WORKING PAPER

"DIAL" – A System for Modeling Multidimensional Demographic Processes

Sergei Scherbov
Vladimir Grechucha

May 1988
WP-88-36

IIASA
International Institute
for Applied Systems Analysis

"DIAL" – A System for Modeling Multidimensional Demographic Processes

Sergei Scherbov
Vladimir Grechucha

May 1988
WP-88-36

INTERNATIONAL INSTITUTE FOR APPLIED SYSTEMS ANALYSIS
A-2361 Laxenburg, Austria

Foreword

Multistate mathematical demography, much of the development of which took place in IIASA, has proven its usefulness in a broad range of applications including the analysis of migration, marriage, fertility, working life and household dynamics. Among other things the method makes possible the calculation of population projections which are disaggregated by region (marital status, parity, occupation, etc.). The following paper serves as a user's guide to a new microcomputer program which greatly facilitates the use of the multistate projection mathematics. A user with access to the appropriate data and modest skills with a computer can explore the future path of a population under a variety of assumptions about the direction of change in key variables. Graphical displays of results, and interactive updating of assumptions, also contribute to the usefulness of the system. The program described here has already been installed and used in a number of research institutes in several countries.

Douglas Wolf
Deputy Leader
Population Program

Acknowledgements

I would like to thank Frans Willekens for his substantial comments and recommendations regarding the text of this paper, Douglas Wolf for his useful comments, Babette Wils for her editing skills, and Susanne Stock for typing the manuscript.

Contents

"DIAL" – A System for Modeling Multidimensional Demographic Processes

*Sergei Scherbov**, *Vladimir Grechucha***

1. INTRODUCTION

The rapid growth of the availability of microcomputers and user-friendly software makes it possible for non-professionals in computer sciences to use application programs. Contemporary demography has a strong quantitative component and this has inspired the development of software for demographic research, both for microcomputers and mainframes (Strong 1987). This paper continues the work started by Scherbov and Grechucha (1986) and presents a DIALOG software system for multiregion (multistate) population projections that resulted from methodological research at IIASA. This system has been implemented on a microcomputer and is intended for use by nonspecialists in computers.

Multistate population models have recently become popular in the study of many aspects of demographic transitions, such as migration, marriage, changes of health status, social status, occupation, etc. (Keyfitz 1980; Yashin 1977).

Computer programs and software packages were developed to realize such models (Willekens and Rogers 1978; Willekens 1979); however, most of these allow analysis of systems only when fertility, mortality, or transition coefficients are constant over time. Some authors have overcome this drawback (Ramachandran 1980; Scherbov and Usbeck 1983) by creating the opportunity to analyze alternative evolutions of the system under various scenarios of natural and mechanical reproduction of the population. However, these programs are not always appropriate for use by the many demographers and health specialists who are not deeply involved in computer modeling. Often the software is not flexible enough to allow choices of and changes in the variables that determine the scenarios, the representation of the results, and the control of modeling itself.

*Population Program, IIASA, A-2361 Laxenburg, Austria
**All-Union Institute for Systems Studies (VNIISI), Prospect 60 Let Octyabria, 9, 117312 Moscow, USSR

The most important disadvantage of these packages is the inability to communicate interactively with the model. As experience shows, interactive work with computers considerably reduces the time spent on model design and debugging. It also creates additional opportunities for model analysis. Thus, there existed a necessity to create a user-friendly system allowing a more effective analysis of demographic processes.

In this paper an interactive system that uses the multistate demographic models is described. The system provides the opportunity to prepare scenarios, change coefficients of the model during the modeling procedure, and present intermediate results. The paper uses some results of research conducted at VNIISI and at IIASA, namely, the design of the man-machine modeling system (Gelovani 1980) and the modeling of multistate demographic processes (Willekens and Rogers 1978; Willekens 1979; Scherbov and Grechucha 1984; and Rogers 1975).

2. STRUCTURE AND ELEMENTS OF THE DIALOG SYSTEM

2.1. Requirements for the DIALOG System of Modeling

The elaboration of the interactive system was based on the following main requirements.

1) The DIALOG system should be simple in that the command language should be as close to natural language as possible. Since the system is oriented towards non-specialists in computer science, the leading role in the DIALOG should be played by the system itself. The user should only answer simple questions or select instructions from a menu.

2) The DIALOG should be organized in a way that the system's reaction is user-friendly when users provide wrong answers, and does not stop operating. The "help" option should be provided.

3) The facility to change DIALOG languages should be provided.

4) The system should automatically adjust to different types of initial data. This may be a different number of regions or states, a different number of sexes, or flexible age group lengths.

5) A visual display of the modeling results should be provided. One should be able to obtain tables and graphs (on a graphic display, or printer) during the modeling.

6) The system should provide opportunities for flexible scenario setting for different controlling variables. For demographic models the opportunity to set scenarios for such variables as total fertility rates, life expectancies, age-specific mortality, age-specific fertility, transition or morbidity intensities, etc., should be provided. The scenarios should be set either in an interactive mode or from a file prepared in advance.

7) The opportunity to easily transfer the model from one computer to another should be provided.

The block scheme for the DIALOG system, which satisfies most of the above requirements, is represented in Figure 1.

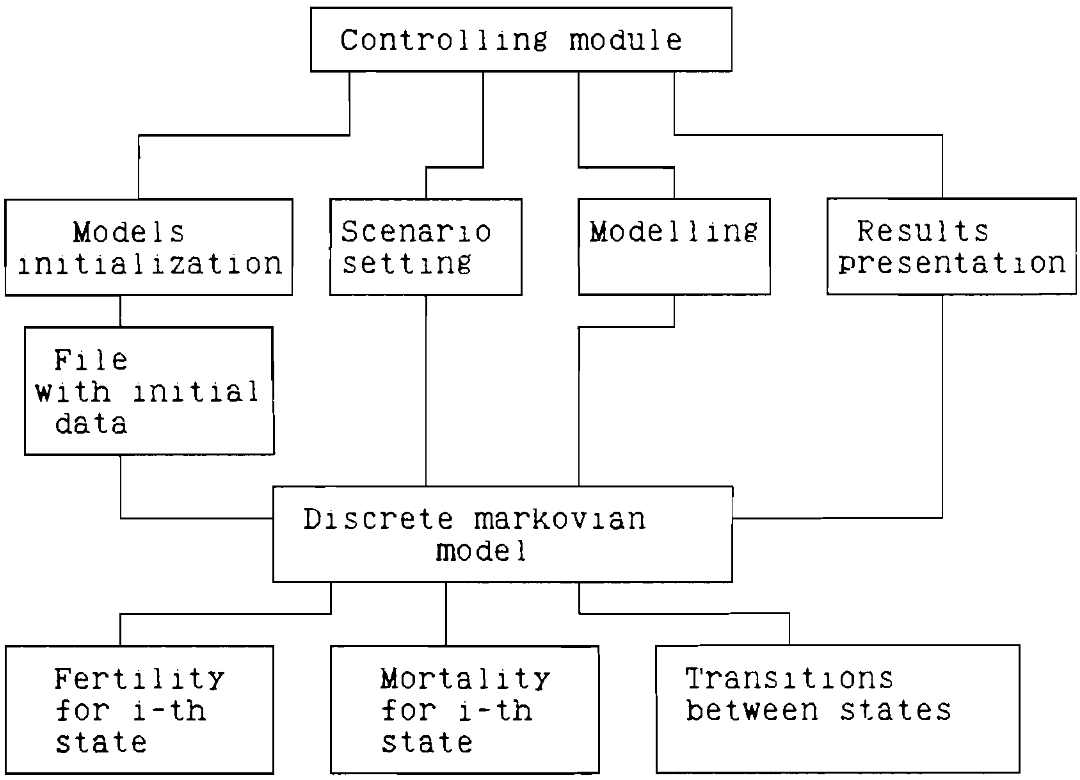


Figure 1. Structure of the DIALOG system.

2.2. Description of the Model

A detailed mathematical description of multistate population dynamics is given by Rogers (1975, 1980). We will discuss this very briefly. Suppose we have a population divided by age and state of residence. Let m be the number of age groups, h the length of the age group, and n the number of states. Then we could describe the distribution of population for any given time t as a vector $K(t)$:

$$K(t) = \begin{bmatrix} k_{11}(t) \\ \vdots \\ k_{1n}(t) \\ k_{21}(t) \\ \vdots \\ k_{2n}(t) \\ \vdots \\ k_{m1}(t) \\ \vdots \\ k_{mn}(t) \end{bmatrix}$$

where $k_{ij}(t)$ is the population of age group i resident in state j at time t. The dynamics of the population in such a system can be described by the equation:

$$K(t+h) = G(t)K(t)$$

where h is the length of a cohort that is equal to the projection step and G is a growth matrix:

$$G = \begin{bmatrix} 0 & 0 & B(a-h) & \cdots & B(b-h) & \cdots & 0 & 0 \\ S(0) & 0 & & & & & \cdot & \cdot \\ 0 & S(h) & & & & & \cdot & \cdot \\ \cdot & \cdot & \cdot & & & & \cdot & \cdot \\ \cdot & \cdot & & \cdot & & & \cdot & \cdot \\ \cdot & \cdot & & & \cdot & & \cdot & \cdot \\ \cdot & \cdot & & & & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & & & & S(z-h) & S(z) \end{bmatrix}$$

where a and b are the first and last ages of childbearing, respectively, and z is the last open-ended age group. The first row of matrix G consists of the matrices $B(x)$ which are of dimension n-by-n

$$B(x) = (h/4)C[P(0) + I][F(x) + F(x+h)S(x)] \quad .$$

$F(x)$ is the diagonal matrix (n-by-n) of fertility rates of people aged x to $x+h$-years old in different states. The C matrix determines the state in which newborns appear in connection with their parents' state. Thus, the rows of the C matrix represent the states of the newborns and the columns represent the states of the parents. If, say, $C(i,j) = 1$, then the parents are in state j and the children will appear in state i. The C matrix satisfies the usual probabilistic constraints

$$C(i,j) \geq 0 \, , \quad \sum_j C(i,j) = 1 \quad .$$

The P matrix contains the survival probabilities from age x until $x+h$. The entries under the main diagonal of the G matrix determine the survival coefficients

$$S(x) = [I + P(x+h)] \; P(x) \; [I + P(x)]^{-1} \quad ;$$

or, allowing multiple transitions (Willekens and Rogers 1978), it may be expressed as

$$S(x) = [I + h/2 \, M(x+h)]^{-1} \; [I - h/2 \, M(x)] \quad ,$$

where I is an indentity matrix. For the last open-ended group z the linear approach (Just and Liaw 1983) was used, thus giving

$$S(z) = [I + h/2 \, M(z)]^{-1} \; [I - h/2 \, M(z)] \quad .$$

The transition probabilities are calculated in the same manner as for increment-decrement life tables. At first the observed coefficients are grouped into the matrix

$$M(x) = \begin{bmatrix} [M_{1d}(x) + \sum\limits_{j\neq 1} M_{1j}(x)] & -M_{21}(x) & \cdots\cdots & -M_{n1}(x) \\ -M_{12}(x) & [M_{2d}(x) + \sum\limits_{j\neq 2} M_{2j}(x)] & & -M_{n2}(x) \\ \vdots & \vdots & & \vdots \\ -M_{1n}(x) & -M_{2n}(x) & \cdots\cdots & [M_{nd}(x) + \sum\limits_{j\neq n} M_{nj}(x)] \end{bmatrix} .$$

And according to (Rogers and Ledent 1976) P is calculated from

$$P(x) = [I + h/2 \; M(x)]^{-1}[I - h/2 \; M(x)] \quad .$$

Just (1983) has shown that to complete the female dominant two-sex model, a diagonal matrix X^f has to be defined. Its elements are the ratios of males to females borne by women in each state. Usually these ratios are independent of the status of the mother. Thus

$$X^f = \begin{bmatrix} v_1^f & \cdots & 0 \\ 0 & \ddots & 0 \\ \vdots & v_j^f & \vdots \\ & & \ddots & \\ 0 & \cdots & v_n^f \end{bmatrix}$$

where $v_j^f = 1/(1 + (b_j^m/b_j^f))$, and b_j^m and b_j^f denote the number of male and female births, respectively.

For the projection of the female population the first row of the growth matrix is given by

$$B^f(x) = (h/4)\ CX^f[P^f(0) + I][F(x) + F(x{+}h)S^f(x)] \quad .$$

The superscript f indicates that the survivorship proportions and probabilities of the female population are used. Superscript m is used for the male population. The projection of the male population is performed in two steps:

1) The total (male and female) population at exact age 0 (births in period $(t,t{+}h)$) is calculated by

$$K_0^*(t) = \sum_x [F(x) + F(x{+}h)S^f(x)]\, K^f(t) \quad .$$

2) The male population in the first age group at time $t{+}h$ can easily be derived by means of

$$K_0^m(t{+}h) = CX^m\,[P^m(0) + I]\, K_0^*(t)$$

where the elements v_j^m of X^m are $1 - s_j^f$. Population by states in other age groups is calculated by

$$K_{x+h}^m(t{+}h) = S^m(x)K_x^m(t)$$

$$K_{x+h}^f(t{+}h) = S^f(x)K_x^f(t)$$

where $K_{x+h}^f(t{+}h)$ is the population distribution by states of female population who are in age group $x{+}h$ at time $t{+}h$.

2.3. The Structure of the Interactive System

2.3.1. *Control Module*

The DIALOG system is designed using a modular approach. This approach allows flexibility in updating the system, changing old modules, or adding new ones to the system to solve different tasks. Also, the demographic model itself is represented as a separate unit. This allows the model to be updated and new variables or subprograms to be entered. The controlling module generates main tables of model variables, switches control between other modules, and checks the memory distribution. All modules are interconnected through the control module. This provides for independence and for flexibility in updating. Data flows between different modules and the model pass through a COMMON block, which is saved in a separate file COMSPAT.FOR and is used during compilation and initialization of the system.

After starting the DIALOG system the control module reads the file COMSPAT.FOR and builds the table with the variable names and their addresses in the COMMON block. This table is used later for scenario setting and presentation of the results. When the table is completed control can be transferred to the other modules.

2.3.2. *Initialization Module*

During model initialization the file with the initial data is read. This file should be prepared in a particular form, described in Appendix A.1 and A.2.

After reading the initial data the initialization unit calls the demographic model and all variables of the model are calculated for the initial year (for the list of variables, see Appendix A.3). This allows the user to analyze demographic indicators for the initial year immediately after initialization.

2.3.3. *Scenario Setting Module*

The scenario setting module provides setting of control variables in the form of a time-series. In DIALOG, the controlling variables (or scenario variables) are the exogenous parameters for the demographic model used by the model at each simulation step. The set of controlling variables which is defined for some time interval is called a scenario.

Scenario setting can be performed in the interactive mode or by calling a previously stored file. There is also an option to store scenarios that were set in interactive mode in the file for future use. Insofar as scenario variables depend on a particular model, a scenario is generally set using variable names. In addition, there is an opportunity to set scenarios for the main demographic indicators such as life expectancy, total fertility rate, and so on, without using special variable names. After the user has defined all the exogenous variables their values are written to a special table. During simulation, at each step all scenario variables are assigned the values in accordance with their definition before calling the demographic model.

Each new initialization of the model overrides the previously set scenario.

2.3.4. *Simulation Module*

The simulation module provides the following functions at each step of simulation: for the current time t it assigns values for scenario variables, calls the demographic model, and controls simulation.

A single simulation step h coincides with the length of the age cohort. The simulation interval may consist of one or more simulation steps. Simulation can be performed for several time intervals. After finishing the simulation of one time interval the user may define the final time for the next time interval.

2.3.5. *Data Representation Module*

The data representation module presents data either at each simulation step (by writing data to the results file), or for the current time after finishing (output is to the terminal's screen). Information is presented in tables and graphics.

3. USING THE DIALOG SYSTEM

3.1. How to Begin Work

To start working with the DIALOG system the user should first have some knowledge about multistate demographic models (a short description was presented above), and second, he should prepare an ASCII file with initial data according to the rules described in Appendix A.1 and A.2.

The current version of the DIALOG system is oriented to use on microcomputers compatible with an IBM PC XT/AT. Computers should have at least 512K RAM memory, a hard disk and coprocessor i8087/i80287, and DOS version 2.1 or above.

The CONFIG.SYS file should include the following line

DEVICE = *<path>ANSI.SYS*

where ANSI.SYS is a device driver provided with DOS, and <path> specifies the full pathname of the file ANSI.SYS.

To use graphic representation the user should have a graphical terminal and an IBM Color Graphics Adapter (CGA) or Enhanced Graphics Adapter (EGA). To print pictures from the computer screen, the user should enter GRAPHICS (using the original DOS diskette) before beginning to work with the DIALOG system.

The user can restore files from the floppy disks using the DOS command RESTORE (to make a backup DOS 3.1 was used):

RESTORE a: c:\DIALOG*.* /s

After execution of this command the new directory DIALOG will contain the following files:

1) DIAL.EXE – executable file

2) COMSPAT.FOR – common block with models variables

3) DIALOGE.DIA – file used for DIALOG in English

4) COUNTRY.DAT – example of initial data file.

The user starts the program by typing DIAL at the DOS prompt. The rest of the DIALOG with the system is continued in the request and answer mode. The system may require the user to enter some information (file name, variable name), or to choose some option from the MENU list. To return to the previous state of the DIALOG the user should enter Ctrl-Z by simultaneously pressing the CONTROL key and Z (with some keyboards the same result may be achieved by pressing <F6> key) and then <ENTER>). In most cases, to return to the previous state of the DIALOG the user can also choose option 0 from the MENU. At any time the user can exit from the system by entering STOP.

After typing DIAL the following MENU appears:

0 – exit
1 – model initialization
2 – modelling
3 – scenario setting
4 – results presentation

The user specifies which of the modules is to have control by entering the subsequent number and the respective module controls into the system.

3.2. Systems Initialization

The initialization procedure is necessary in the following cases:

1) Before starting work with a model.
2) Before a new start of the modeling procedure (starting a new scenario, for example).
3) When choosing a new type of model (medical demographic or demographic) or a new model (another country or region) without exiting the system.

During model initialization the growth matrix $G(t_0)$ for the first simulation step is calculated using the annual rates from the initialization input file. The initial values of gross rates are calculated according to initial patterns of fertility, mortality, and migration.

3.3. Scenario Setting

One of the most important features of the system is the flexible and convenient options for setting the control (scenario) variables.

The scenario variables can be set as a table in the interactive mode in which case the values of the scenario variables correspond to fixed time points, or they can be set in a previously prepared file.

Any model parameter or any exogenous variable can be chosen as a scenario variable, since the scenario determines the number of given variables. During the modeling procedure the values of the scenario variables are calculated in the scenario module, for each time period based on information determined by the user. If for some time point the variable values are missing, then linear interpolation is performed using neighboring values. If a time point is outside the interval for which the scenario variable was determined, then the variable value will stay constant and equal to the closest determined value.

If the scenario is not set for a variable, its value is specified as a default value (calculated from initial data). It remains constant during the process of modeling.

Variables defined as exogenous are assigned values from a scenario table at each simulation step starting from the second one. At the first step the growth matrix $G(t_0)$ is calculated during the model's initialization.

To set the scenario the user enters a number from the following menu:

```
0  -  exit
1  -  fertility (grr)
2  -  life expectancy (exln1)
```

3 – transitions and morbidity (gmr)
4 – variable by name
5 – save scenario to the file
6 – read scenario from the file

3.3.1. *Interactive Scenario Setting*

Choosing options 1, 2, 3, or 4 allows the user to set a scenario interactively.

To set the scenario for life expectancy using option 2, the user should, in advance, define the mode of calculating the age-specific mortality as a function of life expectancy. This mode is determined by a controlling parameter which is defined in the file with initial data (see Appendix A.1). In the system there are two modes to change age-specific mortality rates:

a) equal change of mortality rates in all age groups

$$m_1(x) = m_0(x) + w \quad ;$$

b) proportional change of rates in all age groups

$$m_1(x) = m_0(x)(1. + w) \quad ;$$

where $m_0(x)$ are age-specific mortality rates corresponding to life expectancy e_0; $m_1(x)$ are mortality rates that correspond to new life expectancy e_1 determined by the scenario; and w is estimated numerically in a special module as a function of difference $e_1 - e_0$. (The value of $e_1 - e_0$ should not exceed 4–5 years on one simulation step.)

Scenarios on fertility and mortality are determined for one or several states (regions). For setting transition scenarios the system asks for the state of origin and the state of destination.

The gross fertility (GRRN), and migration (GMRN) rates are taken as scenario variables. In modes 1 and 3 when GRRN and GMRN are set to non-negative values, only the area under the age-specific rates of fertility and migration (RATF and RATM) is changeable; the shape of the age-specific rates is unchangeable. When GRRN and GMRN are set to negative values, they act as switches, which is discussed below.

Using option 4 the system provides an opportunity for scenario setting according to the variable name. For this purpose, the user should have a list of the variable names (identifiers) (see Appendix A.3). For instance, suppose the user wants to change the values of the gross fertility rate for the first region. In Appendix A.3 he will find that the scenario variable for the gross fertility rate for region one is GRRN(1). Then, being in mode 4, he should type the name GRRN(1) and, following the DIALOG, type the desired values.

In option 4 there is also the opportunity to set age-specific fertility and transition rates. After specifying the respective identifier the user can determine the scenario value of the coefficient for given states and age groups. If at the same time a scenario variable for a *gross* rate, say, fertility GRRN for the same state, is set positive ($GRRN > 0$), then two changes occur: first, the form of the curve will change to accommodate the new age-specific rates, and second, the rates will be changed in a way that the area under the age-specific fertility curve will equal the new value of GRRN. If GRRN is negative, then it will no longer be a gross reproduction rate, but act as a switch, that is, the age-specific rates will determine the shape of the curve and the area under the curve.

The scenario variable can have up to 10 values. The time and the values are specified as follows:

time	value
T_1	Y_1
T_2	Y_2
T_3	Y_3

To finish entering data the user should enter Ctrl-Z (simultaneously pressing the CONTROL key and Z (on some keyboards the same result may be achieved by pressing the <F6> key) and then <ENTER>. See also Appendix A.5 for an example of a scenario setting.

Each new setting of a scenario for a particular variable overrides the previously set scenario for the same variable.

3.3.2. *Saving the Scenario*

Creation of a scenario can be time-consuming, especially if the number of control variables is high. Since one sometimes needs to work with the same scenario again, the scenario can be saved using option 5 from the scenario menu. The user should provide the scenario with a name and the information about the scenario will be written to a disk file.

3.3.3. *Reading the Scenario from the File*

To work with a scenario that has already been created by the user or has been saved from the DIALOG system, there is the opportunity to read the scenario from a file. For this purpose the user chooses option 6 from the scenario menu and types the name of the file where the scenario is stored. The file with the scenario is prepared according to the following rules:

1) the string starting with character 'C' is read as a comment;
2) the scenario variable is defined as the following string:

$$<variablename> = TABLE\ (t_1,y_1,t_2,y_2,\ldots,t_n,y_n)$$

   where $t_1,t_2,\ldots,t_n$ are time points, $y_1,y_2,\ldots,y_n$ are values of the scenario variable. It is important that specified values should be real numbers (contain a decimal point).
3) the string can be transferred to any place;
4) multiple scenario variables can be defined in one file;
5) only one scenario can be defined in the file.

After the scenario has been read, the user can deal with it as if it had just been created: he can add the new variable values, change the values of the old variables, etc.

3.4. Modeling

The simulation is performed in the modeling block. Modeling can be done in several time intervals. When the system asks "ENTER TIME SPAN", the last year of modeling should be specified by the user. If the modeling procedure is started after the model initialization, then the time from the file used for initialization is taken as the initial time for the modeling interval. The modeling step coincides with the width of the age group. The modeling interval can consist of one or several steps.

During the modeling procedure the user has the following opportunities:

1) After completing the modeling process for the first time interval, one can continue modeling. For this purpose one should specify the end point of the new time interval. The end point of the previous time interval is considered as the initial point for a new interval of modeling.
2) One can review the results of modeling at the screen after each specified modeling interval (dumping).

3) One can write results to the specified file after each step of the modeling process.

4) One can change the values of the scenario variables and introduce new scenario variables.

5) One can stop modeling and start modeling again with a new model without exiting the DIALOG system.

3.5. Representation of Modeling Results

The data representation module presents the following menu:

0 – exit
1 – printing data during run
2 – current data presentation

A representation module realizes the output of the results produced during the process of modeling (printing data during the run), as well as after completion of the modeling interval (current data presentation).

If option 1 is chosen, information represented in the form of tables is written into the file RESULTS after each step (not the modeling interval) during modeling process. To present information for the initial year as well, option 1 should be specified before model initialization. It is possible to choose which tables should be written to RESULTS. The list of tables is given in the following menu:

0 – exit
1 – age-specific rates
2 – growth matrix
3 – gross rates
4 – expectancies
5 – births, deaths
6 – transition flows
7 – indices of labor activity
8 – summary table
9 – population distribution
10 – percentage population distribution

The growth matrix (option 2) could be printed for a one-sex model. In case of a two-sex model only it will be correctly printed only for the second sex. The other options are self-explanatory. When using option 2 in the data representation menu, the results will appear on the screen. The results will only show the current period data. In this case the choice could be made from the following menu:

0 – exit
1 – age-specific rates
2 – gross rates

3 – expectancies
4 – births, deaths
5 – transition flows
6 – indices of labor activity
7 – summary table
8 – population distribution
9 – percentage population distribution
10 – graphics
11 – variable by name

When working with the model it is sometimes necessary to look at the current value of variables, the representation of which is not provided with tables. Option 11 is used to display the current value of each variable used in the model. To use this mode one chooses identifiers of the variables, which are given in Appendix A.3.

For each request information appears at the user's terminal for the current year of simulation and could be printed simply using the PrtSc (Print Screen) button. The same is valid for option 10 in the last menu which defines graphical presentation of information from the following menu:

0 – exit
1 – population histogram %
2 – age-specific fertility per 1000
3 – age-specific mortality per 1000
4 – pie chart by state
5 – population pyramid

In order to achieve more independence from the available hardware environment (the size of the terminal screen or the width of the printer paper), a special table generator was developed. This helps to easily adapt the representation form for the results of modeling to the particular type and configuration of the computer. It also allows deletion of constraints related to the number of states and age groups in the model, which usually produce problems in representation of the results.

4. REALIZATION

The DIALOG system was created for use on a PC XT/AT or compatible. It was written in FORTRAN 77 and an RM Fortran compiler was used. To create graphics the NOLIMIT library was used. However, at each stage of its design, efforts were developed to make the system as machine independent as possible. The same system was also used on the mainframe VAX 11/780.

Forthcoming versions of the system will allow the graphical setting of scenarios and exogenous variables, including age-specific rates. There will be an option to use model life tables and model fertility schedules for scenario setting. Results presentation options will also be increased. A new version will allow use of the system on PCs without coprocessors.

APPENDICES

A.1 Input Data File

Before using the system the user should prepare the initial data file. The data are read in free format. The name of the input file is requested by the system at the beginning of the DIALOG. This file contains the following information (each paragraph should start with a new line):

1. The title of the data file – one record containing any user defined information about this data file (the form of the note is arbitrary).

2. The model's parameters.

   NA – the number of the age groups, NA $\leq$ 20 in this version of the model.

   NSEXES – the number of sexes (1 or 2).

   NR – the number of regions or states, NR $\leq$ 15 in this version.

   NY – the size of the age cohort, usually NY = 1 or $NY = 5$.

   NU – the number of strings in the title of the system.

   INIT – the initial year of modeling.

   NG – switch of the model type. NG = 0 corresponds to the model type in which a newborn appears in the same state as the parents. In this case the C matrix is identical. When the switch NG = 1 this corresponds to the model where all newborns appear in the same state.

   SEXRAT – the ratio of female to male newborns. (In case of one sex user should set SEXRAT = 1.)

   NEXL – parameter which defines the mode of change of age-specific mortality as a function of life expectancy during scenario setting. NEXL = 1 corresponds to the case when $m_1(x) = m_0(x) + w$; NEXL = 2 corresponds to $m_1(x) = m_0(x)(1. + w)$; where $m_0(x)$ is the age-specific mortality rates corresponding to life expectancy e_0;

$m_1(x)$ is the mortality rates that correspond to new life expectancy e_1 determined by scenario; w is estimated numerically in a special module as a function of difference $e_1 - e_0$. (The value of $e_1 - e_0$ should not exceed 4–5 years on one simulating step.)

NAGE(NA+1) – integer array of age intervals of ages for which initial values are given (on a separate line).

NDAT – indicates the initial data type. NDAT = 1 corresponds to the case when fertility, mortality, and migration are given in absolute numbers. If NDAT = 2 the coefficients of fertility, mortality, and migration are given.

3. The title of the system. These are the records that will be seen on the display after all data from the file have been read. The number of records is defined by parameter NU.

4. The name of the country or the system of the states (groups).

5. Data. For each state (group):

   the data for state 1,
   the data for state 2,
   the data for state NR should be given.

The data for each state includes the following information:

a) The name of the region or state. This may contain up to eight symbols.

b) If NSEXES = 1:

   – The population's age structure at time INIT.

   – Fertility. The age-specific pattern is specified. When NDAT = 1 the births are given in absolute numbers. When NDAT = 2 the birth rates are specified. (In this case rates should be defined as the ratio of the number of newborns for women in a specific age cohort to the total population of the same age cohort.)

   – Mortality. The age-specific mortality is specified. When NDAT = 1 it is given in absolute numbers. If NDAT = 2 the mortality rates are specified.

   – Transitions, or migrations, from one given state to another. When birth and death are specified in absolute numbers, the age structure of transitions should also be specified in terms of absolute numbers. If birth and death are specified by rates, then age-dependent transition rates should also be specified.

c) If NSEXES = 2:

   – Data for female population is the same as in b) except that fertility rates should be defined as the ratio of the number of newborns for women in a specific age cohort to the female population of the same age cohort.

– Data for male population are the same as for female population except that fertility rates are excluded.

A.2. Example of the Initial Data File

Lines which start with the symbol * are comments given to facilitate understanding of the structure of the initial data file. In real initial data files, all such comments are prohibited. Numbers should be separated by one or more spaces.

```
* 1. The title of the data file.

     country 1:    -    2 regions    -    two sexes: male, female

   * 2. Model's parameters.
   * NA NSEXES NR NY NU INIT NG NDAT SEXRAT  NEXL

      16   2    2   5  4 1975 0   1    1.000   2

   * 3. Age intervals for which data are given

     0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80

   * 4. Title in the system.

        MULTIREGIONAL POPULATION SIMULATION
                     for
              hypothetical country
               starting year 1975

   * 5. Name of the country or the states of the system.

     COUNTRY

   * 6. Data for the first state.

        *a) Name of the state.

     NORTH

      *   Data for female.

        *b)Population age structure.

      130449.   164503.   204645.   189898.   188590.   116643.  124046.
      168270.   140364.   117915.    96209.    66299.    95490.   100051.
       84113.    97898.

        *c)Fertility.

        0.        0.        1.     9100.   23550.    9452.    3896.     987.     396.
        3.        0.        0.        0.       0.       0.       0.

        *d)Mortality.

       519.      76.       51.      167.     196.     119.     172.     313.     423.
       501.     676.      651.     1765.    3033.    4544.   12783.
```

*e)Transitions from 1 to 1.

```
  0.       0.       0.       0.       0.       0.       0.       0.       0.
  0.       0.       0.       0.       0.       0.       0.       0.       0.
```

*f)Transitions from 1 to 2.

```
337.     188.     130.     270.     876.     421.     228.     150.      75.
 51.      31.      11.      34.      35.      22.      38.
```

*   Data for male.

*g)Population age structure.

```
130449.    164503.    204645.    189898.    188590.    116643.    124046.
168270.    140364.    117915.     96209.     66299.     95490.    100051.
 84113.     97898.
```

*h)Mortality.

```
919.      96.      81.     267.     296.     219.     272.     413.     523.
601.     776.     851.    1965.    4033.    5544.   15783.
```

*i)Transition from 1 to 1.

```
  0.       0.       0.       0.       0.       0.       0.       0.       0.
  0.       0.       0.       0.       0.       0.       0.       0.       0.
```

*j)Transition from 1 to 2.

```
337.     188.     130.     270.     876.     421.     228.     150.      75.
 51.      31.      11.      34.      35.      22.      38.
```

* 7. Data for the second state.

```
SOUTH
  62653.     79009.     92259.     73156.     72651.     66688.     90215.
 100196.     76735.     62934.     48528.     38995.     56860.     64747.
  52249.     60299.
   0.       0.       2.    3428.    9451.    5840.    3426.     905.     189.
   7.       0.       0.       0.       0.       0.       0.
 221.      30.      30.      53.      69.      62.     112.     164.     212.
 287.     357.     453.    1171.    2194.    3005.    8338.
 108.      83.      99.     105.     216.     132.      87.      60.      36.
  19.      11.       4.      18.      14.       9.      20.
   0.       0.       0.       0.       0.       0.       0.       0.       0.
   0.       0.       0.       0.       0.       0.       0.
  62653.     79009.     92259.     73156.     72651.     66688.     90215.
 100196.     76735.     62934.     48528.     38995.     56860.     64747.
  52249.     60299.
 221.      30.      30.      53.      69.      62.     112.     164.     212.
 287.     357.     453.    1171.    2194.    3005.    8338.
 108.      83.      99.     105.     216.     132.      87.      60.      36.
  19.      11.       4.      18.      14.       9.      20.
   0.       0.       0.       0.       0.       0.       0.       0.       0.
   0.       0.       0.       0.       0.       0.       0.
```

A.3. List of Variables for the Multistate Demographic Model

In the following, indices in brackets denote:

X - age,

I,J - states

IS - index of sex (1 - female, 2 - male, 3 - total)

Population

POPR(X,I,IS) - population by age, state, and sex

PDTOT(X,IS) - population by age and sex

PTOTR(I,IS) - population by state and sex

PRCPDR(I,IS) - percent of population distribution by state and sex

POPTOT(IS) - population by sex

Fertility

RATF(X,I) - fertility rate by age of mother and state

BIRTH(X,I) - number of newborn by age of mother and by state

BR(I,IS) - fertility rate by state and sex of newborn

B(I,IS) - number of newborn by state and sex

BRTOT(IS) - total fertility rate by sex of newborn

BTOT(IS) - total number of newborn by sex

GRR(I) - gross fertility rate by state for one sex model and total fertility rate by state for two sex model

GRRN(I) - scenario gross fertility rate by state for one sex model and total fertility rate by state for two sex model if > 0 or a switch if <0

Mortality

RATD(X,I,IS) - mortality rate by age, state, and sex

DR(I,IS) - mortality rate by state and sex

D(I,IS) - number of dead by state and sex

DRTOT(IS) - mortality rate by sex

DEATH(X,I,IS) - number of dead by age, state, and sex

DTOT(IS) - total number of dead by sex
GDR(I,IS) - gross mortality rate by state and sex
EXL(I,J,IS) - average time spent in state I for those who were born in state J by sex
EXLT(I,IS) - time that an arbitrary individual spent in state I by sex
EXL1(I,IS) - average life expectancy in state I in absence of internal migration.
EXLN1(I,IS) - scenario variable for average life expectancy in state I in absence of internal migration.

Transitions between States

RATM(X,I,J,IS) - transition rate from state J to state I by age and sex
GMR(I,J,IS) - gross transition rate from J to I by sex
GMRN(I,J,IS) - scenario gross transition rate from J to I by sex or a switch if <0
FMGR(I,IS) - number of transitions from J to I by sex
FMGRA(I,IS) - number of arrivals by state and sex
FMGRD(I,IS) - number of departures by state and sex
FMGRR(I,J,IS) - transition rate from state J to state I by sex
FMGRDR(I,IS) - rate of departures by state and sex
FMGRAR(I,IS) - rate of arrivals by state and sex
DFMGR(I,IS) - saldo of transitions by state and sex
FMGRT(IS) - total number of transitions by sex

Indices of Labor Force Participation

ACT(I,IS) - number of individuals of active age (from 15 to 64) by state and sex
RETM(I,IS) - number of elderly by state (65 years and over) and sex
CHLD(I,IS) - number of children up to 15 by state and sex
PASS(I,IS) - number of children up to 15 and elderly of 65 years and over by state and sex
ACTR(I,IS) - proportion of individuals of active age by state and sex

PRPA(I,IS) - dependency ratio (up to 15 and 65+)/(15-65) by state and sex
CHLDR(I,IS) - proportion of children by state and sex
RETMR(I,IS) - proportion of elderly by state and sex
ACTTOT(IS) - total number of individuals of active age by sex
CHLT(IS) - total number of children by sex
RETMT(IS) - total number of elderly by sex
PASST(IS) - total number of dependents by sex
ACTTR(IS) - proportion of individuals of active age by sex
CHLDTR(IS) - proportion of children by sex
RETMTR(IS) - proportion of elderly by sex
PRPTOT(IS) - dependency ratio by sex

Other Indices

AGEM(I,IS) - average age by state and sex
AGEMT(IS) - average age by sex
RNINC(I,IS) - natural growth rate by state and sex
RMINC(I,IS) - growth rate due to transitions by state and sex
RNINCT(IS) - total growth rate by sex

A.4. Examples of Results Presentation

summary table
variant: female year: 1975

	pop. thous.	% tot	mean age	rates of birth	natural death	natural gr. rat.	transit. out	transit. in	transit. net rate
NORTH	2085.383	65.505	34.560	0.011	0.012	-0.001	0.001	0.000	-0.001 -->
SOUTH	1098.174	34.495	37.122	0.011	0.015	-0.005	0.001	0.003	0.002 -->
COUNTRY	3183.557	100.000	35.444	0.011	0.013	-0.002	0.001	0.001	0.000 -->

summary table
variant: female year: 1975

	ext.mig. net	tot. gr. rate
NORTH	0.000	-0.002
SOUTH	0.000	-0.003
COUNTRY	0.000	-0.002

summary table
variant: female year: 1990

	pop. thous.	% tot	mean age	rates of birth	natural death	natural gr. rat.	transit. out	transit. in	transit. net rate
NORTH	2060.495	65.646	36.434	0.012	0.012	0.000	0.001	0.000	-0.001 -->
SOUTH	1078.286	34.354	37.637	0.012	0.013	-0.001	0.001	0.003	0.002 -->
COUNTRY	3138.781	100.000	36.848	0.012	0.012	0.000	0.001	0.001	0.000 -->

summary table
variant: female year: 1990

	ext.mig. net	tot. gr. rate
NORTH	0.000	-0.001
SOUTH	0.000	0.000
COUNTRY	0.000	0.000

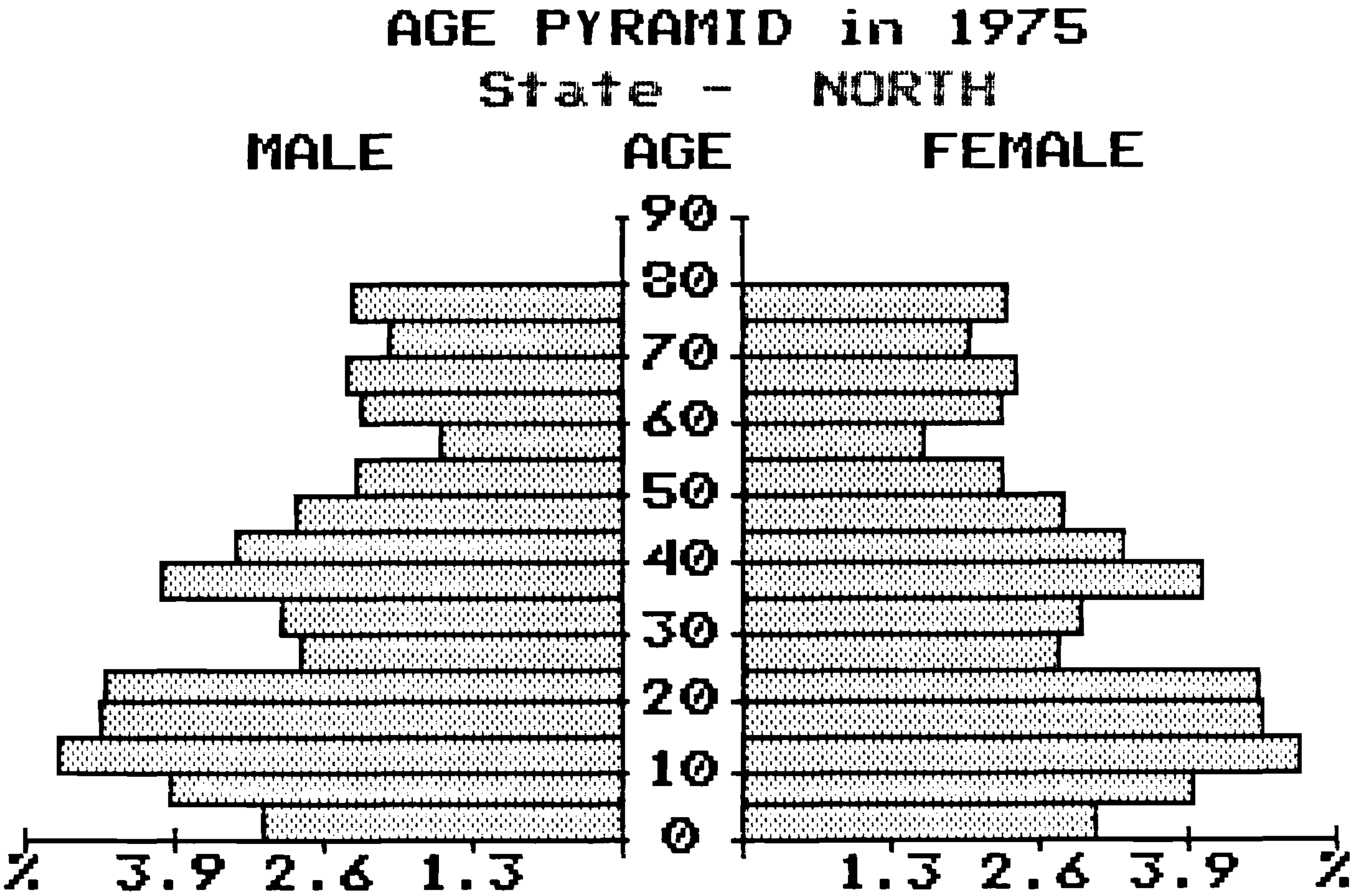
population distributions by age and regions
variant: female year: 1975

	NORTH	SOUTH	COUNTRY
0	130449.	62653.	193102.
5	164503.	79009.	243512.
10	204645.	92259.	296904.
15	189898.	73156.	263054.
20	188590.	72651.	261241.
25	116643.	66688.	183331.
30	124046.	90215.	214261.
35	168270.	100196.	268466.
40	140364.	76735.	217099.
45	117915.	62934.	180849.
50	96209.	48528.	144737.
55	66299.	38995.	105294.
60	95490.	56860.	152350.
65	100051.	64747.	164798.
70	84113.	52249.	136362.
75	97898.	60299.	158197.
total	2085383.	1098174.	3183557.

life expectancies
variant: female year: 1975

	NORTH	SOUTH
NORTH	67.010	3.068
SOUTH	4.180	67.780
total	71.189	70.848

The column denotes the state of birth and the row the average time spent by an individual in a particular state. Thus a woman born in the state NORTH will spend, on average, 67.01 years in the state NORTH, and 4.18 years in the state SOUTH. Her total life expectancy will constitute 71.189 years.


AGE PYRAMID in 1975
State - NORTH
MALE
AGE
FEMALE
90
80
70
60
50
40
30
20
10
0
% 3.9 2.6 1.3
1.3 2.6 3.9 %

A.5. Example of Scenario Setting

```
                  M A I N   M E N U

                0 - exit
                1 - model initialization
                2 - modelling
                3 - scenario setting
                4 - results presentation

                  ====>3

       S C E N A R I O   S E T T I N G

                0 - exit
                1 - fertility (grr)
                2 - life expectancy (exln1)
                3 - transitions and morbidity (gmr)
                4 - varible by name
                5 - save scenario to the file
                6 - read scenario from the file

                  ====>2

         SCENARIO SETTING FOR LIFE EXPECTANCY


1- NORTH        2- SOUTH

Enter state number : 1
Enter number (1 - female; 2 - male): 1

the default value is 71.2456
     1 - do not change default value
     2 - set new scenario

enter number: 2

enter new scenario in format:
  time       value  (to end press ^Z and then <return>)
1975    71.2456
1980    73.
1985    73.5
^Z
```

REFERENCES

Gelovani, V.A. (1980) Man-Machine Modeling System of the Global Development Processes, *System Studies Annual*, 155–173, Moscow: VNIISI (in Russian).

Just, P. (1983) *Two Program Packages for Deriving Multistate-Multiregion Life Tables and Two-Sex (Female Dominant) Population Projections*, Working Paper WP–83–10, Laxenburg, Austria: International Institute for Applied Systems Analysis.

Just, P. and Liaw, L. (1983) *An Improved Method of Computing Multistate Survivorship Proportions for the Terminal Age Groups*. WP–83–65. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Keyfitz, N. (1980) *Multidimensionality in Population Analysis*, Research Report RR–80–33, Laxenburg, Austria: International Institute for Applied Systems Analysis.

Ramachandran, R. (1980) *Simulation of Multiregional Population Change: An Application to Belgium*, Brussels: Brije Universiteit.

Rogers, A. (1975) *Introduction to Multiregional Mathematical Demography*, New York: John Wiley.

Rogers, A. (Ed.) (1980) *Essays in Multistate Mathematical Demography*, Research Report RR–80–10, Laxenburg, Austria: International Institute for Applied Systems Analysis.

Rogers, A. and Ledent, J. (1976) Increment-Decrement Life Tables: A Comment. *Demography* 13:287–290.

Scherbov, S.Y. and Grechucha, V.A. (1984) *Modeling of Multiregional Demographic Development*, Materials of the Third All-Union School Seminar on Systems Analysis of Social-Economical Problems of Regional Development (Novosibirsk) (in Russian).

Scherbov, S. and Usbeck, H. (1983) *Simulation of Multiregional Population Change: An Application to the German Democratic Republic*, Working Paper WP–83–6, Laxenburg, Austria: International Institute for Applied Systems Analysis.

Scherbov, S., Yashin, A., and Grechucha, V. (1986) *Dialog System for Modeling Multidimensional Demographic Processes*, Working Paper WP–86–29, Laxenburg, Austria: International Institute for Applied Systems Analysis.

Strong, M.A. (1987) Software For Demographic Research, *Population Index* 53(2):183–199.

Willekens, F. (1979) *Computer Program for Increment-Decrement (Multistate) Life Table Analysis: A User's Manual to Lifeindec*, Working Paper WP–79–102, Laxenburg, Austria: International Institute for Applied Systems Analysis.

Willekens, F. and Rogers, A. (1978) *Spatial Population Analysis: Methods and Computer Programs*, Research Report RR–78–18, Laxenburg, Austria: International Institute for Applied Systems Analysis.

Yashin, A. (1977) *Methodological Aspects of Modeling and Decision Making in the Health Care System*, Collaborative Paper CP–77–4, Laxenburg, Austria: International Institute for Applied Systems Analysis.

Recent Working Papers Produced in IIASA's Population Program

Copies may be obtained at a cost of US$ 5.00 each from IIASA's Publications Department.

WP-86-01, *Exploratory Analysis of the Umea Data at IIASA* by Arno Kitts. January 1986.

WP-86-02, *Increasing Returns to Scale in Heterogeneous Populations* by Robin Cowan. January 1986.

WP-86-03, *Notes on the Effects of Cohort Size on Intergenerational Transfer* by Robin Cowan. January 1986.

WP-86-06, *A Simulation Study of the Conditional Gaussian Diffusion Process Model of Survival Analysis* by Fernando Rajulton and Anatoli Yashin. February 1986.

WP-86-09, *The Two Demographic Transitions of Finland* by Wolfgang Lutz. February 1986.

WP-86-19, *The Division of Labor for Society's Reproduction: On the Concentration of Childbearing and Rearing in Austria* by Wolfgang Lutz and James Vaupel. April 1986.

WP-86-29, *Dialog System for Modeling Multidimensional Demographic Processes* by S. Scherbov, A. Yashin, and V. Grechucha. June 1986.

WP-86-34, *Culture, Religion and Fertility: A Global View* by W. Lutz. July 1986.

WP-86-37, *The LEXIS Computer Program for Creating Shaded Contour Maps of Demographic Surfaces* by B. Gambill, J. Vaupel, and A. Yashin. August 1986.

WP-86-53, *Population Models Analysis Program (POPMAN)* by A. Lewandowska. October 1986.

WP-86-59, *Cancer Rates over Age, Time and Place: Insights from Stochastic Models of Heterogeneous Populations* by J. Vaupel and A. Yashin. October 1986.

WP-86-60, *Heterogeneity in Composite Link Models* by C. Vanderhoeft. October 1986.

WP-86-63, *Derivative-free Gauss-Newton-like Algorithm for Parameter Estimation* by S. Scherbov and V. Golubkov. November 1986.

WP-86-69, *Modelling Kinship with LISP — A Two-Sex Model of Kin-Counts* by J. Bartlema and L. Winkelbauer. November 1986.

WP-86-74, *Computation of Multi-State Models using GAUSS, A Matrix Based Programming Language* by A. Foster and N. Keyfitz. December 1986.

WP-86-76, *Structural Minimization of Risk on Estimation of Heterogeneity Distributions* by A. Michalski and A. Yashin. December 1986.

WP-86-77, *A Note on Random Intensities and Conditional Survival Functions* by A. Yashin and E. Arjas. December 1986.

WP-86-78, *Cause Specific Mortality in Japan: Contour Maps Approach* by B. Gambill, A. Yashin, J. Vaupel, Z. Nanjo, and T. Shigematsu. December 1986.

WP-86-81, *Kinship and Family Support in Aging Societies* by D. Wolf. December 1986.

WP-87-12, *Comparative Anatomy of Fertility Trends: The Aging of the Baby Boom* by W. Lutz and A. Yashin. January 1987.

WP-87-13, *Using the INLOGIT Program to Interpret and Present the Results of Logistic Regressions* by D. Wolf. January 1987.

WP-87-46, *The Multistate Life Table with Duration-Dependence* by D. Wolf. May 1987.

WP-87-51, *The Concentration of Reproduction: A Global Perspective* by W. Lutz. June 1987.

WP-87-58, *A Simple Model for the Statistical Analysis of Large Arrays of Mortality Data: Rectangular vs. Diagonal Structure* by J. Wilmoth and G. Caselli. June 1987.

WP-87-59, *Sibling Dependences in Branching Populations* by P. Broberg. June 1987.

WP-87-87, *The Living Arrangements and Familial Contacts of the Elderly in Japan* by K. Hirosima. September 1987.

WP-87-92, *The Demographic Discontinuity of the 1940s* by N. Keyfitz. September 1987.

WP-87-104, *A Random-Effects Logit Model for Panel Data* by D. Wolf. October 1987.

WP-87-116, *Some Demographic Aspects of Aging in the German Democratic Republic* by T. Büttner, W. Lutz, and W. Speigner. November 1987.

WP-88-10, *On the Concentration of Childbearing in China, 1955-1981* by W. Lutz. February 1988.

WP-88-13, *Beyond "The Average American Family": U.S. Cohort Parity Distributions and Fertility Concentration* by M. King and W. Lutz. March 1988.

WP-88-23, *Understanding Medical and Demographic Trends with MEDDAS* by M. Rusnak and S. Scherbov. April 1988.

WP-88-32, *Kinship Patterns and Household Composition of the Elderly: Hungarian Women, 1984* by D. Wolf. April 1988.