THE MICROSIMULATION SYSTEM

Balázs Végh

Department of Information and Knowledge Management
Budapest University of Technology and Economics
H–1521 Budapest, Hungary

Abstract

The microsimulation procedure examines social and economic changes by assessing the effect of each provision by small units. It is routineley used throughout the industrialized world when analyzing for example the distributional impact of taxes and government cash benefits. According to the international applications we have created a SAS Software based microsimulation modelling system. This microsimulation system lets economic professionals easily do their social or economic analysis and it can be a very effective tool for decision makers in the social and economic world.

1. Introduction

In the past, computational simulation was dominated by scientific and engineering problems which had simple data structures which required straightforward manipulations. Today not only scientists and engineers use simulation but professionals of economic and statistic too. Of course these economic simulations differ from the previously mentioned simulations. The subjects are not the events or processes of nature physics or mathematics. The main subjects of economic simulaions are social and economic data, changes and processes. The microsimulation is one kind of these economic simultions.

The microsimulation procedure examines social and economic changes by assessing the effect of each provision by small units. The description of the overall effects is derived from these. The relevance of the results for the society or the economy as a whole is ensured by the database being a national representative sample of the units of such size that guarantees the required statistical reliability. Naturally, the range of social and economic changes that can be modeled this way is defined by the information available on the microsimulation units in a predefined avaiable database. In addition to the generally used economic, social and demographic variables a microsimulation database must contain detailed information on these units. These units can be for example firms and households.

Microsimulation is routineley used throughout the industrialized world when analyzing for example the distributional impact of taxes and government cash benefits. In the past few years Hungarian Central Statistical Office (KSH), Ministry of Economic Affairs (GM), Ministry of Finance (PM) and other departments of the government wasn't able to analyse large data sets and to establish policy proposals.

An effective microsimulation system could be widely applied in Hungary too. The main data sources which can be used in economic simulations in Hungary are:

- the Household Budget survey of Central Statistical Office,

- the (anonymous) tax returns from the APEH (Internal/Inland Revenue Service) database.

Now let's see the common requirements of the microsimulation system in the next chapter.

2. The common requirements of the microsimulation

The common requirements of an economic simulation system:

- Data,
- Security,
- Efficiency,
- Flexibility,
- Economic profitability.

The data sets are indispensable. We must provide datas for the simulation. One possibility is collecting new data. Differerent surveys are the main sources of economic simulations. We must mention that the execution of such surveys is not only very costly and time consuming, but would not be reasonable due to reliability aspects, either.

A new data set can be created for the simulation by merging various already available data sets. It is an other possibility to rewrite available and newly created data sets using economic and statistic methods. These processes can be solved with a sorely planed microsimulation system.

Data handling and the security of data aren't negligible. Economic simulations use data of the society and the economic. These data are confidential, therefore a third person must be prevented to get intern information. The simple spreadsheet and analysis programs can't provide the reasonable level of security. Today the data security is an elementary demand on data-handling systems. This is especially important in economic and governmental areas.

Building new statistic and economic models isn't easy either. The demands are continually changing and the applied model must follow the new requirements. This situation is very wasteful of money and effort. It is more significant if we want to build up a new model from the beginning. This is the problem of flexibility.

The efficiency can be increased with the reuseability of resources. This can be solved with the maintenance of the once previously created functions.

A simulation system with all these solutions enhance the economic profitability too.

The most important tasks of microsimulation:

- Iproving quality of statistical surveys (by re-weighting),
- Updating statistical surveys,
- Making forecasts,
- Investigating effects of measures.

3. The common and functional requirements of the system

3.1 The features

The economic professionals need an effective, userfriendly tool to make their work efficient. The computer and software technology make these requirements achievable. Software technology gives all the properties of such a system. The good points of an effective userfriendly universal microsimulation system are the follows:

- "Point and click" userfriendly interface,
- Variety of languages,
- Easily maintain and process even the largest data sets,
- Project based,
- Significantly reduce development time for new models,
- Quickly create wide range of analyses,

- Easy exporting to other applications,
- Support many platforms,

If the professional have the effective tool with the requirements mentioned before, he must do the economic simulation. These are the main tasks of a microsimulation system:

- data access,
- data management,
- analysis,
- presentation.

3.2 Data Access

The most important is the subject of the simulation, the data. A microsimulation system requires data. The sources of data can be very different. Data must be imported from the available data sources. These sources can be any file which is in any format like variable-length records, binary files, free-formatted data - even files with messy or missing data.

3.3 Data Management

After data have been imported into the system, the data must be registered and managed in the system. It means that the available attributes needed by the further functions and descriptions of each data and data sets must be stored in the system too. This is the role of Meta system. This meta system is a subsystem of a microsimulation system.

The data management must realize data security. The questions of data security is one of the most important tasks of the data management. It must be guaranteed that only that users are allowed to work with the data sets of the system who really have the required rights for them. These rights can be different. The users do different works on the data, so they should have only that level of rights which is just enough for his work.

3.4 Data Analysis

The main task of the economic professionals and the microsimulation system is modeling economic processes and changes, and analyze their effects on the different part and levels of the society. Computing must be easy for economic analysts. Therefore the system should provide powerful data analysis tools. It should be able to create analyses to concisely summarize the data, such as frequency counts and cross-tabulation tables. An analyst likely want to compute a variety of descriptive statistics, including the mean, sum, variance, standard deviation and more. Compute correlations and other measures of association, as well as multi-way cross-tabulations and inferential statistics. The analysts have the demand of design, development, test and use of new models and methods too. Above all the system must provide tools for many kind of business need: planning, forecasting, quality improvement, project management, decision support and more.

3.5 Data Presentation

Of course it is essential that presentation must be prepared. This is the need of the decision makers too. In the final steps the system produce reports that range from a simple listing of a data set to customized reports of complex relationships. Nowadays a developed system must be able to present certain information and results for the world via Internet.

4. The realization

4.1 The reasons

Today standards such as Excell and Access programs give a very simple solution to the user and the portability between different windows platforms and Pc's is easy too. On the other hand the software efficiencies are low.

There is a high demand for an appropriate microsimulation model. In 2001 at the Budapest University of Technology and Economics a more-year project started with the purpose of constructing a microsimulation system. The microsimulation group was established in 2001 at the Department of Information and Knowledge Management.

According to the international applications we have created the technical and methodological conditions for analyzing large data sets and we have developed a SAS Software based microsimulation system.

This system realize a dynamic microsimulation model as an appropriate tool for economic analyses. Dynamic models subject the individuals in a large population sample to the chances of various events occurring. Such events include: demographic events, such as birth, marriage or death. As a result, dynamic microsimulation models have the capacity to generate highly detailed projections of changing population characteristics, as well as the lifepaths of individuals. This capacity makes dynamic microsimulation a particularly powerful tool for applications where there is an interest in people's financial situation through life. Dynamic microsimulation models have thus been used in areas such as superannuation, age pensions and education. Other areas where this modeling approach promises added insights include long-term issues in labour market, health, aged care and housing policy, and broad long-term distributional issues within the population and across generations.

The usefulness of this model for assessing the cost and distributional implications of possible policy changes for existing clients has also been greatly increased by the development of a user-friendly interface, making the model and data accessible to non-programmers.

- This system can access local SAS and non-SAS data types.

- Import free-formatted data - even text files with messy or missing data.

- It delivers maximum flexibility for dynamic economic and business requirements.

- Business needs can be met and changed based on demand.

4.2 The system is based on SAS

The system has the need for SAS software to be installed on a user's desktop PC. The development tools are Base SAS software and SAS/AF software.

The Base SAS software provides tools essential for mastering the four data-driven tasks common to virtually any application: data access, management, analysis and presentation under a powerful applications development environment. This helps in bringing all application needs and organizational data into a single system. The powers of Base SAS software can be extended to other SAS Solutions.

The SAS/AF software extends the power of the SAS System throughout the organization in the form of customized, point and click applications. This interactive applications development environment lets to harness the power of the SAS System for data access, management, analysis, and presentation. SAS/AF software gives users quick, easy access to current information, develops point-and-click applications that can be invoked using only a mouse or function keys. Both new and experienced users can put the SAS System to work without writing any code or knowing the SAS language. The developers need only create an application once, as SAS System is portable across hardware platforms. Now the development work can be done on a single platform, then port the completed application to other environments with little or no modification.

4.3 Languages

The language of the program is optional. It is a good point of this system. So not only Hungarian economic professionals but the professionals of other countries can use this system too. At the moment English and Hungarian versions are available. The system uses a dictionary with its own structure and one macro function to

assign the current label or phrase to the matching object. If the system is working this function is called every time before the presentation of the next frame. The function get the current parameters of the system. Using these parameters the function opens the dictionary and gets the all the data which are needed to the presentation of the graphical user interface.

The extension of the languages and the words in the system is easy. This system gives a userfriendly graphical user interface for editing the dictionary of the languages. The compilation doesn't demand any special knowledge, it is very easy to translate the contents of the dictionary into any other language. It is a big advantage of our system that it has no limit in the number of applied languages. It is only the question of the interpreters talent.

4.4 The projects

The data sets, the users, the tasks are different. The system groups all these into projects to guarantee the transparency and the order of the work. A project provide the required environment for the users of microsimulation.

One project contains and describes the follows:

- Dictionary structure,
- Data sets,
- Users,
- Warrant,
- Micro modules,
- Parameter determinations,
- Parameters and settings of simulation models,
- Analytic functions,
- Analytic results.

A strictly structured library system belongs to each project. The user can do his work after selecting one project. The data, the sources of the simulation are in the data source library of the selected project. There are many other important specific data sets in the libraries of the current project. These data sets contain all required information and the main parameters of the system.

4.5 The security

The system must be reliable and safe. This is the task of a security system. Therefore the microsimulation system has his own security module and policy. The security is solved in the frame of the project too. The different users have different access rights to the data sources and this rights cab be different in the different projects. Not only the data sets but the system and the functions of the system must be managed in the aspect of access. There are different levels of the user rights, accessing the data and using the functions.

There are three different subjects of security:

- User,
- Data,
- Function.

There are data sets in this module and these data sets contain the common and security information of the users, the data and the functions. Of course these are three independent data sets. These data sets describe the level and the parameters of the encryption too. For example there are the identifiers, passwords, encryption codes, keys, paths in encrypted formats. The security module manages all the rights in the system. There are some specified administrative users with special rights. Only they can use the userfriendly graphical user interface and they can easy determine and set the rights using this interface.

4.6 The Meta system

After selecting or creating the Project data from an external data source could be read and written to a SAS data set ("Data Handler"), and the Meta database could be viewed or changed.

The first step in model building have to be the creation of data sets. These data sets must have a sufficient detailed variable structure. This very important function is the role of the own Meta database system. Briefly, the role of the Meta database system is containing all the information about the data. There is a Meta dictionary in this system, which contains these information. These information describe not only the data but the structures and hierarchies of the data too on a strictly predefined way.

There are information in the meta system from:

- Data,

- Data sources, files,

- Structures of data in each data set,

- Structure of data sets,

- Hierarchy,

- Inheritance of data sets.

The meta always must be up to date and it must contain all the required attributes and descriptions. These attributes are identifiers, types, lengths, names, paths, descriptions, validations, relations. All these are indispensable for the microsimulation system.

The row data must be prepared sorely before the microsimulation system can use it. The data handling functions can create the appropriate data sets for the microsimulation system. This step uses the meta system. The main task of the creation of data set is managing the problem of missing data. The data in the microsimulation system can be manipulated in many predefined way. Powerful tools must help to maximize data processing efficiency, with facilities for sub setting and merging a wide range of well prepared data sets.

4.7 Data handling

Naturally the subjects of microsimulation as any other simulations are data sets. These sets naturally contain the data of different samples. The most difficult technical task is creating a new database. We can import a new data set from a source data file. The other possibility is the merging different already avaiable data sets.

Let's see a typical example, the creation of a new data set of households. This involve a procedure during which attributes can be assigned to the current household of one data set that is not available in the given data set, but is available on many other households of another data set. The criteria of how well the matching works is whether the attributes assigned to the given household come from a household of the other data set that is similar to the one we are assigning to in the most significant aspects.

The datas imported from surveys are not realiable. The reweighting of the avaiable data is a need. Since the start of this project different approach has been taken to calculating weights. Prior to this the reweighting process involved using parameter tables to calculate weights for people in the micro data.

4.8 The micro module building module

The micro modules realise the parts of the economic models which are described by the economic professionals. These are complicated complex mathematical and statistical functions with many parameters.

The most important goal of the development is to replace the SAS programs with micro modules. This is the task of the micro module building module. This is a graphical development tool in the system. This module gives a userfriendly tool for the development of micro modules. This is a great help for economic professionals to develop their functions and models. Of course the new model can use only the resources avaiable in the microsimulation system. The other advantage is that the user can reduce the most syntactical and many semantical errors. So the model develped by the user must be more reliable.

4.9 The parameter table handling module

During the model building the economic professionals describe the main rules of the model. This simulation models have many parameters. The most of them must be able to alter. There is a module for this purpose too.

The parameter table handling module has to perform the following tasks:

- Provide a graphical interface for the user.

- Different types of tables and different structured tables have to be handled.

- With using the system, the user has to be able to create new tables, to set the name, the dimensions and the values of the cells of the new table, or to modify already existing ones.

- Provide the contents of the suitable cell of the suitable table for the estimation algorithms during the run of the simulation.

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.

4.10 The simulation module

The simulation module is the main part of the system. The previous modules helped us to provide the conditions for the embraceable and reliable work via projects, the row data able to work with them, to build the flexible new economic or statistical model. With the help of these modules we have the opportunity to modify input data in different ways and to compare the results (for example to analyse the tax rate plans).

The simulation module provides a userfriendly interface too. The interface includes simple parameter modification, choice of year for modeling and standard output tables for distributional model. The system allows the analysts to change the rules or the parameters of the model. So the existing model can be refined. The analyst can compile his new model with the previous defined data, rules, functions, models and parameters. Of course he can change the parameters or the micro modules of one previous completed simulation too. This reusability is a great advantage of the system, and all parts of the model can modified and reused in a new simulation.

The main feature of this microsimulation model in relation to social and economic policy analysis is that they produce results that can be analyzed, repeated and refined.

4.11 Analyze

After running the simulation it is very important to have the opportunity to analyse the input and output data and all the results of the completed economic or statistical simulations. The analyse module of the system can help analysts, who are not experienced in SAS programming, because it contains already avaiable SAS programs that meet users' requirements and the analyst can define and realize his own analyze functions too.

One of the most important goals of the system is to complete the mathematical algorithms to provide opportunity for making appropriate analysis. The microsimulation system must be able to present the results according to the demands of the social and economic world. The system is able to present this result in many format. One of these is the Html format, the format of the Internet, so the results can be automatically presented via Internet with the required delimitation of access.

5. Summary

As a result of this project a SAS Software based Microsimulation Modeling System was developed which is suitable for modeling 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.

This system with all the advantages of economic simulations and userfriendly graphical interfaces gives an universal tool of Microsimulation.

This Microsimulation system can be a very effective and economic tool for decision makers in the social and economic world.

6. References

[1] Dr. Zafír Péter:: A háztartási mikroszimuláció. Koncepció, rendszerleírás. Központi Statisztikai Hivatal, Budapest,1987

[2] IIASA Workshop on Demographic Microsimulation. Központi Statisztikai Hivatal, Budapest,1988