td>
<td>37.7</td>
</tr>
<tr>
<td>1952</td>
<td>49.</td>
<td>58.7</td>
<td>80.8</td>
<td>76.1</td>
<td>39.8</td>
</tr>
<tr>
<td>1953</td>
<td>54.</td>
<td>61.2</td>
<td>82.8</td>
<td>71.5</td>
<td>41.0</td>
</tr>
<tr>
<td>1954</td>
<td>59.</td>
<td>65.6</td>
<td>85.9</td>
<td>70.9</td>
<td>46.4</td>
</tr>
<tr>
<td>1955</td>
<td>64.</td>
<td>69.5</td>
<td>88.1</td>
<td>75.8</td>
<td>51.6</td>
</tr>
<tr>
<td>1956</td>
<td>68.</td>
<td>74.8</td>
<td>89.6</td>
<td>90.1</td>
<td>58.0</td>
</tr>
<tr>
<td>1957</td>
<td>71.</td>
<td>75.3</td>
<td>90.4</td>
<td>107.2</td>
<td>65.0</td>
</tr>
<tr>
<td>1958</td>
<td>71.</td>
<td>77.7</td>
<td>89.9</td>
<td>101.6</td>
<td>66.8</td>
</tr>
<tr>
<td>1959</td>
<td>78.</td>
<td>81.1</td>
<td>92.0</td>
<td>91.9</td>
<td>68.3</td>
</tr>
<tr>
<td>1960</td>
<td>87.</td>
<td>87.6</td>
<td>95.4</td>
<td>95.2</td>
<td>75.8</td>
</tr>
<tr>
<td>1961</td>
<td>90.</td>
<td>89.9</td>
<td>97.6</td>
<td>89.1</td>
<td>83.7</td>
</tr>
<tr>
<td>1962</td>
<td>95.</td>
<td>94.9</td>
<td>98.4</td>
<td>95.2</td>
<td>90.9</td>
</tr>
<tr>
<td>1963</td>
<td>100.</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
</tr>
<tr>
<td>1964</td>
<td>110.</td>
<td>107.7</td>
<td>100.6</td>
<td>116.7</td>
<td>115.8</td>
</tr>
<tr>
<td>1965</td>
<td>116.</td>
<td>114.0</td>
<td>102.8</td>
<td>125.5</td>
<td>127.7</td>
</tr>
<tr>
<td>1966</td>
<td>122.</td>
<td>120.8</td>
<td>102.5</td>
<td>138.5</td>
<td>142.3</td>
</tr>
<tr>
<td>1967</td>
<td>126.</td>
<td>132.2</td>
<td>101.5</td>
<td>137.0</td>
<td>154.7</td>
</tr>
<tr>
<td>1968</td>
<td>139.</td>
<td>147.7</td>
<td>101.9</td>
<td>134.9</td>
<td>168.0</td>
</tr>
<tr>
<td>1969</td>
<td>153.</td>
<td>165.8</td>
<td>94.8</td>
<td>149.0</td>
<td>195.4</td>
</tr>
<tr>
<td>1970</td>
<td>166.</td>
<td>179.5</td>
<td>95.5</td>
<td>166.1</td>
<td>225.9</td>
</tr>
<tr>
<td>1971</td>
<td>172.</td>
<td>193.1</td>
<td>95.3</td>
<td>163.9</td>
<td>253.7</td>
</tr>
<tr>
<td>1972</td>
<td>181.</td>
<td>212.3</td>
<td>94.3</td>
<td>147.3</td>
<td>280.7</td>
</tr>
</tbody>
</table>

A.1 Index of the industrial production in the manufacturing sector.
A.2 Index of inputs of capital services in the manufacturing sector.
A.3 Index of inputs of labor services in the manufacturing sector.
A.4 Index of the user cost of capital.
A.5 Index of the remuneration of labor services in the manufacturing sector.

* For more details and sources, the interested reader is referred to Schim van der Loeff (1974).