Developing the Microsimulation Service System

Josef Csicsman - Cecília Fényes
Department of Information and Knowledge Management
Budapest University of Technology and Economics
H-1521 Budapest, Hungary

HU ISSN 1418-7108: HEJ Manuscript no.: INF-030530-C

Abstract

Microsimulation models are used for forecasting, policy advice and development of economic theory. This paper begins with a short description of microsimulation, then we summarise the concepts and results of the development.

1 Introduction

In the past few years the Hungarian Central Statistical Office (KSH), the Ministry of Economic Affairs (GM), the Ministry of Finance (PM) and other departments of the government have had no opportunity to analyse large data sets to estimate costs of proposed legislation.

The available analytical systems - usually based on EXCEL - are, in methodological respect, questionable, because it is difficult to accept the reports based on small data sets both in mathematical and in economical consideration.

The research group was founded jointly by the Research Centre for Financial Economics (BME[1]) and Új Calculus Bt. The aim of the project is to develop and maintain a model system which makes it possible for economists to analyse data collections and databases available in the administration. Naturally, this methodology can also be used for processing other types of data.

In accordance with international applications the group creates technical and methodological conditions for analysing large data sets and developing a SAS Software based microsimulation modelling system.

2 The term of microsimulation

Microsimulation is well known as a tool for analysis of tax and transfer policies, and for the generation of cost estimates for proposed legislation. There are also several examples of efforts to develop models of socio-economic and demographic outcomes in multiple domains, such as births, deaths, marriages and divorces, education, labour force behaviour, incomes and savings.

Microsimulation can be described as some type of algorithm for data analyses, for which the unit of observation is micro. It means that microsimulation models operate at the level of the individual behavioural entity, such as a person, family, or firm. Such models simulate large representative populations of these low-level entities in order to draw conclusions that apply to higher levels of aggregation such as an entire country.

All the elements of the population must be changed according to the conditions of the model. The generated data will look like the original data, and can, therefore, be analysed and summarised just like the original data. The specification of the model must precede the microsimulation procedure, and

[1] Budapest University of Technology and Economics

parameter values must be obtained either by statistical estimation or other means (including assumption, borrowing from other sources, laws, rules or empirical facts).

It is also very important to mention that we can only expect an exact, reliable result if the observation has appropriate size, because the procedure is based on the Monte Carlo Method.

We have to draw a distinction between "statistical" models and "simulation" models. The former consists of one or more mathematical expressions that include parameters, the numerical values of which are obtained through estimation based on empirical data. Assessment of a statistical model can depend, in part, on a comparison of the estimated model's predictions with their real-world counterparts. In contrast, a simulation model may take the form of a computer program, and the output of the model might consist of artificial data.

2.1 What is microsimulation used for?

- Correcting survey errors

  Surveys often contain errors, and may not be considered representative. The reason for this can be that the sample is not representative (people having a modest income fill out questionnaires more helpfully), or the question is a sensitive one (for example it is related to alcohol consumption or smoking), so the answers may not correspond to the facts.

- Bringing data of former surveys up-to-date

  Making a survey is very expensive (in particular in case of a large sample). Microsimulation models can often substitute repeated surveys (naturally, accurate parameter values are required).

- Generation of cost estimates for proposed legislation

  As mentioned above, microsimulation is suitable for modelling the decisions of economic and social policy. Microsimulation models are used in various departments of the Canadian federal government, where the policy analysis generates inputs to the government decision-making process.

  European tax-benefit microsimulation model (EUROMOD), that provides quantitative estimates for Germany, the Netherlands and the UK, or the Danish Micro Simulation Tax model can also serve as an example.

2.2 Why is a Microsimulation Service System needed?

We had two alternative possibilities of building the Microsimulation Service System: offer a universal modelling system, or implement a software to answer special questions.

The advantage of a special software is that it is developed definitely to solve a concrete problem, so it can be a very efficient solution. The disadvantage is that a new development has to be started, if a new problem arises.

Building a universal microsimulation system requires a great deal of effort, but we decided to give priority to flexibility rather then to the expense of programming time.

3 The Microsimulation Service System

As a result of this project a Microsimulation Modelling System is developed which is suitable for modelling the decisions of economic and social policy, and - according to the national and international requirements - well-founded analyses can precede the policy proposals of the government.
The main tasks of the system:

- Meta dictionary;
- Data protection;
- Micromodules;

- Estimation algorithms and parameter charts;
- Data analysing.

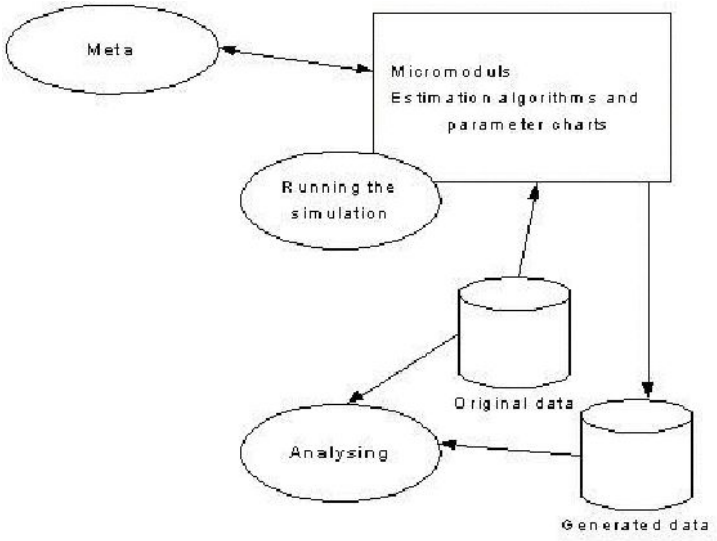

Figure 1:

The system has to serve the needs of the data analyst that will build the micro modules, tune the parameters of estimations and analyse input and output data.

3.1 The language of the system

In 2001 a Microsimulation Conference was organized at the Budapest University of Technology and Economics (V. Conference on Financial Informatics, 15-16 October 2001) which proved that the international cooperation is very important. For this reason the language of the program is adjustable. At the moment English and Hungarian versions are available, but because of the system works from a dictionary and the compilation doesn't demand any special knowledge (only the command of the language), it is very easy to translate it into any other language.

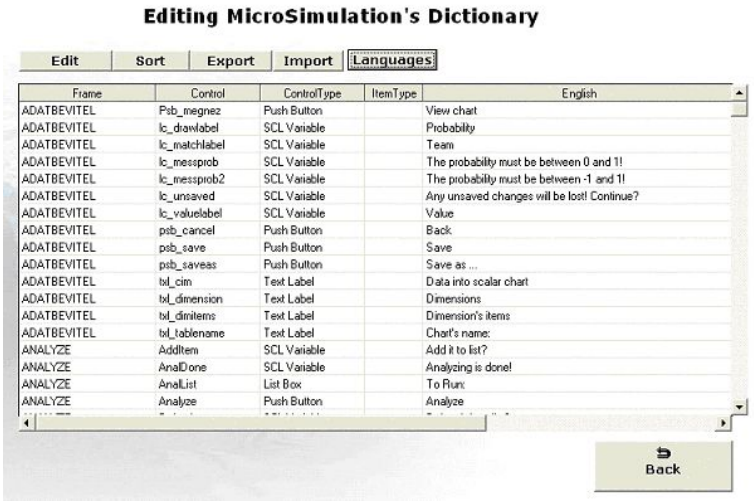

Figure 2:

3.2 Meta dictionary

The Meta dictionary contains all the information about the data: identifier, type, length and name of nomenclatures and pointers. The Meta dictionary makes data handling easier, because in this way data is easy to use both for analyser and for computer programmers.

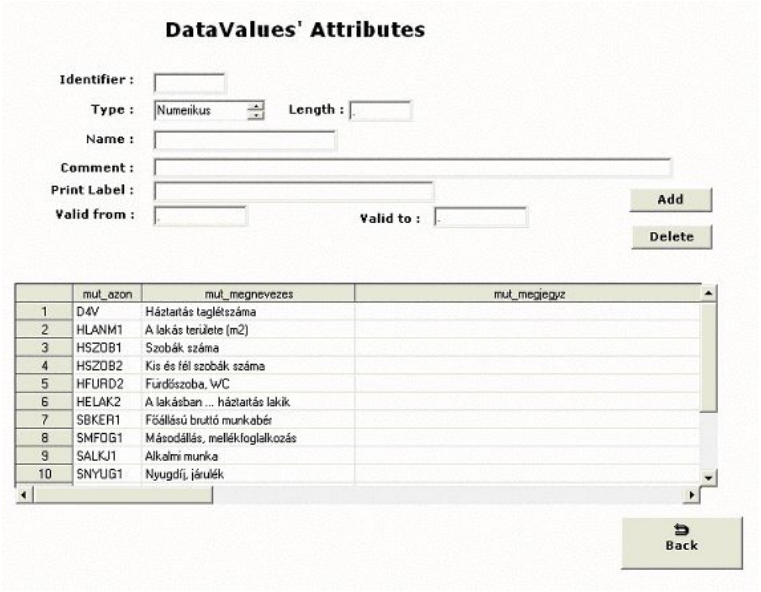

Figure 3:

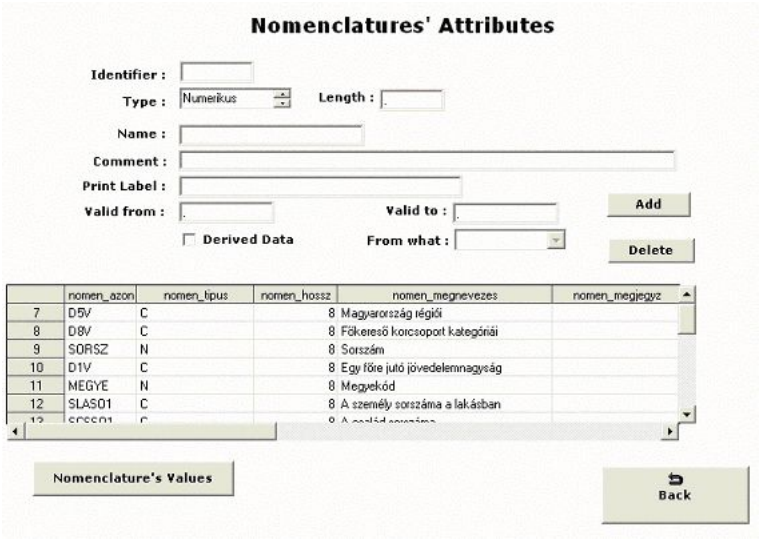

Figure 4:

3.3 Datahandling

Manipulating the input data

The system can handle text and SAS files, expanding the types of importable external files is one of the tasks of the next developing period.

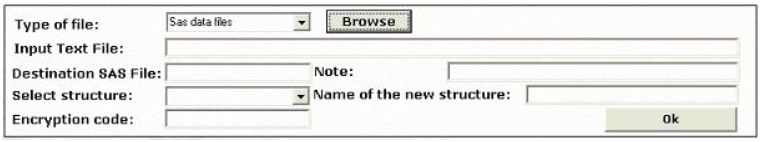

Figure 5:

3.4 Data protection

The purpose of data protection legislation is to ensure that personal data is not processed without the knowledge and, except in certain cases, the consent of the data subject, to ensure that personal data which is processed is accurate, and to enforce a set of standards for the processing of such information.

The Data Protection Act imposes data protection controls over computer and manual records, relating to personal data, access to personal data, processing sensitive data, correction of inaccurate data, criminal records, health records.

Because of these obligations high responsibility is shifted to the holders of data, and data became difficult to access for research centres. For this reason providing adequate data handling is very important.

4 Micromodules

Since several economists (who may be specialist in different fields) use the simulator, elementary algorithms, micromoduls make up the models. For example birth, death, marriage, rise in wages or tax calculation.

From the point of view of the analysts the Micromodul builder is a very important part of the system. Users of the microsimulation system are economists who probably don't have programming skills, but have knowledge of specification of the tasks of micromoduls. With the help of the Micromodul builder the algorithms can be planned, and the appropriate order can also be determined.

4.1 Estimation algorithms and parameter charts

The parameter charts of estimation algorithms can be filled with the help of a graphical user interface, so economists can determine the internal algorithms without any SAS programming knowledge.

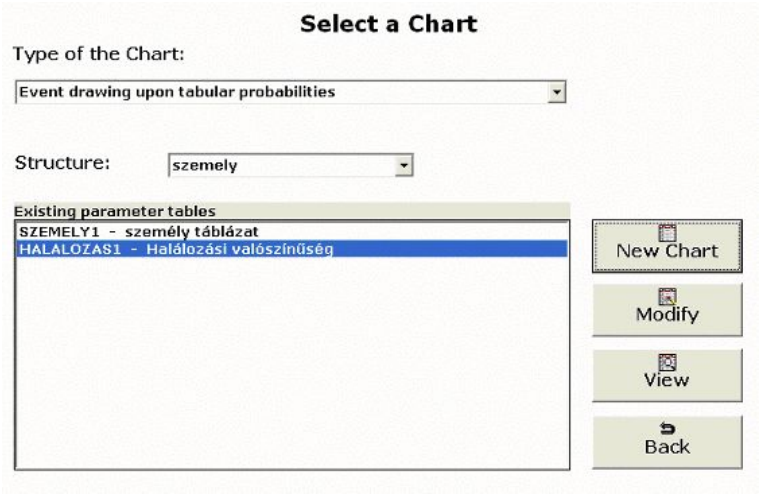

Figure 6:

One of our most important goals is to complete these mathematical algorithms to provide opportunity for making appropriate analysis.

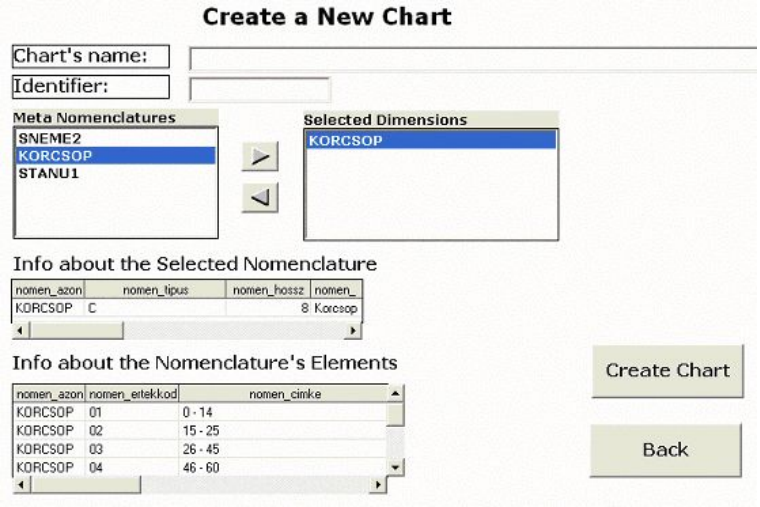

Figure 7:

5 Data analysing

After running the simulation it is very important to have an opportunity to analyse the input and output data. SAS environment provides excellent opportunity for this, because it has a huge range of analytical tools and solutions. The Analyse function of the system can help analysts, who are not experienced in SAS programming, because it contains SAS programs that meet users' requirements.

References

[1] Csicsman József: A BME PIKK Mikroszimulációs projektjének célkitűzései és meggondolásai (Presented at the Budapest University of Technology and Economics, V. Conference on Financial Informatics, 15-16 October 2001)

[2] Daniel Morris: A Micro-Simulation Modelling System for HM Treasury (Presented at the Budapest University of Technology and Economics, V. Conference on Financial Informatics, 15-16 October 2001)

[3] Herwig Immerwoll: The Impact of Inflation on Income Taxes and Social Insurance Contributions (Presented at the Budapest University of Technology and Economics, V. Conference on Financial Informatics, 15-16 October 2001)

[4] Frederik Hansen: The Danish Micro Simulation Tax model (Presented at the Budapest University of Technology and Economics, V. Conference on Financial Informatics, 15-16 October 2001)

[5] Csicsman József: A mikroszimulációs rendszer számítástechnikai hátterének kialakítása. Központi Statisztikai Hivatal, Budapest, 1987. (A Háztartási Mikroszimulációs Rendszer munkálatai, Ts-3/8/8 tanulmánysorozat, 1. kötet)

[6] Dr. Zafír Mihály: A háztartási mikroszimuláció. Koncepció, rendszerleírás. Központi Statisztikai Hivatal, Budapest,1987. (A Háztartási Mikroszimulációs Rendszer munkálatai, Ts-3/8/8 tanulmánysorozat, 1. kötet)

[7] IIASA Workshop on Demographic Microsimulation. Központi Statisztikai Hivatal, Budapest,1988. (A Háztartási Mikroszimulációs Rendszer munkálatai, Ts-3/3/8 tanulmánysorozat, 7. kötet)

[8] A jövedelem modulblokk. Központi Statisztikai Hivatal, Budapest,1988. (A Háztartási Mikroszimulációs Rendszer munkálatai, Ts-3/3/8 tanulmánysorozat, 8. kötet)