

The Computer Implementation of Econometric Model

Haitao Sun

School of Management Bohai University

Jinzhou, China

womenhao700@sohu.com

Abstract—Based on the application of software like MATLAB and Excel etc., the article gives focus on the detailed implementation of economic model. The article introduces the basic knowledge of econometric model and model building method, elaborates on the application of econometric model in survival model, and gives example on how to estimate undetermined parameters with a complete set of data.

Keywords—Economic model, Survival model, Computer application

I. INTRODUCTION

All manuscripts must be in English. These guidelines include complete descriptions of the fonts, spacing, and related information for producing your proceedings manuscripts. Please follow them and if you have any questions, direct them to the production editor in charge of your proceedings at Conference Publishing Services (CPS): Phone +1 (714) 821-8380 or Fax +1 (714) 761-1784.

Mathematics is a science to describe the movement of the objective world in the form of abstract mathematical relations. Before the middle of the 20th century, mathematics mainly served for natural science and productive technology. About 50 to 60 years ago, economic mathematics emerged to meet the requirement of economic science. Economic mathematics is a discipline to analyze economic affairs using mathematical method and to make special research on the mathematical problems in economics, which enhances analysis and research on quantity relation [1-3].

Model is a method of analysis which describes the actual world very concisely. The actual world is very complicated with various main variables and subordinate variables. Therefore, the subordinate elements need to be excluded; otherwise strict analysis will not be possible or even cannot be carried out [4]. To make hypothesis can exclude many subordinate elements and establish model. Then the model can be used for analysis on the special occasion set by the hypothesis.

An econometric model consists of one or more equations. It concisely and effectively describes and summarizes the quantitative features of an actual economic system, and reveals the quantitative change rule of this economic system. It is composed of systems or equations, and equations are composed of variables and factors, in which factors are also composed of equations. The econometric model reveals the quantitative relation between factors of the economic activity and describes it in random mathematical equations. In the

wide sense, it refers to all the models containing economics, mathematics and statistics. In the narrow sense, it means the models using mathematical statistical method like parameter estimation and hypothesis examination to analyze experience data [5]. An important feature of econometric model is that it is based on statistical data, i.e., the econometric model cannot be established without statistical data. Econometrics is an influencing discipline in the field of economics and has formed relatively independent discipline system with rich content [6-7].

The article discusses on economic model from the aspect of econometric model and gives example of its application and computer implementation. The software for the values in this article includes MATLAB7.0 and Excel2007.

II. ESTABLISHMENT OF ECONOMETRIC THEORETIC MODEL

First select the factors that shall be contained in the model as per purpose of research on the basis of deep analysis on the economic phenomenon. Choose the proper variables to represent these factors and set up mathematical expression $f(t; a_1, a_2, \dots, a_m)$ (in which a_1, a_2, \dots, a_m are parameters), i.e. theoretical models to describe the relation between the factors as per the relation between the variables displayed by economic behavior theory and sample data. The design of the theoretical model mainly consists of 3 parts, selection of variables, determination of the mathematical relation between the variables and formulation of the range of value of the undetermined parameters in the model.

To determine the variables contained in the model: In a single equation model, there are two types of variables: the variables as research object, i.e. the result in cause-result relation, for example, the yield in production yield is the explained variable in the model; and the variable as the cause, i.e. the cost, labor and technology are the explaining variable in the model. To determine the variables in the model mainly refers to determine the explaining variables. The following types of variables can be used as explaining variables: exogenous economic variable, exogenous condition variable, exogenous policy variable and lagged explained variable.

To be strict, the yield, cost, labor and technology etc. in the production function can only be called as factors. There is cause and result relation between these factors. To set up econometric model, we must choose the proper variables to represent these factors. These variables must have data accessibility. Therefore, we can use the total output to represent the yield, the fixed asset original value to represent cost, the number of employees to represent labor and time to

represent technology. Thus the model finally set up is the mathematical expression about the relation between total output, fixed asset original value, number of employees and time variables.

The key point is how to choose the explaining variables correctly after determination of the explained variables.

First, we need to understand the economic theory and economic behavior rule hidden in the economic phenomenon being analyzed. This is the basis for correct selection of variables. For example, as production is concerned, when the problem of shortage of supply is clearly indicated, the factors influencing on the yield must be the input factors. Generally, the input factors mainly include technology, asset and labor. In case of lack of demand, the factors influencing on the yield must be in the aspect of demand instead of input factors. Then if the research object is production of consumers' product, variables like income of residents should be selected as the explaining variables. If the research object is production of production material, variables like fixed asset and total investment and so on should be chosen as the explaining variables. This indicates that the selection of variables is different depending on the economic environment and industry.

Secondly, data availability shall be considered for the selection of variables. This demands thorough understanding of economic statistics. The model of econometrics is to estimate the parameters using certain mathematical method with the support of sample data, i.e. sample observation value of the variables to reveal the quantitative relation between variables. Therefore, the variables selected must be existing in the statistical index system with reliable data source.

Thirdly, the relation between all the selected variables shall be considered for the selection of variables to make each explaining variable independent. This is demanded by the econometric model technology. Of course, this is difficult at the beginning. If there are related variables in the selected variables, it can be checked out and deleted during the establishment of model.

Determination of the mathematical form of model: After selecting the proper variables, we need to select the proper mathematical form to describe relation between these variables, that is, to set up the theoretical model.

The main principle for selecting the mathematical form is economic behavior theory. In mathematical economics, there has been a large amount of research on the mathematical form for models of production function, demand function, consumption function and investment function etc. The research output can be referred to. Here we must notice that modern economics pays much attention to empirical study. Any theoretical model on the basis of certain economic theoretical hypothesis cannot be acceptable if it fails to explain the past well, especially the historic statistic data. Thus the establishment of theoretical model demands repetitive revision during the course of parameter estimation and model examination to obtain a mathematical model that can have better economic explanation and better reflect the relation between the historic variables. It is not correct to ignore any aspect. This also indicates that econometrics is to

some extent dependent on mathematical economic model. We can also make a scatter diagram to show the relation between the explaining variables and explained variables as per the sample data of the variables, and use the function relation between variables shown by the scatter diagram as the mathematical form of the theoretical model. This method is commonly used for establishment of model.

Application of econometrical model-survival model

Survival model: We observed the survival status of a group of white mice injected with a kind of carcinogen. The injection time $t=0$. The living time of these mice after injection is observed. Obviously T is a random variable. The difference from general random variables is that here we are more interested in survival function instead of distribution function of T .

$$S(t)=1-F(t)=P(T>t)$$

It is the survival rate of the white mice at any moment t .

Obviously,

$$(a) T>0$$

$$(b) S(0)=1$$

$$(c) S(t) \text{ is the decreasing function of } t \text{ and } \lim_{t \rightarrow \infty} S(t)=0$$

The so called survival model is the model to study survival function $s(t)$.

There are other similar survival models: to examine the reliability of mechanical equipment and so on. Generally, $S(t)$ is the function only about t and has relation with the survival time of the object before $t=0$. For example, for a machine with an average campaign life of 3 years, the survival function is definitely different to be observed after 1 year and after 2 years. However, in some cases, there is survival function that only relates to t . For example, as to the patients diagnosed with a fatal disease, the survival distribution only depends on time t to be observed from the time the diagnosis is confirmed.

Type of survival model: A classification is to divide survival model into parameter survival model and non-parameter survival model. When the survival function $S(t)$ is represented by a mathematical function of t , it is called parameter survival model. For example, if the random variable T abides by index fraction number, then

$$S(t)=e^{-\lambda t}$$

Here, λ is the parameter. Generally, single parameter model (e.g. even distribution or index distribution) cannot describe well the precise survival model. Usually double parameter Gomperts distribution or Weibull distribution or even tri-parameter Makeham distribution is used for better result.

The survival model can be classified as table survival model and non-table survival model as per the way the survival model is described. However, the table can only give the value of finite discrete $S(x)$, e.g. the life table gives only the value when $x=0, 1, 2$, and when x is not an integer, we can get the value of $S(x)$. Therefore, in the life table, we make several hypotheses on the end age, e.g. linear hypothesis, constant value dead force hypothesis and Balducci hypothesis etc. When these hypotheses are applied to the whole model, $S(x)$ value can be obtained. Survival model is classified as survival model with accompany

variables and survival model without accompany variables as per whether there are accompany variables or not. We often attach certain hypothesis to the survival model with accompany variable, and there are often undetermined parameters in these hypotheses. Therefore, such survival model is also called as the parameter model with accompany variables [8].

As per the property of the sample data, we classify survival model as survival model with complete sample data and survival model without complete sample data. The so called complete sample data means that the observer observes the object in each small unit with enough time until the object dies and the object is not allowed to leave till it dies. Usually the complete sample data is difficult to get and can only be obtained in laboratory or in other strictly controlled conditions. The observation of each object starts from a certain moment (no observer is added in the middle of the observation) until all the objects die and no object leaves the observation alive. In practice, there are not many such sample data. Generally during certain observation period, the observation of an object sometimes starts from the beginning and sometimes from the middle. And some leave the observation alive, some live till the end of the observation and some die during the observation period. Such sample data is called incomplete data.

Research method: Generally, there are mainly two methods. The research on relatively large sample space is called as horizontal research. During the research, we first determine the group for research. This group can be population of a city or a country or the insurance policy holder of a life insurance company or the members of a pension scheme. And we also need to select an observation period. At the beginning of the period, many observation objects are already members of the research group and receive observation from the beginning of the period. Some other objects join the group of research at any moment during the observation period and some objects leave the research during the observation period. The behavior of joining and leaving the observation is called migration. The death probability of each age group can be estimated by classification and data treatment, especially the classification of the number of deaths as per designated program. The probability value constitutes a transitive table survival model.

Contrary to the horizontal study of large sample, most clinical studies first select a group as research object and carries out research on each of the member of the group till the death. Such research is called vertical research. This method is usually adopted when the time of living is short. Apparently, the observer must have the control ability to make the objects stay with the research until the death for this type of research. Therefore, such research is called controllable data research sometimes.

In some circumstances, vertical research can terminate before all the objects die to avoid excessive study cycle. In this case, the corresponding data is incomplete data. Generally there are two ways to terminate research. One way is to terminate the study before a reserved date, where data ends on this date. The other way is to terminate the study

when the estimated number of deaths is reached, where data is cut off.

III. COMMENTS ON ECONOMETRIC MODEL

The 3 important elements for setting up econometric model are theory, method and data.

a) Theory, i.e. economic theory and behavior theory of the economic phenomenon being studies. Theory is the basis for econometric study.

b) Method, mainly including model method and calculation method. Method is the tool and means for econometric study and is the prominent difference of econometric economy from other economic branch disciplines.

c) Data, to reflect the activity, relation in between and external environment. In the wide sense, data is information and is the raw material for econometric study. These 3 elements are indispensable for econometric study.

Usually, the research on method is the focus of the researchers in econometric study. The level of method normally becomes the measurement of a research output. This is quite normal. The research on the method for econometrics is the obligation of the researchers. However, the discussion and exploration on the theory of economics also cannot be ignored. A person without the knowledge of economic theory and economic behavior is not qualified for econometric study and cannot set up an even simplest econometric model. Therefore, an econometric specialist first should be an economist. People obviously have not attached enough importance to the study of data quality. Applying for a research project or reviewing a research output, people lack consideration and speculation on the availability, accessibility and reliability. In case of problems during the research, people seldom look for the reason from the aspect of data quality. And the actual situation is that data has become an important issue limiting the development of econometrics.

IV. CONCLUSION

The article classified economic model as econometric model, optimum economic model, input and output economic model and mathematical economic model from the aspect of applied method for discussion, and gives example of application and computer implementation. The software used for the values in the article includes MATLAB7.0 and Excel 2007.

REFERENCES

- [1] Batten, R.W., Mortality Table Construction. Englewood Cliffs: Prentice-Hall, Inc,1978
- [2] Bowers J Actuarial Mathematics[M]. Chicago, Illinois: The society of Acturiers, 1994,1:1-6
- [3] Alan B Krueger, Bruce D Meyer, Labor Supply Effects of Social Insurance, NBER Working PpPers, No.9014
- [4] Mata J, Portugal P, Life duration of new firms[J]. The Journal of Industrial Economics, 1994, 42(3): 227-245.
- [5] Allen L N, Rose L C, Financial survival analysis of defaulted debtors[J]. Journal of the Operational Research Society, 2006,57(6): 630-636.

- [6] Luoma M, Laitinen E K. Survival analysis as a tool for company failure Prediction[J]. Omega International Journal of Management Science,1991,19(6): 673-678.
- [7] Lee E T. Statistical method for survival data analysis[M]. New York: John Wiley&Sons Inc. 1992.
- [8] F Black M Scholes, The Pricing of Options and Corporate Liabilites, Journal of Political Econmics, 1973