DEVELOPING A MICROSIMULATION SERVICE SYSTEM

Joseph Csicsman and Cecilia Fenyes

Department of Information and Knowledge Management

Budapest University of Technology and Economics

H-1111, Budapest, Hungary

E-mail: csicsman@calculus.hu

fenyesc@itm.bme.hu

ABSTRACT
The term of microsimulation

The microsimulation procedure examines social and economic changes by assessing the effect of each provision with small units and the description of the overall effects is derived from these assessments. Relevance of the results relating to the society as a whole is ensured by the database which is a national representative sample of the units or households of such size that guarantees the required statistical reliability. Naturally, the range of social and economic changes that can be modelled in this way is defined by the information available on the microsimulation units in the database.

Objectives of the project

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 establish policy proposals.

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

A research group was founded for the project jointly by the Research Centre for Financial Economics of the Budapest University of Technology and Economics and Új Calculus Bt. The aim of the project is to develop and maintain a model system which allows analysing data collections and databases available in the administration for economists. Naturally, this methodology can also be used for processing other types of data too.

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

Till now a methodology to assess the impact of government programs has not been available. By utilizing research experience of the university more exact analyses could be prepared to qualify and quantify policy options.

As an outcome of this project a Microsimulation Modelling System will be developed which will be suitable for modelling the decisions of economic and social policy, and – complying with the national and international requirements – well-founded analyses can support the policy proposals of the government.

The following description will present the recent results of the project in a framework in which great number of university theses have been developed.

THE MICROSIMULATION SERVICE SYSTEM
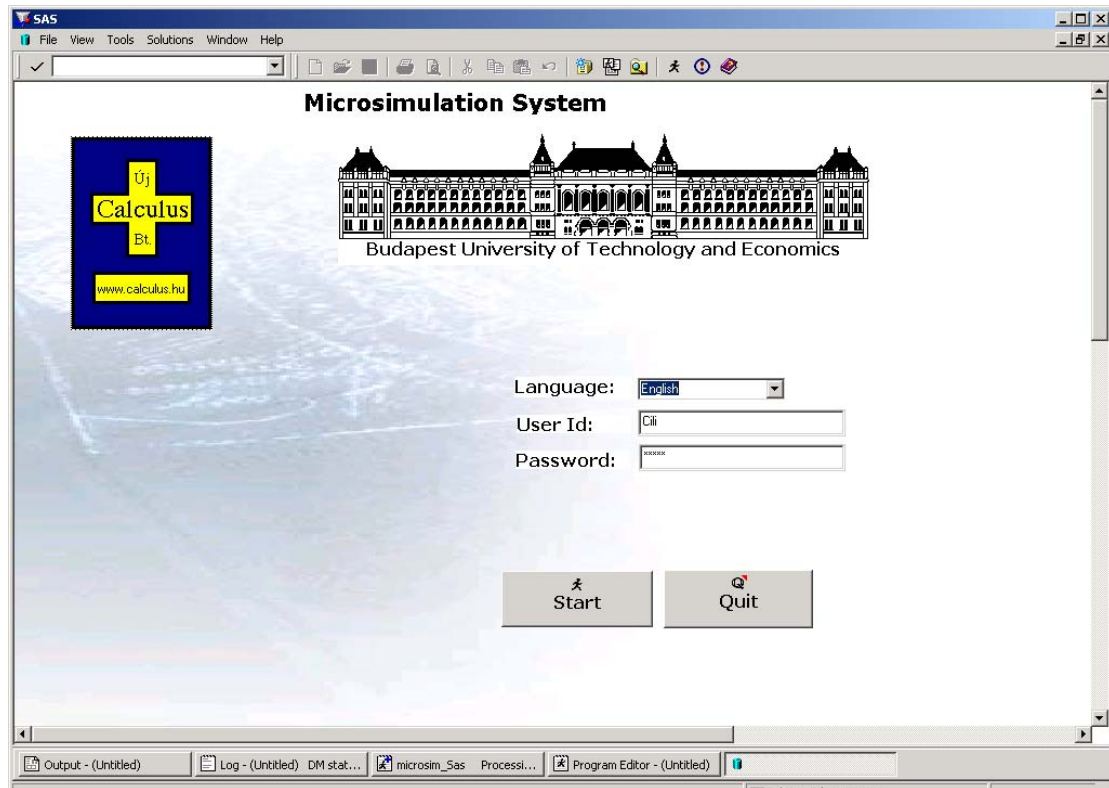
As it can be seen in the Main menu, the language of the program is adjustable. At the moment English and Hungarian versions are available, but because the system works from a dictionary and the compilation doesn't need any special knowledge (only the command of the language), it is very easy to translate it into any other language.

Project selection

Prior to running a particular simulation a Project is to be selected (or created, if it does not exist), because the next steps are related to the data we want to use.
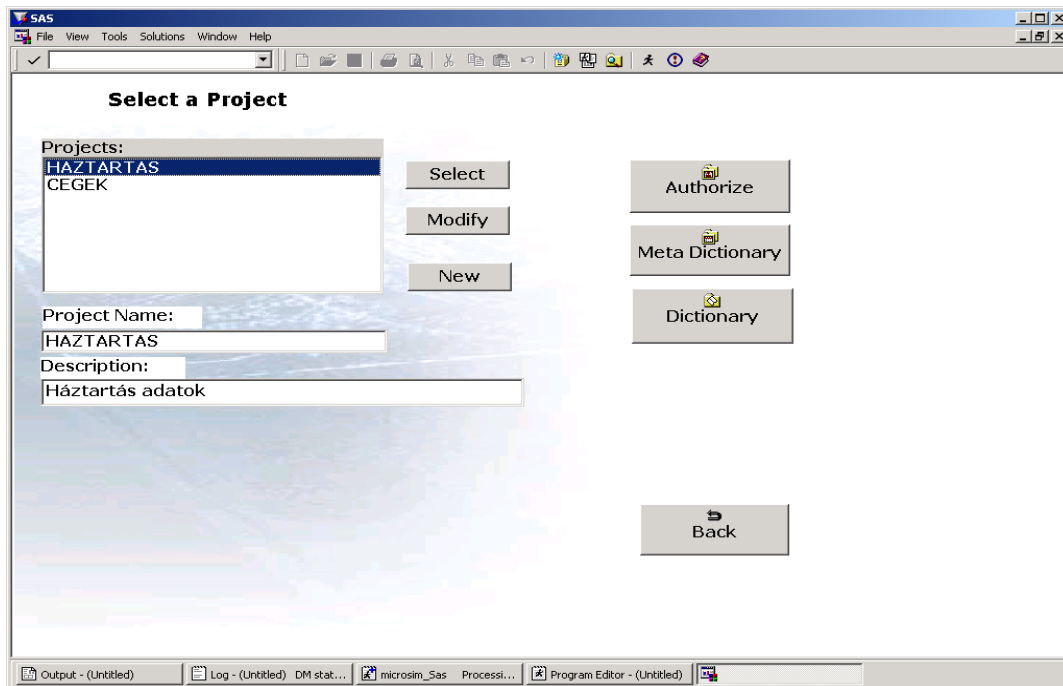
Without selecting a project 3 tasks can be chosen. By selecting 'Authorize' rights of the users can be set. By pushing the button 'Meta dictionary' the data's of meta can be read or set, which are independent in any project (nomenclatures and numeric values). By pressing Dictionary the dataset of languages can be found. After selecting or creating the Project or selecting a task, the Project Handling frame will appear:
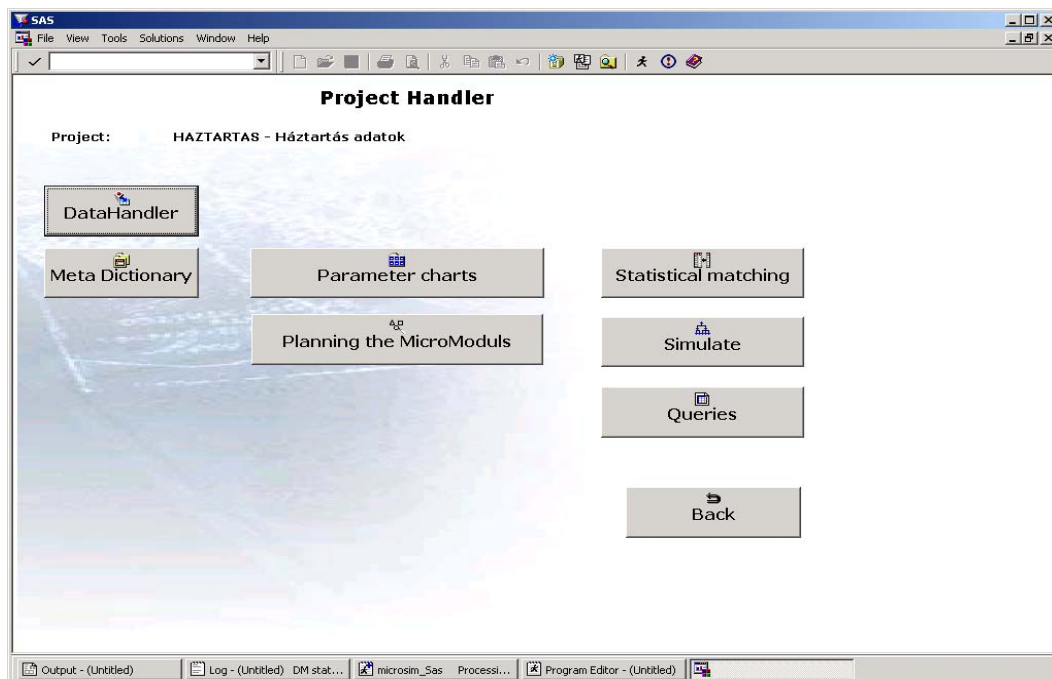
Datahandling

Manipulating the input data

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

Data protection

Due to the public environment use of a password entry and different levels of users is unavoidable.

Meta Dictionary

The Meta Dictionary contains all the information about the data and datasets: identifier, type, length and name of nomenclatures and pointers, structures of input datasets. Also the file catalog is the part of the meta. On the following screenshot the list of nomenclatures can be seen with a user friendly screen to view or

modify it. Numeric values, structures of datasets, and the file catalog will be shown in similar frames.

Estimation algorithms

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.

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

One of the frames, which helps to fill in the 'Value assignment upon distribution' parameter chart can be seen below:

Micromodules

Micromodules are Base Sas codes. During the simulation the selected micromodules will run on every record of the input dataset. This means that micromodules set the changes, which will occur. Here

we prepared the platform for making micromodules without the knowledge of Base Sas. The code will be generated from the rows of modul steps (down, left) made by the user using almost only the mouse.

Project: HAZTARTAS – Háztartás adatok
Structure: SZEMELY

Micromodul modification

Micromodul: ado2003   **Description:** Adószámítás   ☐ Use technical file

☐ Called step

Steptype
◉ 1 – Simple condition
○ 2 – Complex condition
○ 3 – Set value

Operators
AND
OR
NOT
<
>
=
+
-
*
**
/
(
)
!!
"
,
SUM

Input variables
MEGYE -- Code of county
RND -- Véletlen szám
RUN --
SBKER1 -- Főállású bruttó
SCSAP2 -- Family status
SFEOR1 -- FEOR number
SNEME2 -- Sex
SNYUG1 -- Nyugdíj, járulé
SNYUJ1 -- Le: nyugdíjjáru

Stepnumb 10
Identifier: a
Description:

☑ Expression   SBKER1 <650000
☐ True   SZJAD1 = SBKER1 *0.2   [Continue]

[New step]   [Save step]

Set of values

Called stepid:
Value: .
Text:

Estimating tables
TANULMPENZ
FIZEMELES
KER_96_03

Global variables
DEPENDENTS -- Number
AGE -- The age of the p

Local variables

Technical variables

Modul steps

Stepnumber	Step ID	Steptype	Expression
20	b	1 -- Egyszerű feltétel	SBKER1 >650000 AND SBKER1 <1350000
30	c	1 -- Egyszerű feltétel	SBKER1 >1350000
40	d	1 -- Egyszerű feltétel	SBKER1 >0 AND SBKER1 <2368850
50	e	1 -- Egyszerű feltétel	SBKER1 > =2368850

[Delete step]
[Modify step]
[New variable]
[Save]
[Cancel]
[↩ Back]

Running the simulation

This frame is to set all the parameters of the simulation. Input file(s) can be chosen, name of the output file must be set, and here the micromodules can be selected to run from the list of the premade micromodules.

Versions: Set parameters and Run

Project Name: HAZTARTAS   Háztartás adatok

Versions:
1 Done
2 Done
3 Done

[New Version]   **Version:** 1   háztartás szimuláció
☑ Works with a Structured File   **Actual Year:** 2003

Files
Layer: 1
Input file: haztartas
Output file: htout
Technical file: techht
Structures:
Technical Files techht

MicroModuls
Layers: 1
◉ Run Before Next
○ Run After Next

MicroModuls
DEPNULL
FAMALL

Selected

△   □ Clear
✕   1,2,3. ..6,7.. ReNumber
▽   Random

[Save]   [Save as New]   [Save & Run]   [Run]   [↩ Back]

SAS — File View Tools Solutions Window Help

☐ Output - (Untitled)   ☐ Log - (Untitled) DM stat...   microsim_Sas   Processi...   Program Editor - (Untitled)
C:\WINNT\system32

Analysing

After running the simulation it is very important to have opportunity to analyse the input and output data. The

Statistical Matching

Statistical Matching is to pair the records of two data sets without having any key variable. The records of the secondary data sets are separated into groups by their selected attributes (they can be defined in the section of parameter charts, by selecting 'statistical matching - parameter chart of teams' from the list of type of table). The statistical matching goes through on every records

Analyse function of the system can help analysts, who are not experienced in SAS programming. For Analysis the procedures of SAS can be used very well.

of the primary data set. By the attributes of the record the member of the appropriate group (which is a record of the secondary data set) will be paired with this record.

If there is only one data set and the records of the same dataset are intended to be paired (like the simulation of marriage), it is also solved in the microsimulation system.

PRACTICE

Real application of the Microsimulation System started in the Hungarian Central Statistical Office last year.
The known errors of Household Statistics 2002 were corrected using standard simulation technics.
Income Distribution was modified with statistical matching, based on administrative income registration system.