ETLA ELINKEINOELÄMÄN TUTKIMUSLAITOS
THE RESEARCH INSTITUTE OF THE FINNISH ECONOMY
Lönnrotinkatu 4 B 00120 Helsinki Finland Tel. 609900 Telefax 601753

Keskusteluaiheita - Discussion papers

No. 387

Juha Kettunen*

A PROGRAMME FOR ESTIMATING NONLINEAR MAXIMUM LIKELIHOOD MODELS WITH AN APPLICATION TO DURATION MODELS OF UNEMPLOYMENT

* A grant from the Yrjö Jahnsson Foundation is gratefully acknowledged.

ISSN 0781-6847 27.11.1991

KETTUNEN, Juha, **A PROGRAMME FOR ESTIMATING NONLINEAR MAXIMUM LIKELIHOOD MODELS WITH AN APPLICATION TO DURATION MODELS OF UNEMPLOYMENT**. Helsinki : ETLA, Elinkeinoelämän Tutkimuslaitos, The Research Institute of the Finnish Economy, 1991. 25 p. (Keskusteluaiheita, Discussion Papers, ISSN 0781-6847; no. 387).

ABSTRACT: This paper provides a useful tool for estimating parameters of a general user specified Maximun Likelihood model. Statistical packages often have restrictions regarding the parametrisation of the likelihood function and the development of statistical tests. These practical problems are avoided by a programme, which is written using the SAS/IML matrix language. Applications to unemployment duration models are provided using Finnish microeconomic data. An exponential model of unemployment is estimated and the correction for neglected heterogeneity is made assuming that the effect of omitted variables has a gamma distribution.

1. Introduction

This paper deals with the estimation of unemployment duration models using microeconomic data. The Finnish data is collected from persons becoming unemployed in 1985. It is compiled from various registers and it is more reliable than the data from surveys based on interviews. However, in an econometric analysis relevant variables will often be omitted, either because they are unmeasurable or because their importance is unsuspected. It is well known that omitted variables bias the parameter estimates of duration models. The purpose of the econometric application in this paper is to take the neglected heterogeneity into account in estimation. An exponential model allowing for neglected heterogeneity is estimated assuming that the effect of omitted variables has a gamma distribution across individuals.

An estimation of a user specified maximum likelihood model is often difficult or impossible using standard statistical software. This paper provides the needed statistical tool, which enables estimation of the unknown parameters of the user specified model with an application to duration models. However, the programme is a useful tool to a wider class of maximum likelihood models and a basis for further development. The programme is written using the SAS/IML (1985) matrix language, which is close to matrix algebra notation. The programme can be easily translated with a little effort to any other matrix language.

The paper is organized as follows: An exponential model of unemployment duration is estimated in section 2 with the allowance for gamma heterogeneity across individuals. In section 3 an overview of the structure of the programme and the algorithm is presented. Appendix 2 presents the programme and section 4 concludes the study.

2. Parametric duration models of unemployment

2.1. A general form of the duration model

A general form for the likelihood function of parametric duration models with censored data is provided before the parametrisation of the distribution. Let us consider independent pairs of independent random variables T and Z, where T is the duration variable of primary interest and Z is a censoring variable. A censoring time or a duration and a censoring indicator are observed, $t = \min(T, Z)$, with the indicator for complete spells c. If $T < Z$, then $c = 1$ and otherwise $c = 0$.

The survivor function of T is equal to one minus the distribution function of the duration variable, which can be written as

$$S(t) = e^{-I(t)} \tag{1}$$

and the density function can be written

$$f(t) = h(t)e^{-I(t)} \tag{2}$$

for $t \geq 0$. I(t) is the integrated hazard

$$I(t) = \int_0^t h(\tau)d\tau. \tag{3}$$

The likelihood contribution of an individual can then be written in view of the above definitions as

$$(4) \qquad \ell(t) = h(t)^{c} e^{-I(t)},$$

which is a general form for the duration models with right censored data. The distribution of unemployment duration has to be specified. To estimate the unknown parameters the hazard function and integrated hazard are substituted into the likelihood function (4).

2.2. An exponential model allowing for gamma heterogeneity

Econometric duration models are specified in terms of the hazard function h(t), which is the conditional probability that the person leaves unemployment at t given that he still is unemployed. A commonly applied specification is the proportional hazard model, where the hazard function factors into the product of a function of duration time t and a function of the regressors **x**

$$(5) \qquad h(t) = h_0(t)h_1(\mathbf{x}),$$

where $h_0(t)$ is called the baseline hazard.

The simplest form of parametric duration models is considered in this paper. The base-line hazard function of an exponential model $h_0(t) = 1$. There are a lot of generalizations of the exponential model assuming different specifications for the base-line hazard e.g. Weibull, gamma and Gompertz models. The hazard function of an exponential distribution is constant and it can simply be written

$$(6) \qquad h(t) = \phi.$$

The explanatory variables are introduced into the model by a log-linear form $\phi = e^{x\beta}$. The survivor, density and hazard functions of the exponential distribution can be written as

$$(7) \qquad S(t) = e^{-I(t)}$$

$$(8) \qquad f(t) = e^{x\beta - I(t)}$$

$$(9) \qquad h(t) = e^{x\beta},$$

where I(t) is the integrated hazard

$$(10) \qquad I(t) = \int_0^t h(\tau)d\tau + C.$$

The constant C is chosen so that $I(0) = 0$, then the integrated hazard of the exponential distribution can be written as

$$(11) \qquad I(t) = te^{x\beta}.$$

It is inevitable that in an econometric analysis relevant variables will be omitted, either because they are unmeasurable or because their importance is unsuspected.

Unobserved heterogeneity is widely discussed in the econometric literature. Lancaster (1979) assumed a parametric functional form for the pattern of heterogeneity. The gamma mixing distribution was chosen because it is analytically simple to use and it provides quite a flexible model for the distribution of the heterogeneity component. Lancaster found that the estimated parameters were biased towards zero if the unobserved heterogeneity was not controlled. Even if the omitted variables are uncorrelated with those which are included in the model, the parameters will be biased towards zero as noted by Nickell (1979). The method of correcting for gamma heterogeneity has been widely used with duration distributions [e.g. Kooreman and Ridder (1983), Newman and McCulloch (1984), Narendranathan, Nickell and Stern (1985), Engström and Löfgren (1987)].

Suppose the individuals of the sample differ to some degree with respect to some unobservable variable, say, motivation v. Each individual has his own v and hence his own hazard function h(t). The conditional hazard in a Weibull model allowing for gamma heterogeneity is

$$(12a) \qquad h(t|v) = vh(t),$$

where v has a gamma density

$$(12b) \qquad g(v) = \frac{\epsilon^{\mu}}{\Gamma(\mu)} v^{\mu-1} e^{-\epsilon v} \quad \text{with} \quad \Gamma(\mu) = \int_0^{\infty} w^{\mu-1} e^{-w} dw.$$

The expected value of the heterogeneity component $E(v) = \mu/\epsilon$ is normalized to one by setting $\epsilon = \mu$ and its variance, i.e.

$\sigma^2 = 1/\mu$, is estimated. The marginal survivor function, not conditional on v, is obtained by integrating over the assumed mixing distribution. The density function is got from the survivor function by differentiating $f(t) = - \partial S(t)/\partial t$ and the hazard function is got as a ratio $h(t) = f(t)/S(t)$. The exponential distribution allowing for unobserved gamma heterogeneity across individuals gives the following survivor, density and hazard functions

$$(13) \qquad S_g(t) = [1 + \sigma^2 I(t)]^{-1/\sigma^2}$$

$$(14) \qquad f_g(t) = h(t)[1 + \sigma^2 I(t)]^{-1/\sigma^2 - 1}$$

$$(15) \qquad h_g(t) = h(t)[1 + \sigma^2 I(t)]^{-1},$$

where I(t) is the integrated hazard of the original exponential distribution (3). The integrated hazard with gamma heterogeneity can be written as $I_g(t) = -\log[S_g(t)]$, which can be rewritten

$$(16) \qquad I_g(t) = 1/\sigma^2 \log[1 + \sigma^2 I(t)].$$

The integrated hazards I(t) and $I_g(t)$ are the generalized residuals of these models in the sense of Cox and Snell (1968).

To write the likelihood functions and estimate the unknown parameters, the hazard functions and the integrated hazards of the two models presented are substituted into the likelihood contribution (4). For completeness the log likelihood functions which are maximized are presented. The log likelihood function of the exponential model can be written

$$(17) \qquad L(t) = \sum_{i=1}^{n} [c_i x_i \beta - t_i e^{x_i \beta}]$$

and the log likelihood function of the exponential model with gamma heterogeneity can be written as

$$(18) \qquad L_g(t) = \sum_{i=1}^{n} [c_i x_i \beta - (c_i + 1/\sigma^2)\log(1 + \sigma^2 t_i e^{x_i \beta})].$$

The expected length of unemployment given by the exponential model is obtained by integrating the survivor function S(t) from nought to infinity, which gives

$$(19) \qquad E(T) = e^{-x\beta}.$$

Let $u = 1 + \sigma^2 t e^{x\beta}$, then the corresponding integration of $S_g(t)$ gives the expected length of unemployment given by the model with gamma heterogeneity

$$(20) \qquad E_g(T) = e^{-x\beta}/(1 - \sigma^2).$$

Data on 2077 Finnish unemployed persons has been taken for this study from the register of the Ministry of Labour. 40 % of the observations are right censored, i.e. the completed spells of unemployment were not observed. In order to guarantee that the sample would be randomly generated, every hundredth individual was picked from the flow into unemployment during 1985. The individuals were then followed until the end of their unemployment spells but at most until the end of 1986. The data set is fairly rich as regards individual and market specific information. The description of the variables of the models are in Appendix 1 and reference for further details regarding the data should be made to Kettunen (1989, 1990).

The results of estimations are in Table 1. It is well known that uncontrolled unobservables bias the estimated hazards towards negative duration dependence (Heckman, Singer, 1984, 1986). The constant of the model, where the effect of omitted variables is captured, decreases and the absolute values of nearly all the statistically significant parameter estimates increase when gamma heterogeneity is introduced into the model, as was expected.

Many of the explanatory variables have significant effects on the re-employment probability. Married persons seem to leave unemployment earlier than single persons. Age is a very significant factor. Older people are more likely to have problems in finding jobs. Training for further employment has a significant and positive effect on the re-employment probability. Members of the UI funds, i.e. members of the labour unions in the Finnish system, are often skilled workers and therefore they become employed earlier than the non-members. The persons leaving school or the army have generally not very big problems. They leave unemployment clearly earlier than the others. The persons who have come from house work find it very difficult to find a job. The effects of unemployment benefits are measured using the benefit replacement ratio. The benefits decrease significantly the probability of re-employment as is expected by search theories. The number of children, gender, level of education, demand variables and taxable assets do not have statistically significant effects on the re-employment probability.

The expected lengths of unemployment for an average person in the sample given by the two exponential models without and with gamma heterogeneity are 22.2 and 35.0 weeks respectively. The effects of explanatory variables change substantially when correction for heterogeneity is made. From a policy viewpoint it may be of interest to see whether an increase in unemployment benefits has any influence on the expected duration. The increase of the replacement ratio for example by 0.1 will increase the expected length of unemployment by 3.2 and 6.6 weeks in these two models.

Table 1. Exponential models of unemployment duration

Dependent variable: The length of the spell of unemployment

(A) An exponential model
(B) An exponential model with gamma heterogeneity

Std.errors in parentheses

	(A)	(B)
Variance of heterogeneity		0.540
		(0.066)
Constant	-1.779	-1.213
	(0.105)	(0.181)
Number of children	-0.013	-0.007
	(0.046	(0.068)
Married	0.187	0.148
	(0.059)	(0.088)
Sex	-0.009	-0.041
	(0.051)	(0.078)
Age	-0.046	-0.050
	(0.003)	(0.004)
Level of education	0.066	0.052
	(0.051)	(0.082)
Training for employment	0.180	0.253
	(0.065)	(0.101)
Member of UI fund	0.204	0.288
	(0.054)	(0.082)
Came from schooling	0.326	0.313
	(0.069)	(0.111)
Came from house work	-0.762	-0.801
	(0.115)	(0.154)
Geographical demand	0.278	0.201
	(0.218)	(0.298)
Occupational demand	0.768	0.238
	(0.544)	(0.828)
Taxable assets	1.200	0.838
	(1.002)	(1.290)
Replacement ratio	-1.337	-1.723
	(0.137)	(0.193)
Mean $x\beta$	-3.102	-2.779
Log likelihood	-4987.7	-4929.2

3. A programme for estimating nonlinear maximum likelihood models

3.1. General features of the programme

This section deals with the issues that arise in estimating nonlinear structural models using the method of maximum likelihood. The practical problem of estimating the models with a large number of parameters and likelihood function specified by the researcher is solved in this section. The data handling and estimation of this kind of models using micro computers are awfully slow. Mainframe computers can speed the computation, but the estimation of user specified models using standard statistical packages may be difficult or even impossible. Writing programs using low-level languages like Fortran and using libraries of subroutines may be tedious because programming every detail is a waste of time and the subroutines may not allow for changes by the user. Especially this work is motivated by the estimation and test development of duration models, where the low-level programs may be very long as can be seen in Kalbfleish and Prentice (1980) or Lee (1980). In that area there is clearly a need for a flexible and powerful high-level programme, which allows the user to specify the likelihood function and which provides an environment for the development of specification tests. The programme is written using the SAS/IML matrix language.

The version of the programme presented in Appendix 2 estimates the unknown parameters of an exponential model allowing for gamma heterogeneity. It reads an ASCII data file including the duration t, censoring indicator c and matrix of explanatory variables x and saves it to a SAS data set. The user can control the programme using a set of requirements for the iteration and solution in the beginning of the SAS/IML programme. The set of explanatory variables can be changed using an indicator. The limits for the maximum number of iterations and linear search can be given as well as the criteria for the convergence. The requirement of the precision of solution can be controlled by the user using the accuracy requirements and the proportion of step length.

The estimates of structural parameters of an exponential model can be used as starting values for those parameters of the corresponding model with gamma heterogeneity and the starting value for σ^2 can be randomly allocated. A safeguard against the possibility of convergence to a local maximum that is not a global maximum is to choose several initial values of the parameters. If the iterations do not converge to the same solution, the shape of the log likelihood function should be investigated with care until the global maximum is located.

During the iterations the iteration monitor prints the number of iteration, values of the likelihood function, parameters and gradients. In the case of linear search the number of search, the value of likelihood function, the step adjustment and parameters are provided. However, it is suggested to use the options to suppress the printings of the iteration monitor and linear search during the iterations if they are not needed. If the accuracy requirement is achieved, the programme prints to a listing file e.g. the starting values of parameters, the log likelihood function, the parameter values and their standard errors at solution and the number of iterations. More printings can be easily added.

There are two links which are written at the end of the programme but which are called and executed during the iterations. The likelihood function can be changed by the user. It is written in LINK LIKEF. The algorithm requires the evaluation of first partial derivatives of the log likelihood function with respect to the parameters to be estimated. They are written in LINK LIKED. The Berndt, Hall, Hall and Hausman (BHHH) (1974) algorithm is used to estimate the unknown parameters, but it can easily replaced by an other algorithm.

3.2. The algorithm

The convergence of the BHHH algorithm is guaranteed by the theory unlike the method of scoring and some other statistical maximisation procedures. The ideal is to reach the values of parameters $\theta = (\sigma^2, \beta)$ such that the gradient $L_\theta = 0$. The likelihood equations in a case of an exponential model allowing for gamma heterogeneity can be written

(22) $$L_{\sigma^2} = \Sigma\ [\log K/\sigma^2 - (c + 1/\sigma^2)\log K/\sigma^2 K] = 0$$

(23) $$L_\beta = \Sigma\ \{[c - (c + 1/\sigma^2)(K - 1)/K]\#x\} = 0,$$

where $K = 1 + \sigma^2 te^{x\beta}$ and # indicates elementwise multiplication. Iterations will move uphill along the likelihood function. Each iteration consists in computing the log likelihood function $L(\theta)$ and the gradient L_θ, which is used to derive a direction of increase of $L(\theta)$. According to the classical Gradient Theorem, which is proved e.g. in Jacoby, Kowalik and Pizzo (1972), any vector d with $L_\theta d > 0$ is a direction of increase of $L(\theta)$ in the sense that $L(\theta + \alpha d)$ is an increasing function of the step length α for small enough values of α. The directions d can be derived from the gradient by multiplying it by a positive definite matrix Q. The convergence is speeded by a choice of Q such that it is the inverse of the Hessian matrix of second derivatives of $L(\theta)$. The use of the information matrix identity $- E(L_{\theta\theta}) = E(L_\theta)E(L_\theta)'$ avoids the need of analytical or numerical calculation of second derivatives. The updated estimates of parameters calculated during the i^{th} iteration can be written

$$(24) \qquad \theta^{i+1} = \theta^i + \alpha^i L_\theta^{i\prime}(L_\theta^i L_\theta^{i\prime})^{-1}.$$

If α is too large leading to a decreasing value of $L(\theta)$ the linear search subiteration is used. There are plenty of methods of linear search, as noted by Quandt (1986). As a simple method to avoid computation during the iterations it is suggested that the step length α is halved as many times as needed to find an increasing value of $L(\theta)$ and then new values of θ^{i+1} are calculated.

There are many criteria for the stopping rule of iterations [see e.g. Quandt (1986)]. In the neighbourhood of a maximum the algorithm takes small steps in the sense that $|L(\theta)^{i+1} - L(\theta)^{i}|$ is small. The ideal of reaching values of θ such that $L_\theta = 0$ is not attainable in practice. In the neighbourhood of a maximum $(\Sigma\ L_\theta{}^2)^{1/2}$ is likely to be small. Both of these conditions are used as a stopping criteria.

4. Conclusions

An exponential model of unemployment duration was estimated using Finnish microeconomic data collected from various registers. Even though the data is rich in explanatory variables and more reliable than the data from surveys, there is reason to assume that relevant variables have been omitted from the model. Neglected heterogeneity across individuals was taken into account in estimation. An exponential model allowing for gamma heterogeneity was estimated assuming that the effect of omitted variables has a gamma distribution across individuals. Comparing the results of the two models shows that the model without correcting for heterogeneity gives lower estimates of parameters. The absolute values of parameter estimates and effects of explanatory variables increase when heterogeneity is introduced into the model.

Many of the explanatory variables have significant effects on the re-employment probability. Married persons, members of labour unions, school graduates and persons with training for employment leave unemployment earlier than other persons. Old people, persons who have come from house work and those who get high unemployment benefits have longer unemployment spells than the others.

The estimation of the microeconomic models with a large number of parameters and likelihood function specified by the researcher may be difficult in practice using micro computers or standard statistical packages. The needed statistical tool with an application was presented in this paper. It is written using the SAS/IML matrix language, which is close to matrix algebra notation. The use of this programme is not limited to this particular application, but with slight modifications it is a useful tool to estimate a wide class of maximum likelihood models. Furthermore, the programme is a basis for further development and an environment for developing specification tests according to the needs of the users.

References

Berndt, E.R, Hall, B.H., Hall, R.E. and Hausman J.A. (1974), "Estimation and Inference in Nonlinear Structural Models, Annals of Economic and Social Measurement", 3, 653-666.

Cox, D.R. and Snell, E.J. (1968), "A General Definition of Residuals", Journal of the Royal Statistical Society, Series B, 30, 248-275.

Engström, Lars and Löfgren, Karl-Gustaf (1987), "Disguised and Open Unemployment Intensified Employment Service and Unemployment Durations", Trade Union Institute for Economic Research, Working Paper Series, No. 39, Stockholm.

Heckman, James J. and Singer, Burton (1984), "Econometric Duration Analysis", Journal of Econometrics, 24, 63-132.

Heckman, James J. and Singer, Burton (1986), "Econometric Analysis of Longitudinal Data", in Griliches, Z and Intriligator, M.D. (eds.), Handbook of Econometrics, Vol. III, North-Holland Publishing Company, Amsterdam.

Jacoby, S.L.S., Kowalik, J.S. and Pizzo, J.T. (1972), "Iterative Methods for Nonlinear Optimization Problems", Prentice-Hall, Englewood Cliffs, New Jersey.

Kalbfleisch, J.D. and Prentice, R.L. (1980), "The Statistical Analysis of Failure Time Data", John Wiley and Sons, New York.

Kettunen, Juha (1989), "Työttömyysturvan vaikutukset työnetsintään", The Research Institute of the Finnish Economy, C 49, Helsinki.

Kettunen, Juha (1990), "Työllistyminen, työvoiman liikkuvuus ja työttömän taloudellinen asema", The Research Institute of the Finnish Economy, B 67, Helsinki.

Kooreman, P. and Ridder G. (1983), "The Effects of Age and Unemployment Percentage on the Duration of Unemployment", European Economic Review, 20, 41-57.

Lancaster, Tony (1979), "Econometric Methods for the Duration of Unemployment", Econometrica, 47, 939-956.

Lee, Elisa T. (1980), "Statistical Methods For Survival Data Analysis", Lifetime Learning Publications, Belmont, California.

Narendranathan, W., Nickell, S. and Stern, J. (1985), "Unemployment Benefits Revisited", The Economic Journal, 95, 307-329.

Newman, John L., McCulloch, Charles E. (1984), "A Hazard Rate Approach to the Timing of Births", Econometrica, 52, 939-961.

Nickell, Stephen (1979), "Estimating the Probability of Leaving Unemployment", Econometrica, 47, 1249-1266.

Quandt, R.E. (1986), "Computational Problems and Methods", in Griliches, Z., Intriligator, Michael M.D. (eds.), Handbook of Econometrics, Vol. I, North-Holland Publishing Company, Amsderdam.

SAS/IML User's Guide, Version 5 Edition (1985), SAS Institute Inc., Cary, North Carolina.

Appendix 1. Variables of the data

Duration of unemployment is calculated in weeks and it is the difference between the date of entry into unemployment and the date of returning back to work. Mean = 15.03.

Number of children is the number of unemployed person's children who are younger than 18 years old. Mean = 0.23.

Married is a dummy variable, 1=yes. Mean = 0.37.

Sex is a dummy variable, 1=male. Mean = 0.54.

Age is measured in years. Mean = 31.2.

Level of education is a dummy variable, 1 = at least 12 years of education. The level of education is based on the education code of the Central Statistical Office of Finland. Mean = 0.45.

Training for employment is a dummy variable, 1 = The person has got training for further employment. Mean = 0.15.

Member of UI fund is a dummy variable, 1 = yes. Mean = 0.42.

Came from schooling is a dummy variable, 1 = The person has come from schooling or from the army. Mean = 0.13.

Came from house work is a dummy variable, 1 = The person has come from house work or elsewhere outside the labour force. Mean = 0.07.

Geographical demand describes the geographical rate of jobs available. It is the number of vacancies divided by the number of job seekers in the area. Mean = 0.10.

Occupational demand describes the occupational rate of jobs available in the whole country. It is the number of vacancies divided by the number of job seekers in the occupation group. Mean = 0.12.

Taxable assets has been compiled from the tax register and it is measured in millions of marks. Mean = 0.011.

Replacement ratio is unemployed persons average replacement ratio of unemployment benefits during the unemployment period after tax. Weekly unemployment benefits after tax have been divided by the weekly income after tax.
Mean = 0.17.

Appendix 2. A SAS/IML programme for estimating an exponential duration model with gamma heterogeneity

```
*SAS/IML PROGRAMME FOR ESTIMATING AN EXPONENTIAL MODEL WITH GAMMA
HETEROGENEITY IN THE VMS OPERATING SYSTEM;
OPTIONS LS=80 PS=500;
LIBNAME SASLIBR '[JKETTUNEN.SASFILES]';
FILENAME RAWDATA 'GL.DAT';
DATA SASLIBR.ADATA;
  INFILE RAWDATA;
  INPUT T 1-7 .3 C 9 CONST 11 KIDS 13 MARRIED 15 SEX01 17 AGE 19-20
        EDU38 22 UEDU01 24 MEMBER01 26 CAME1 28 CAME2 30
        REGDEM 32-35 .3 PROFDEM 37-40 .3 ASSU 42-46 .4 RUN 48-54 .6;
PROC IML;
START;
*POSITION OF T DATA;        IND1=1;
*POSITION OF C DATA;        IND2=2;
*POSITION OF X DATA;        IND3={3 4 5 6 7 8 9 10 11 12 13 14 15 16};
*LIMIT OF ITERATION;        LIT =100;
*LIMIT OF LINE SEARCH;      LLS =100;
*ACCURACY OF REL FUN;       ACC1=0.001;
*ACCURACY OF GRADIENT;      ACC2=0.01;
*STEP ADJUSTMNT;            ALPHA=1.0;
*SUPRESS LINE SEARCH;       SUPRESLS=1;
*SUPRESS ITERATION;         SUPRESIT=1;
*STARTING VALUES;           B={0.540  -1.213  -0.007   0.148
                              -0.041  -0.050   0.052   0.253
                               0.288   0.313  -0.801   0.201
                               0.238   0.838  -1.723};
*NAMES OF PRINTINGS;
 NAMES1={'SEARCH NO' ' OLD L' 'NEW L' 'ALPHA'};
 NAMES2={'OLD B' 'NEW B'};
 NAMES3={'ITER NO' L};
 NAMES4={PARAMETR GRADIENT};
 NAMES5={PARAMETR 'S.ERROR' 'T STAT' GRADIENT};
 NAMES6={ACC1 ACC2 ALPHA 'NO OBS' 'NO PARS'};
 NAMES7='VARIANCE';
 NAMES8={T C CONST KIDS MARRIED SEX01 AGE EDU38 UEDU01 MEMBER01
         CAME1 CAME2 REGDEM PROFDEM ASSU RUN};
USE SASLIBR.ADATA;
READ ALL INTO A;
T=A(¦,IND1¦); C=A(¦,IND2¦); X=A(¦,IND3¦);
FREE A;
OBS=NROW(T);
PAR=NCOL(X);
PAR2=1+PAR;
VARNAME=NAMES7¦¦NAMES8(¦1,IND3¦);
PRINT 'MAXIMUM LIKELIHOOD ESTIMATION';
PRINT 'WRITTEN BY JUHA KETTUNEN';
PRINT 'ACCURACY REQUIREMENT:STEP LENGTH PROPORTION:NO OF OBS &
PARAMETERS';
PRINT
'**********************************************************************';
PRT=ACC1¦¦ACC2¦¦ALPHA¦¦OBS¦¦PAR2;
```

```
PRINT PRT (¦ COLNAME=NAMES6 ¦);
PRINT 'DEFAULT STARTING VALUES PROVIDED';
PRINT '********************************';
PRINT B;
NIT=1;
BOLD=B;
LINK LIKEF;
LOLD=LL;
MAR4:  LOLD=LL; BOLD=B;
       LINK LIKED;
       IF SUPRESIT=0 THEN DO;
          LOGLIK =NIT¦¦LL;
          PARAMS =B`¦¦DL;
          IF NIT=1 THEN DO;
             PRINT '*****************';
             PRINT 'ITERATION MONITOR';
             PRINT '*****************';
             END;
          PRINT LOGLIK (¦ COLNAME=NAMES3 ¦);
          PRINT PARAMS (¦ COLNAME=NAMES4 ¦);
          FREE LOGLIK PARAMS;
          END;
       NLS=1;
       FAC=DL`*INV(CL);
MAR3:  B=BOLD-ALPHA#FAC;
       IF NLS > 1 & SUPRESLS=0 THEN DO;
          NLS1=NLS-1;
          IF NLS=2 THEN DO;
             PRINT 'LINE SEARCH SUBITERATION';
             PRINT '************************';
             END;
          LOGLIK =NLS1¦¦LOLD¦¦LL¦¦ALPHA;
          PARAMS =BOLD//B;
          PRINT LOGLIK (¦ COLNAME=NAMES1 ¦);
          PRINT 'PARAMETER VALUES';
          PRINT PARAMS (¦ ROWNAME=NAMES2 ¦);
          FREE LOGLIK PARAMS ;
          END;
       LINK LIKEF;
       IF LL > LOLD THEN GOTO MAR1;
       NLS=NLS+1;
       IF NLS < LLS THEN GOTO MAR2;
       PRINT '*****************************';
       PRINT '*SORRY - LINE SEARCH FAILURE*';
       PRINT '*****************************';
       PRINT BOLD B DL CL FAC LOLD LL NLS NIT;
       GOTO MAR7;
MAR2:  ALPHA=ALPHA/2;
       GOTO MAR3;
MAR1:  NIT=NIT+1;
       IF ABS((LL-LOLD)/LOLD) < ACC1*ALPHA
          & SQRT(DL(¦##,¦)) <= ACC2 THEN GOTO LEN;
       IF NIT > LIT THEN GOTO MAR5;
       GOTO MAR4;
MAR5:  PRINT '**************************';
       PRINT '*ITERATION LIMIT EXCEEDED*';
       PRINT '**************************';
       PRINT BOLD B DL CL FAC LOLD LL NLS NIT;
       GOTO MAR7;
```

```
LEN:   PRINT '***********************************************';
       PRINT '*CONGRATULATIONS ACCURACY REQUIREMENT ACHIEVED*';
       PRINT '***********************************************';
MAR7:  LINK LIKED;
       PRINT 'LOG LIKELIHOOD FUNCTION AT SOLUTION';
       PRINT '***********************************';
       PRINT LL;
       VAR=INV(-CL);
       SERR=SQRT(VECDIAG(VAR));
       TSTAT=B`/SERR;
       SOLN =B`¦¦SERR¦¦TSTAT¦¦DL;
       PRINT 'PARAMETER VALUES AT SOLUTION';
       PRINT '****************************';
       PRINT SOLN (¦ COLNAME=NAMES5 ROWNAME=VARNAME ¦);
       PRINT 'NUMBER OF ITERATIONS';
       PRINT '********************';
       PRINT NIT;
       MEANXB=XB(¦+,¦)/OBS; PRINT MEANXB;
       STOP;
LIKEF: *THE LIKELIHOOD FUNCTION;
       SIGMA=B(¦1,1¦); BETAS=B(¦1,2:PAR2¦);
       XB=X*BETAS`; K=1+SIGMA#T#EXP(XB);
       H=EXP(XB)/K;
       IH=1/SIGMA#LOG(K);
       L=C#LOG(H)-IH;
       LL=L(¦+,¦);
RETURN;
LIKED: *THE FIRST DERIVATIVES;
       V1=LOG(K)/SIGMA/SIGMA-(C+1/SIGMA)#(K-1)/SIGMA/K;
       V2=(C-(C+1/SIGMA)#(K-1)/K)#X;
       V=V1`//V2`;
       DL=V(¦,+¦);
       CL=-V*V`;
RETURN;
FINISH;
RUN;
```

ELINKEINOELÄMÄN TUTKIMUSLAITOS (ETLA)
THE RESEARCH INSTITUTE OF THE FINNISH ECONOMY
LÖNNROTINKATU 4 B, SF-00120 HELSINKI

Puh./Tel. (90) 609 900 Telefax (90) 601 753
Int. 358-0-609 900 Int. 358-0-601 753

KESKUSTELUAIHEITA - DISCUSSION PAPERS ISSN 0781-6847

No 359 JUHA KETTUNEN, Occupational Mobility of Unemployed Workers. 25.03.1991. 29 p.

No 360 JUSSI RAUMOLIN, Logistiikan näkymiä yhdentyvässä Euroopassa. 25.03.1991. 17 s.

No 361 MARKKU OLLIKAINEN, Kestävä kehitys - ongelmia ja tulkintoja. 08.04.1991. 24 s.

No 362 PEKKA ILMAKUNNAS, Working Time, Productivity and Labor Demand in Finnish Manufacturing. 22.04.1991. 40 p.

No 363 JUHA KETTUNEN, Time-Dependent Effects of Unemployment Benefits. 24.04.1991. 29 p.

No 364 GEORGE F. RAY, Long-Term Prospects for Industrial Materials. 30.04.1991. 45 p.

No 365 ROBERT HAGFORS - TOIVO KUUS, The Structure and Distribution of Income in Estonia and Finland. 07.05.1991. 30 p.

No 366 T.R.G. BINGHAM, The Foreign Exchange Market: Structure, Intervention and Liquidity. 10.05.1991. 29 p.

No 367 Elinkeinoelämän Tutkimuslaitoksen toiminta vuonna 1990. 24.05.1991. 30 s.

No 368 JUHA KETTUNEN, The Effects of Education on the Duration of Unemployment. 06.06.1991. 38 p.

No 369 ROBERT HAGFORS - ROLF MAURY, Suomalaisten kotitalouksien rakennekehityksen arviointia vuosille 1960-1985. 18.06.1991. 47 s.

No 370 VESA KANNIAINEN - JAN SÖDERSTEN, Undervaluation of Corporate Shares Revisited: A Note. 27.06.1991. 9 p.

No 371 MARKKU RAHIALA - TIMO TERÄSVIRTA, Forecasting the Output of Finnish Forest Industries Using Business Survey Data. 29.07.1991. 13 p.

No 372 MARKKU KOTILAINEN, Exchange Rate Unions: A Comparison to Currency Basket and Floating Rate Regimes - A Three-Country Model with Endogenous Prices. 29.07.1991. 47 p.

No 373 EIKE HINDOV, On Population Development in Estonia and Finland. 20.08.1991. 47 p.

No 374 JUHA KETTUNEN, A Search Theoretical Analysis of the Finnish Unemployment Insurance System. 28.08.1991. 25 p.

No 375 JUHA KETTUNEN, Regional Mobility of Unemployed Workers. 28.08.1991. 19 p.

No 376 JUHA KETTUNEN, Työttömyysturvajärjestelmän vaikutukset työnetsintään: Tutkimusaineistokuvaus. 28.08.1991. 13 s.

No 377 WIDGRÉN MIKA, Voting Power in the EC Decision Making and the Consequences of two Different Enlargements. 04.09.1991. 24 p.

No 378 JUKKA LASSILA, Income Tax Indexation in an Open Economy. 11.09.1991. 17 p.

No 379 KARI ALHO - KATARIINA LEIKAS, Arvio kotitalouksien varallisuudesta 1960-1990. 14.10.1991. 28 s.

No 380 TIMO TIAINEN, Viennin markkinaosuudet ja kilpailukyky Suomen teollisuuden eri toimialoilla 1970-1985. 21.10.1991. 148 s.

No 381 JUSSI RAUMOLIN, Prospects for Logistics in Europe in the 1990s. 21.10.1991. 24 p.

No 382 SYNNÖVE VUORI, Teknologiapanosten vaikutukset ja teknologian diffuusio: Kirjallisuuskatsaus. 21.10.1991. 29 s.

No 383 PEKKA VALKONEN - RAIMO LOVIO: Suomalaiset patentoijat Yhdysvalloissa 1980-luvulla: Menestyksen suhteellisuus. 21.10.1991. 19 s.

No 384 ROBERT HAGFORS, Kulutusmenojen arviointia kotitaloustyypeittäin Suomessa vuosille 1960-1985. 04.11.1991. 30 s.

No 385 GEORGE F. RAY, Innovations Diffused - A Random Walk in History. 15.11.1991. 46 p

No 386 LUUKKANEN, Harri, Diffusion of Numerically Controlled Machine Tools in Finnish Engineering Industry. 18.11.1991 37 p.

No 387 KETTUNEN, Juha, A Programme for Estimating Nonlinear Maximum Likelihood Models with an Application to Duration Models of Unemployment. 27.11.1991. 25 p.

Elinkeinoelämän Tutkimuslaitoksen julkaisemat "Keskusteluaiheet" ovat raportteja alustavista tutkimustuloksista ja väliraportteja tekeillä olevista tutkimuksista. Tässä sarjassa julkaistuja monisteita on rajoitetusti saatavissa ETLAn kirjastosta tai ao. tutkijalta.
Papers in this series are reports on preliminary research results and on studies in progress; they can be obtained, on request, by the author's permission.

E:\sekal\DPjulk.chp/27.11.1991