

Household and education projections by means of a microsimulation model

J.H.M. Nelissen

Two modules of the microsimulation model NEDYMAS are presented: the demographic module and the education module. After a description of the ins and outs of microsimulation, both modules are presented. Next both modules are used to reconstruct the past and to generate the future Dutch household structure and the level of education of the Dutch population. Standard deviations are also shown for the year 1981.

Keywords: Microsimulation; Household projection; Education projection

Household and education projections are becoming more and more important. This holds not only for the projections alone, but also – or perhaps especially – for their distributional aspects. One reason for this is the increased emphasis on micro-oriented economic research, which has made it clear that household processes have a large influence on economic processes. Klevmarken [28], for example, shows that in Western countries a number of demographic processes – like marriage, divorce, leaving the parental home or childbirth – have more influence on the household's well-being than an economic process like becoming unemployed. All demographic decisions have economic implications. Marriage, for example, leads to advantages in the sphere of a division of labour in both household work and a paid job. What is more, in a household of two (or more) people, economies of scale arise.

This effect has been identified by the Netherlands Central Bureau of Statistics (NCBS). The NCBS introduced so-called dynamic income statistics to measure these kind of effects. A description can be found in van de Stadt [105]. It is not usually normal to take demographic changes into account – age distribution excluded – even in long-term projections.

Another reason for increasing interest in household dynamics can be found in the ever growing need for

more detailed information on the socioeconomic characteristics of the population, not simply from the point of view of planning but also as a tool for evaluation. There is an increasing demand for population projections which has strongly grown, not only from their traditional users but also from elsewhere. We can think of the effect of household composition on the labour supply of married women (see Kapteyn and Woittiez [23]), of its effect on consumer expenditures (see Alessie *et al* [1]), of the role of household composition in the determination of the level of various social security benefits (see eg Nelissen and Vossen [94]), or of its impact on the discussion of reforms in the social security system and the tax system.¹

Moreover, more information is needed on distributional aspects. We need knowledge about education, income, transition probabilities between various (socioeconomic) positions and so on, if we want to determine, for example, the effects of changes in the social security system. As a consequence, problems arise if we want to use the traditional macro- or meso-oriented models, because these types of models are not able – or are able only in a very limited way – to provide us with adequate information about distributional issues.² Such information is one of the great advantages of microsimulation, a method which has not been fully used in the Netherlands until now.³

J.H.M. Nelissen is with the Department of Sociology, Tilburg University, 5129, PO Box 90153, 5000 LE Tilburg, The Netherlands.

Final manuscript received 5 December 1990.

¹ See the discussion in the Tweede Kamer (Dutch Parliament) with respect to the Oort proposals in the Netherlands.

² See van de Stadt, Huigen and Zeelenberg [106].

³ The microsimulation model of the Social and Cultural Planning Agency [104] is a static one.

Table 1. Characteristics of three types of models.

	Micro	Meso	Macro
Possibility of incorporating:			
Direct effects	Good	Good	Average
Behavioural effects	Good	Good	Average
Cyclical effects	Bad	Good	Good
Distribution	Good	Average	Bad
Model dynamics	Bad	Bad	Good

Source: van de Stadt *et al* [106].

This is not really surprising in view of its large costs. But these costs are diminishing as a consequence of decreasing computer costs. Microsimulation enables us to model demographic information (like age, marital status, sex and household situation), economic characteristics (like income, labour force participation and non-labour income) and personal attributes (like health, level of education and social class) in a consistent and very flexible way. Micro-units (individuals and households or – in other applications – firms) are the starting point of the process, which produces very detailed information.

This article discusses the microsimulation approach, which is the subject of the next section. An application of microsimulation in the field of household and education dynamics is given in the fifth section. The demographic and education module of a microsimulation model, developed by the author, are first described in the preceding sections. The paper ends with some concluding remarks.

Microsimulation

Table 1 shows some characteristics of three types of models: micro-, meso- and macromodels. It appears that micromodels are the most appropriate instruments for an analysis of the type of problems we have discussed above. With the aid of micromodels, we can incorporate first order effects (direct effects as a

consequence of changes in eg institutional relations) and second order effects (behavioural effects which are a consequence of the reallocation of activities by the micro-units). This method also yields ideal information on the distribution of the effects. However, this type of model cannot directly be used to process so-called third order effects (the consequences of the spill over to other markets, and the cyclical consequences). These can be solved in practice by constructing a (limited) macromodel around the micromodel (see, for example, Orcutt *et al* [98]; Hansen *et al* [15]), although problems with regard to the simultaneity and consistency of the model remain. The model dynamics are also often limited because the structure is recursive. This problem can be alleviated by restricting the period of analysis to a week or a month instead of a year (see, for example, Bennett and Bergmann [5]).

Common meso- and macromodels cannot be disaggregated to provide the information needed, in particular the distribution of effects if knowledge of the other household members is needed. A comparison between the microsimulation approach and other models used for policy analyses can be found in Greenberger, Crenson and Crissey [14], Chapter 4. The microsimulation approach was proposed by Orcutt [96] and [97], as was the first – limited – application (see Orcutt *et al* [101]).

Microsimulation can be described as an analysis of the behaviour of the system under investigation, using characteristics of the micro-units distinguished in the system, whereby, with the aid of a model which reproduces the functioning of the system under study, individuals' characteristics are adjusted each period. So a microsimulation system acts on a microdata base, generally containing a number of thousands of households (and their household members), and this microdata base is adapted – or aged – from period to period. The process is shown in Figure 1.

Assume that the system under study is the social security system in the Netherlands. The micro-units in the microdata base are then individuals, living in

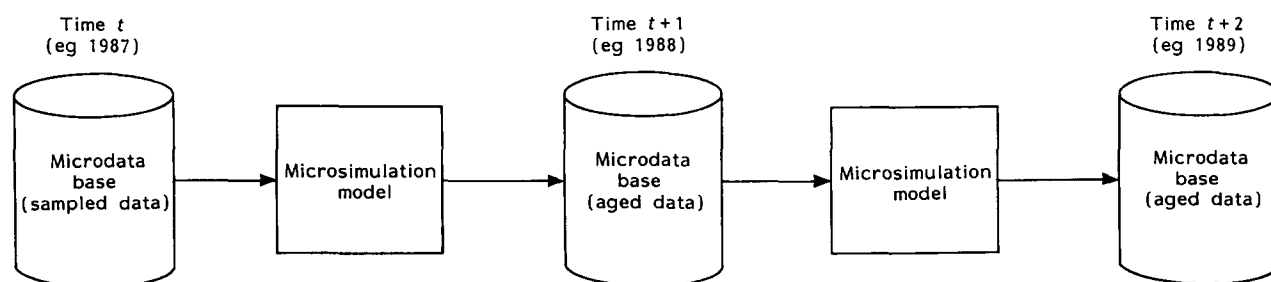


Figure 1. The principle of microsimulation.

Source: Hellwig [18], p 3.

Table 2. A list of individuals and some of their attributes.

Label (1)	Age group (2)	Specific age (3)	Sex (4)	Marital status (5)	Race (6)	Education (7)	Occupation (8)	Part- or full- time job (10)	Number of weeks worked in previous year (11)	Wage earned in previous year (£) (12)	Multiple of wage median (13)
1058.1	4	30	Male	Divorced	White	3	11	1	52	5110.11	1.40
2069.1	2	21	Male	Single	White	2	19	0	0	0.00	0.00
2187.1	3	25	Male	Married	White	4	16	1	52	4842.06	1.27
2187.2	2	23	Female	Married	White	4	3	1	52	3635.18	1.11
2206.1	8	54	Male	Married	White	4	3	1	52	8914.73	1.25
2206.2	9	55	Female	Married	White	2	3	1	52	7567.65	1.05
2206.3	3	28	Female	Single	White	2	21	0	52	0.00	0.00
2206.4	4	33	Male	Single	White	2	16	1	52	4069.80	0.98
2206.5	1	20	Male	Single	White	2	21	0	0	0.00	0.00
2728.1	7	46	Male	Married	White	4	11	1	52	3461.27	0.91
2728.2	6	41	Female	Married	White	2	18	1	52	4225.86	1.24
2728.3	2	21	Female	Single	White	2	6	1	52	3875.89	1.22
3031.1	11	69	Male	Married	White	4	20	0	0	0.00	0.00
3031.2	10	60	Female	Married	White	4	21	0	0	0.00	0.00
3039.1	7	46	Male	Married	White	2	14	1	52	6804.96	1.34
3039.2	6	42	Female	Married	White	3	7	1	52	5142.55	1.28
3039.3	1	12	Female	Single	White	1	0	0	0	0.00	0.00
4092.1	6	44	Male	Divorced	White	3	9	1	52	7495.00	1.33
4107.1	4	34	Female	Divorced	Black	4	20	0	0	0.00	0.00
4107.2	1	7	Male	Single	Black	1	0	0	0	0.00	0.00
4521.1	7	45	Male	Married	White	3	15	1	52	3693.00	0.91
4521.2	6	42	Female	Married	White	3	20	—	—	—	—
4521.2	1	11	Female	Single	White	1	0	—	—	—	—
4521.4	1	7	Female	Single	White	1	0	—	—	—	—

Source: Clarke [9], p 249.

the Netherlands, and emigrants. These individuals are related to other members in their households, if present. The microdata base contains the characteristics of all these micro-units. Characteristics which are recorded are age, marital status, level of education, number of children, economic activity, labour supply, wage income, various social security benefits and payments, year of immigration, year of emigration, and so on. The model is implemented in a computer program, which describes (preferably by applying behavioural assumptions) the changes in the individuals' characteristics and describes the conditions which, for example, determine whether an individual has to pay social security premiums or has to receive social security benefits; in both cases equations determine how much has to be paid or received. Each characteristic has to be able to be fixed – sex, year of birth – or change with time by means of a behavioural model which ages it. The behavioural hypotheses can be formulated in various ways: by tables, regression equations, probability distributions, decision trees, transition probabilities, institutional rules and so on. The required parameters generally have to be estimated from data. These data are often derived from the microdata base or are obtained externally. By applying behavioural assumptions we can determine how the

situation of each of the micro-units changes as a consequence of internal and external factors which influence the system in question. We must realize that the model is 'solved' by simulation. No analytical solution is obtained, so the result is not unique. Other initial values for the random generator (see the next section) will give other solutions. The reliability of the results therefore depends on the number of micro-units distinguished in the microsimulation system.

The essence of microsimulation lies in the program which focuses on multiactor, multilevel and multiprocess synthesis (Caldwell [8]). For this, a bottom up strategy is applied. The state representation of the components of the system of interest forms the basis of each modelling strategy. As the name suggests, the microlevel representation of individuals in terms of social, economic and demographic characteristics, together with other possible relevant spatial and activity attributes, forms the basis of microsimulation (Clarke [9]). The information with respect to individuals, for example, is stored in the form of lists which are updated (list processing). An example of such a list is given in Table 2.

This storage method forms the key to the application of microsimulation. The advantage of list processing as compared with storage via an occupancy matrix,

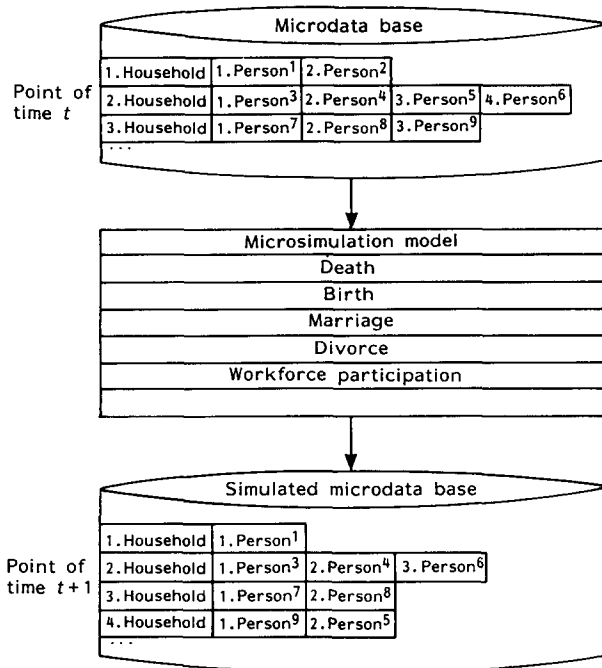


Figure 2. The principle of microsimulation.

for example – the method which is used for macrosimulation, among other applications – is that the storage space needed is relatively very small, with a limited number of individual characteristics. The number of records to store via list processing in microsimulation – which uses a discrete *event* model – increases linearly with the number of micro-units, whereas storage via an occupancy matrix in macrosimulation – which uses a discrete *time* model – is multiplicatively related to the number of states. This is the consequence of working with decision unit probabilities instead of with transition probabilities (see also Hayes [16]). Considering only the demographic part, and using 100 000 individuals, our model NEDYMAS requires 9 900 000 bytes or about 10 megabytes.⁴ To get the same information by macrosimulation would need about 1.6×10^{24} states, and consequently a matrix of about 2.5×10^{48} cells!⁵ Besides, the macrosimulation approach cannot tell us who is whose partner or child and so on, whereas that is no problem in the microsimulation approach. When using more than a very limited number of variables – and certainly in the case of continuous variables, like income – the micro/macro storage ratio will be close to zero.⁶ So, we can state that microsimulation is able to provide information which cannot be obtained even from such a macro-approach.

⁴ The model is called NEDYMAS, the NETHERLANDS' DYnamic Micro-Analytic Simulation model.

⁵ See Nelissen [93].

⁶ See also Keilman [24].

So how is microsimulation applied in practice? As we said above, a microsimulation model comprises a microdata base of micro-units and a micromodel. With the aid of the micromodel a representative continuation of the sample is obtained at the end of each period. This process is called ageing of the population, as given schematically in Figure 2.

To illustrate this, we will take the modelling of mortality. The decision as to whether an individual will or will not undergo a potential transition is simulated with the aid of the Monte Carlo method. In view of this, the conditional probability of an individual undergoing an event has to be given. For example, for a 77-year-old divorced woman the probability of dying was 6.75% in 1968. We then randomly draw a number from the uniform [0, 1] distribution. If this number is smaller than or equal to the probability of dying of 0.0675, the woman is expected to die. If the number is larger than 0.0675, the woman will remain alive. If she dies, we then investigate whether she had children (who have lost their mother). So, decisions taken at the level of an individual can have implications for other individuals.

Microsimulation creates a synthetic database which reflects (developments in) the demographic and economic structure of the population. A stylized example is given in Table 3. Hendricks, Holden and Johnson [21] used this possibility to produce a simulated set of family and earning histories from 1960 to 2000, using the DYNASIM model, in order to assist the staff at the Office of the Secretary of Health, Education and Welfare.

The simulated population can be considered as the realization of a stochastic process, as a random sample from the real population. The simulated and observed population will differ randomly for two reasons. The first is the variance related to the specification of the initial sample population. The second is the variance associated with the sampling approach for selecting the particular members of the population whose characteristics have changed each year of the simulation. This is called Monte Carlo variability (see Orcutt *et al* [98]: Chapter 11).

We can distinguish a number of types of microsimulation models.⁷ First of all we can distinguish standard and empirical simulation models. Standard simulations refer to stylized calculations (eg for specific types of households or persons). Empirical simulations use empirical data, so that we can better determine the number of persons to which some calculations refer and how relevant changes will be. In the following, we limit ourselves to empirical models. A first

⁷ See Social and Cultural Planning Agency ([104], pp 46–47) and Merz ([31], pp 4–9).

Table 3. An example of microsimulation.

1986 (sample data)							1987 (aged data)						
Household 1							Household 1						
	ID	Age	Sex	Marital status	Job	Income	ID	Age	Sex	Marital status	Job	Income	
P1	1	47	M	Married	Yes	38 000	P1	1	48	M	Married	Yes	38 000
P2	2	44	F	Married	No	0	P2	2	45	F	Married	No	0
P3	3	20	M	Single	Yes	23 000	P3	4	15	M	Single	No	0
P4	4	15	M	Single	No	0							
Household 2							Household 2						
P1	5	79	F	Widowed	No	14 000	P4	3	21	M	Married	Yes	25 000
							P5	8	19	F	Married	No	0
Household 3							Household 3						
P1	6	37	M	Married	Yes	32 000	P6	6	38	M	Married	Yes	35 200
P2	7	38	F	Married	No	0	P7	7	39	F	Married	Yes	14 175
P3	8	18	F	Single	No	0							
Household 4							Household 4						

Source: Hellwig [18].

differentiation is based on the question of whether the time aspect in the population is ignored or not. This results in static and dynamic microsimulation models respectively. The static ones measure only first order effects: it is assumed that the sample structure of the micro-units does not change. Weighting factors adjust the microdata base to the population in the future. These models are only suitable for short- and medium-term simulations. In dynamic models, the sample characteristics are aged in each period and micro-units are allowed to change their behaviour as a consequence of changes in the system. Children and immigrants can be added to the sample and household situations are allowed to change by eg marriage or divorce. So the number of micro-units in dynamic models generally changes in time, whereas the number remains constant in the static models. Dynamic models are as originally proposed by Orcutt and they explicitly model demographic processes. This does not imply that static models do not contain behavioural relationships; only the demographic part is assumed not to change.

Within the dynamic type, two subtypes can be distinguished: longitudinal and cross sectional models respectively. In the longitudinal approach the complete life history of each separate micro-unit is simulated in one run, without interaction with other micro-units. In the cross sectional approach all micro-units are aged for the first year, and then all these micro-units pass through the second year of the simulation, and so on. In this way interaction between micro-units is continuously possible. The life history of one individual is only known if all micro-units have passed through the whole simulation period. The absence of interaction in the longitudinal approach is a very great drawback. The disadvantage of the last type of model

is that it requires more information, and using this information is also more expensive in terms of computer program development, computer time and so on. But its advantage is clear: a more realistic approximation of reality. A discussion can also be found in Hayes [16]. Because of their costs, the number of applications of dynamic cross sectional models is rather limited. Extensive models – covering not only household dynamics, but also socioeconomic characteristics – are only available in the USA and FR Germany. An overview of these three types of microsimulation is given in Figure 3.

The most important advantages of microsimulation are as follows:⁸

- (i) Microsimulation takes place at the level at which most decisions are made, namely that of micro-units. The micro-unit is the natural starting point for analysis. Aggregation is therefore unnecessary and the concomitant problems can be avoided. So there is no need to translate behavioural relations at the microlevel to the macrolevel. This is especially important in the field of social sciences because aggregation of theories to the macrolevel has not been carried out except under very restrictive assumptions (see, for example, Kirman [26]). This also implies that no information is ever lost through aggregation. Different types of institutional rules – eg in the field of taxes and social security – can also be modelled.
- (ii) Modern social scientists break problems down

⁸ See Social and Cultural Planning Agency ([104], pp 48–49), Orcutt, Glazer, Harris and Wertheimer ([100], p 86), Arrow ([2], pp 258–260), Orcutt and Glazer ([99], pp 124–125) and Hellwig ([19], pp 30–32).

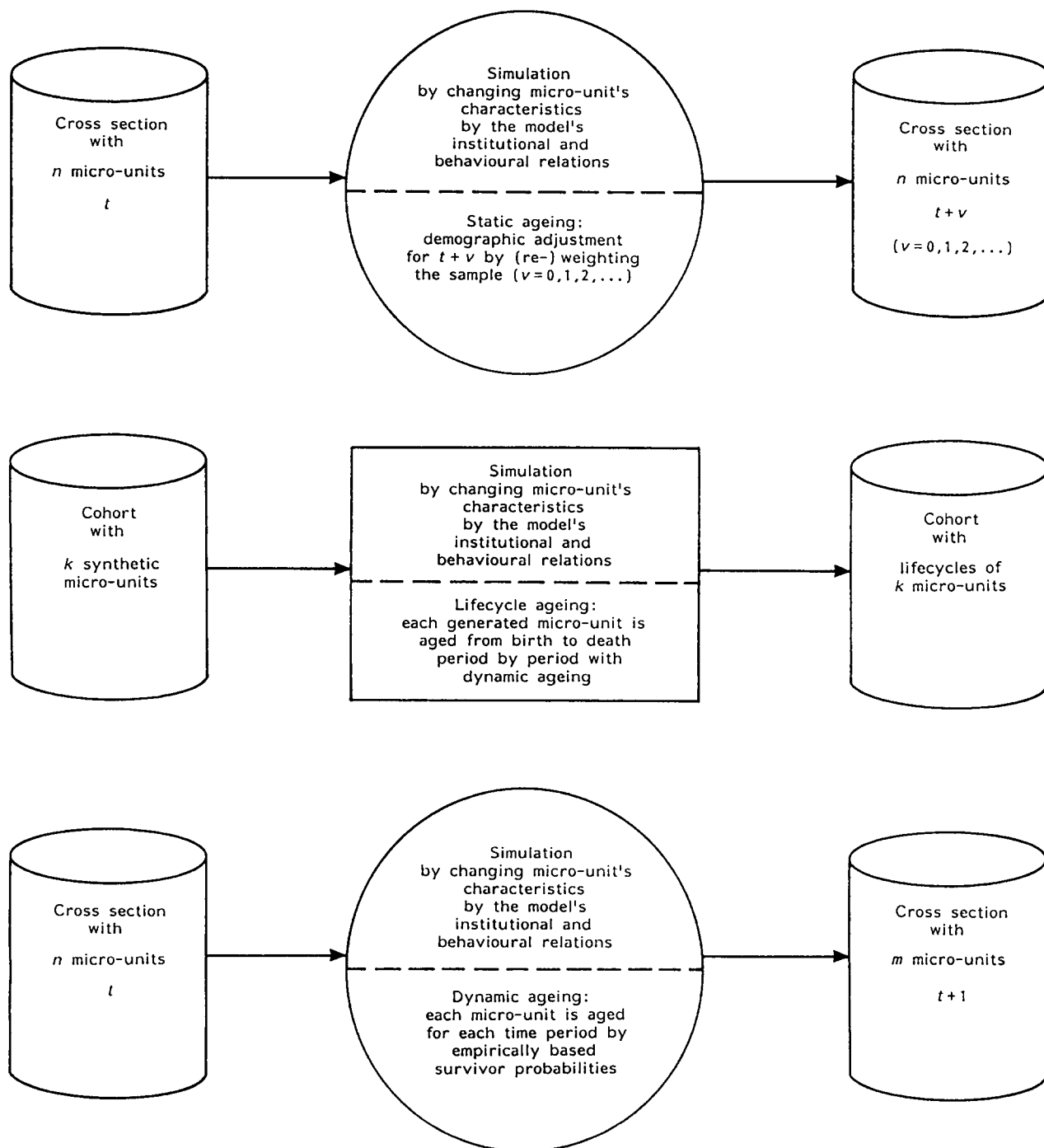


Figure 3. Static, dynamic and longitudinal microsimulation.

Source: Merz [31], p 6.

into manageable parts. This leads to problems when trying to put the parts together again. In many cases, the quantitative revolution in social sciences has resulted in a strengthening of the imbalance of analysis over synthesis. Microsimulation can contribute to the reverse (see Caldwell [8], 1990).

(iii) There is hardly any loss of data. In principle, all microdata can be used. As a result, there is a great deal of available information, and there is full representation and treatment of the heterogeneity in household and individual characteristics. It is possible, for example, to generate the income distribution of a particular

year; this is virtually impossible with macromodels. In this manner we can also gain insight into the question of how many and which households are affected by specific measures of, for example, the social security system, and how this influences further behaviour (see, for example, Bekkering *et al* [4], for such processes in the field of tax legislation).

- (iv) The most complicated interrelationships can be described; in such cases paramount importance is attached to the explicit treatment of dynamics in simulating the system. For example, the determination of eligibility for a social welfare benefit or for a rent rebate is relatively easy with the aid of the microsimulation model, which can also determine its effect on, for example, the labour supply of various household members.
- (v) It is relatively easy to maintain consistency, which is a general methodological problem in, for example, modelling household and family dynamics (see Galler [12]).
- (vi) Assumptions and hypotheses with respect to the micro-units can be introduced or changed at any time. The model is very flexible in view of its use of modules: each process is generally implemented in a separate module. These modules can be developed independently of other modules. In a microsimulation model alternative policies can be modelled and calculated very quickly.
- (vii) The method is intuitively appealing because it uses existing units (eg individuals).

It goes without saying that microsimulation also has a number of drawbacks.⁹ However, it is preferable to speak of (practical) problems than of disadvantages.

- (i) The behavioural hypotheses are often based on insufficient knowledge (such as the behavioural hypotheses which form the basis of demographic decisions). As a consequence all microsimulation models have only a limited number of behavioural responses, particularly in the field of household formation.
- (ii) Large databases are needed, or have to be simulated. In the Netherlands such databases are limited. One of the available databases is the Socio-Economic Panel, but in view of the costs involved, it could not be used for our purpose.
- (iii) The construction and maintenance of microsimulation models requires large-scale investment in manpower, computer capacity and computer time. In view of the high costs of microsimulation in comparison with other methods (case studies,

mental models, regression models, demographic and social accounts) it is amazing that the method has been applied at all. It is estimated that the development of the SPES/Sfb3 models required 100 to 150 man years and the Darmstad model about 15 man years.¹⁰

- (iv) In addition to variance related to the specification of the initial sample population, Monte Carlo variability results from the sampling approach to selecting the members of the model population whose characteristics change. By repeating the runs, we can gain deeper insight into the magnitude of this phenomenon (see Orcutt *et al* [98] and Nelissen [93]). Microsimulation does not give a unique solution.

Looking at the drawbacks, we can simply state that they are also particularly applicable to other methods which try to give the same type of output (lacking behavioural hypotheses and the required database). Microsimulation is certainly expensive in terms of computer capacity, manpower in construction and so on. This is, however, inherent to its nature: using large input to give very detailed results. The Monte Carlo variability problem can be reduced by using larger microdata bases or by repeating the runs, so that a distribution for the desired result can be derived. But this is a relative problem in a certain sense: the solution to the problems treated with microsimulation models cannot be solved analytically. So, the choice is between having no solution or a solution subject to Monte Carlo variability.

All in all, we can conclude that a micro-approach has to be preferred for our purpose and that, given certain sorts of problems, microsimulation is the appropriate method to use.

Our model NEDYMAS is a dynamic cross sectional model. Demographic processes are explicitly simulated, which implies that the size of the microdata base changes during the simulation period. The sample passes through time year by year. For each person in the microdatabase we investigated which personal characteristics changed, and to what extent, each year. We stated above that the heart of microsimulation modelling is formed by its state representation of the components of the system of interest. This is executed by drawing up a list of attributes for each individual in the sample. The storage of this list of attributes requires 512 bytes per individual. Each time one of the attributes is changed, the attribute in question is overwritten. If some data need to be kept for further analysis, they have to be transferred to tape.

The next step, after the adaptation of a micro-

⁹ See Social and Cultural Planning Agency ([104], pp 49–51), Klevmarken [27], Hoschka [22] and Hellwig ([19], pp 32–34).

¹⁰ See Hellwig ([20], p 6).

representation, is the specification of an initial population. We would prefer to use a real sample of individuals and households, together with their attributes. However, such a sample is not available. A first usable sample can be derived from the 1947 Census data.¹¹ The construction of this initial population is described below. As a consequence, the first birth generation which can be followed over time is the generation born in 1930.

Once the initial population has been determined, our analysis can be carried out by updating the characteristics of individuals (and thus households) during the period under investigation. Here, the Monte Carlo sampling procedure is applied. NEDYMAS has been organized modularly. So we do not have a massive program, but rather a set of subprograms, which can be used – partly or together – at will. For example, the demographic modules can be used to generate the household structure.

A problem in programming the model is the simultaneity of the processes. Generally, the periodicity of the data is one year, whereas the periodicity of processes in the real world is much less. The processes can take place at any time within the year, also within one block eg the demographic module. Therefore, more processes can take place simultaneously within one year. This simultaneity influences the calculation of the transition probabilities, when opting for a sequential ordering of events. In NEDYMAS the problem is solved by using the appropriate occurrence exposure rates, so that the sequence in which the processes are simulated has been taken into account. In this way, a random determination of the sequence in which the processes and the individuals to be simulated are executed – as proposed by Hellwig ([20], pp 11–12) – is not necessary.

Carrying out an analysis with the aid of NEDYMAS requires some simple control modules. The specification of a simulation requires the writing of a number of these control modules (to determine the simulation process and the desired output), compiling them (translation of the program into machine code) and linking them with the other modules needed. Unlike macroeconomic models, no comfortable packages for microsimulation exist. The existing models are implemented in an *ad hoc* manner. This is understandable, since microsimulation programming requires more efforts than the management of the structure of time series data. We also opted for an *ad hoc* approach. Both an ALGOL68 and a FORTRAN version of NEDYMAS are available.¹²

Analyses with NEDYMAS can be run under batch, the FORTRAN version also interactively. However, the method of organization is not really user friendly. An experienced programmer, who knows the program, is needed to carry out modifications. The advantage, on the other hand, is that program development is rather simple. Moreover, flexibility is greater, and users can write their own routines.

The external data required – eg for the calculation of transition probabilities – are stored sequentially, whereas the list of attributes of the sample population members is stored using random access capability. This random access capability makes it very easy to execute interactions. If the receipt of a benefit depends on the characteristics of other household members, it is almost impossible to determine the eligibility for a benefit when applying sequential storage.¹³ Its price, however, is that the process is relatively expensive in input–output operation in comparison with sequential storage, and that disks have to be used instead of tapes, which is also more expensive. To save computer time, the input–output operations for the list of attributes is written in machine code.¹⁴

The demographic and education module of NEDYMAS

Introduction

The demographic module contains the following submodules: immigration, emigration, mortality, marriage, cohabitation, divorce, the breaking up of a cohabiting household, henceforth called ‘dehabitation’, flows into and out of old people’s homes, splitting off of children and fertility. The way in which the demographic module is passed through is indicated in Figure 4.

Each year starts with the determination of the number of new immigrants. By new immigrants, we mean people who have never before lived in the Netherlands. New immigrants lead to an increase in the number of personal records. Next, the individuals in our database who are still alive successively go through the following processes: family reunification of immigrants, emigration and return immigration. Then, a check on immigration takes place to ensure that all new immigrants have been dealt with.

Thereafter the people of our database are subjected to the possibility of moving into or out of an old

¹¹ See NCBS ([59], [60], [62] and [65]).

¹² The FORTRAN version was implemented by Willem Spoelتمان and Paul Vermaseren.

¹³ For that reason benefit eligibility can only be treated in the first operating version of DYNASIM – DYNASIM/MASH – which uses a random access method correctly and is less sophisticated in the second version – DYNASIM/MASS – which uses a sequential method (see Devine and Wertheimer, [10], p 55).

¹⁴ This was done by Paul Vermaseren.

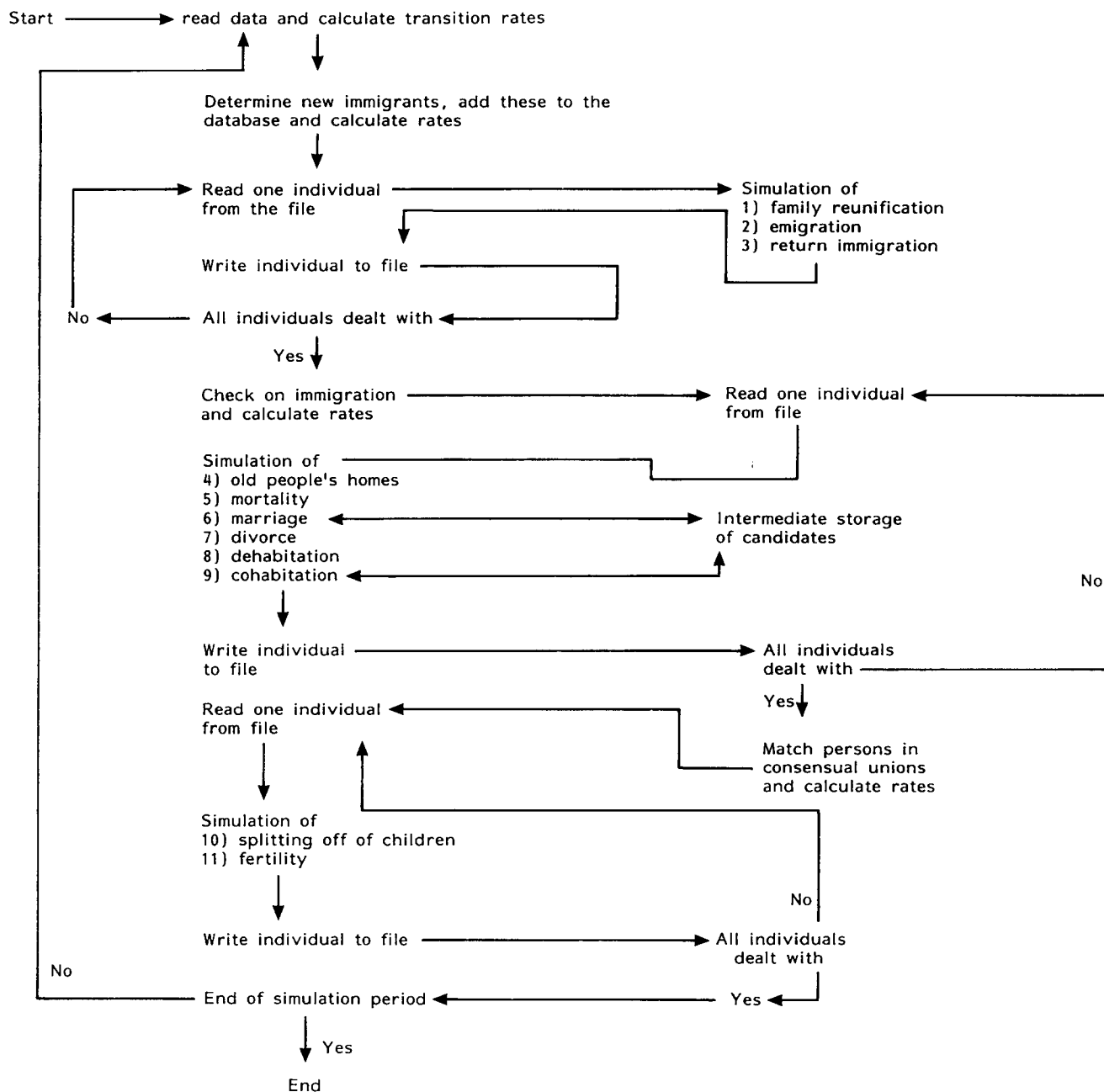


Figure 4. The demographic module.

people's home, of death, of being a candidate for marriage, of divorce, of dehabitation and of being a candidate for cohabitation. When all persons are dealt with, the matching of persons takes place to complete the simulation of marriage and cohabitation. The last steps in each year are the simulation of splitting off of children and fertility.

During the passage through the demographic module, different occurrence exposure rates are calculated. First, the family reunion rates and return immigration rates are determined. After the addition of new immigrants the return emigration rates and the first emigration rates are calculated. After the check

on immigration, the rates for the flows into and out of the old people's homes, the death rates, marriage rates, divorce rates, dehabitation rates and cohabitation rates are calculated. The rates for leaving the parental home and fertility are determined after the matching of persons in consensual unions has taken place. In calculating these rates, we used, among other data as generated by the simulation model, NCBS data on the population by age, sex and marital status for the years 1947-85¹⁵ and the same forecast

¹⁵ These data were provided in the form of tables by the Netherlands Central Bureau of Statistics (NCBS). See also NCBS [63] and [73].

data from the NCBS 1984 forecast.¹⁶ From 2035, the rates are held constant at the 2034 level.

The size of a number of simulated subpopulations is determined to calculate some of the above mentioned rates. This concerns the number of new male and female immigrants and emigrants by year of immigration and emigration, respectively; the female population, aged 15 up to 50 years, by parity, marital status and cohabiting or not; the number of dehabitations plus the number of cohabitations that ended in the death of one of the partners plus the number of cohabiting men emigrating without their families; the number of married men living abroad without their families; the number of immigrant married men living in the Netherlands without their families and the number of male and female persons, aged 15 up to 40, still living with their parents.

All demographic events are assumed to take place on 1 July each year. Divorce and dehabitation do not occur in the year of marriage or cohabitation. And after a divorce or dehabitation, the possibility of remarrying or recohobiting does not exist until the next year. If a person dies, he or she is, in spite of this, subjected to the other possible demographic events. However, the probability that these events will occur is halved. In the case of emigration, the person's record is maintained. In the year of emigration other possible demographic events also occur with a 50% chance. In the next years, no demographic changes are allowed until the person decides to return to the Netherlands, with the exception of mortality. This means that, in spite of the recursive nature of the model, competing risks are dealt with to a large extent. Only the occurrence of marriage, cohabitation, dehabitation and divorce in the same year are excluded. The resulting error, however, is very small.

Technically, the problem could be solved by introducing the possibility of remarriage or recohobitation in the year of divorce or dehabitation. Its disadvantage is that eg the remarriage rates are not specified by duration since divorce. Were we only to use an age specific rate then remarriage probability would be strongly overestimated. The same holds for a divorce in the year of marriage. Therefore, the chosen option seems to be the optimal one in view of the absence of the required rates by duration.

In the following subsections we will describe these elements of the demographic module. The last subsection describes the construction of a representative initial population.

Immigration

Immigration is simulated in four steps. The first step is the determination of the number of new immigrants and the number of return immigrants. The starting point is the total number of immigrants. For 1947 we know the total number of male and female immigrants.¹⁷ Using the 1949 distribution of immigrants by age group and the distinction between married and non-married persons and using the distribution by sex, age and marital status in 1960, these figures are transformed into age, sex and marital status specific numbers. For 1948 we know the number of immigrants by sex, by six age groups and by the distinction between being married and being non-married.¹⁸ These are transformed into sex, age and marital status specific numbers in the same way as those for 1947. For the years 1949 up to and including 1959 we know the total number of male and female immigrants for the age groups 0–14, 15–19, 20–24, 25–29, 30–34, 35–39, 40–49, 50–64, 65 and older, all – with the exception of the first age group – for married and non-married people.¹⁹ These numbers are transformed into age, marital status and sex specific numbers, using the distribution of immigrants by marital status and sex for the distinguished age groups in 1960 and the localization coefficients within these groups in the same year.

For the period 1960–85 we obtained age, sex and marital status specific number of immigrants from the NCBS,²⁰ and from 1986 to 2010 we use the NCBS 1984 forecast. After 2010 the number of immigrants has been set equal to zero, in accordance with the NCBS 1984 forecast.

The number of return immigrants has been calculated on the basis of the return immigration rates by number of years passed since the year of emigration and the number of emigrants in those years who are still living abroad. We used the return immigration rates for the years 1950 (covering the period 1947–52), 1955 (1953–57), 1960 (1958–62), 1965 (1963–72), 1975 (1973–77) and 1980 (covering the period from 1978).²¹ In this, we limited ourselves to people who emigrated between the current year and six years earlier. However, we do not know the people who emigrated before 1947, in the sense that we do not know their personal attributes. They are not in our sample, because our initial population only contains people who were living in the Netherlands on 31 May 1947.

¹⁷ Source: NCBS tables.

¹⁸ See NCBS [58].

¹⁹ Source: NCBS [34].

²⁰ Tables supplied by the NCBS.

²¹ These rates can be derived from NCBS [64], [67], [69], [72], [78] and [83].

¹⁶ Also available from tables from NCBS, see also NCBS [88].

For that reason, in 1947 we first determined only the number of return immigrants, who had emigrated from the Netherlands, in the same year between 1 June and 31 December. Later a proportion (based on the return immigration rates) of the new immigrants were called return immigrants and supposed to have emigrated between 1 January 1941 and 31 May 1947. In 1948 this procedure was applied for the determination of return immigrants who had emigrated between 1 January 1942 and 31 May 1947 and so on. This procedure is necessary to find out the real number of new immigrants, because return emigration only holds for new immigrants and does not apply to reimmigrated persons.

Subtracting the resulting number of return immigrants from the total number of immigrants, we get the number of new immigrants (by age, sex and marital status).

The second step is the addition of new immigrants to our sample. For that purpose the number of new immigrants is divided by the proportion of the sample population in the real population (the sampling ratio), in which rounding occurs by lot, using the uniform $[0, 1]$ distribution. This resulting number of new immigrants is added to our sample. In this process, we distinguish six subgroups:

- (i) Married women with or without children, who immigrate without a partner. We assume that these are cases of family reunification. The proportion of these women in the total number of immigrating married women is calculated using the probability that family reunification takes place with immigrated married men, living in the Netherlands without their family. These probabilities are derived by assuming that married females, immigrating without their husbands are cases of family reunification.²² For the future, we assume that the number of women involved in family reunification as a proportion of the total number of married women will equal the proportion in 1980 (being about 65%).²³ The number of women in this subgroup is related to the number of immigrated married men living in the Netherlands without a partner (see (v)). This gives us the probability of family reunification. Using a Monte Carlo process, we determined for each married man without a partner whether or not family reunification takes place. If this occurs, one female person is added to the sample. The number of persons in the database is raised by one and she gets this number as her personal identification number. The sex of this person is

of course 'female', the marital status 'married'. The pointer to her husband equals the identification number of the man in question and conversely, the pointer of the man for his wife is set equal to the identification number of the female in question. Her year of immigration is the current year. The age of the wife is set equal to the male's minus three, unless such a woman is not available in the set of immigrating married women. In the latter case the age difference is minimalized. The age at which the woman is assumed to have left her parental home is put at the year in which she became eighteen years old. If no more married women are present in the population of new immigrants, no family reunification takes place for the man in question, otherwise the number of new immigrating married women of a given age is decreased by one. If family reunification did take place, the number of minor children involved in the reunion is simulated. For that purpose we use the distribution of immigrating family heads by number of minor children. We assume that immigrating women aged 64 years and older do not have minor children in their households. The number of minor children is determined by a Monte Carlo process. The age of the children is dependent on the age of the woman and on the number of children by age in our population of new immigrants, if available. If a child is added to the immigrating female, the number of persons in the database grows by one and this number becomes the child's personal identification number. The sex of each child is assigned by lot, his or her marital status is unmarried, not cohabiting. The child gets a pointer for his father and mother, and vice versa. Also, the number of siblings is attached to the child's attributes. The number of children of the father and mother is increased by one and the number of siblings of the other children is also increased by one. Every time a child is allocated to a woman, the number of children of that age and sex in the population of new immigrants is diminished by one. If no more children are available, no children are attached in the case of family reunification.

- (ii) Never married people older than 17 years. We assume that these persons immigrate as single persons, without children. The number of this kind of person in the population of new immigrants is added to our sample. These persons get their personal attributes in the same way as the immigrating married women.
- (iii) The same applies for the group of widowed and divorced men.

²² See NCBS [64], [67], [69], [72], [75], [78] and [83].

²³ NCBS [83] and own calculations.

- (iv) Widowed and divorced women: these groups are also added to our sample. In this case, children may be present. The presence of children is determined in the same way as under (i), on the understanding that now the number of children is determined by a draw from the distribution of female-headed immigrating households by number of children.
- (v) Married men who immigrate without their families and consequently leave behind their wives and children, if present, in the country of origin. The probability that a married man will immigrate without his family is derived from NCBS [34]. A Monte Carlo process determines whether this will happen. If this is the case, the man is added to the sample, otherwise we have to do with a person of category (vi).
- (vi) Married men, immigrating with their families. The probability that this will occur is determined above. A wife is attached to a married man, analogously to men to women in category (i) if present. Children are attached in a similar manner.

When this process has been carried out all never married men and women older than 17 years, all widowed and divorced persons, and all married men out of our population of new immigrants are added to our sample. However, we are not sure that this also holds for married women and for minor children. Now, the third step, if necessary, contains the addition of the remaining married women and minor children to the sample.

When the population of newly immigrated married women does not equal zero, then first the remaining women are added to the sample and a husband is drawn from the set of married men living in the Netherlands without a partner. Both wife and husband get a pointer to their partner.

Analogously, if minor children are present in our set of new immigrants, a father and/or mother are drawn for each child out of the immigrated married people in the current year and the child is added to his father and/or mother. Pointers are added and so on.

Now all new immigrants are inserted in our sample. The fourth and last step is the simulation of return immigration. Earlier, we fixed the number of return immigrants by year of emigration up to and including six years ago. Dividing these numbers by our sampling ratio and the number of emigrants (who have not yet 'return immigrated' or died) in the last seven years in our sample, gives us the probability of return migration. These probabilities are used to determine whether a person who emigrated during the last six

years will return to the Netherlands. In this, we assume that married women immigrate with their husbands and that minor children immigrate with their mothers, or in case of divorce or widowhood, with their fathers. Adult children are assumed to decide independently whether they will immigrate or not. If parents decide to immigrate, whereas their children (older than 17 and living with their parents) do not, the consequence is that these children will decide to leave their parental home and form a one-person household. Analogously, if these children decide to immigrate, whereas their parents stay in the Netherlands, they will also decide to leave the parental home. Their personal attributes are then adjusted.

Emigration

Emigration is largely simulated analogously to immigration. However, no persons are added to our sample population. Now also, the first step is the determination of return emigrants and new emigrants. The starting point is the total number of emigrants. For 1947 we know the total number of male and female emigrants.²⁴ Using the 1949 distribution of emigrants by age group and the distinction between married and non-married persons, and using the distribution by sex, age and marital status in 1960, these figures are transformed into age, sex and marital status specific numbers. For 1948 we know the number of emigrants by sex, by six age groups and whether they are married or unmarried. These are transformed into sex, age and marital status specific numbers in the same way as those for 1947. For the years 1949 up to and including 1959 we know the total number of male and female emigrants for the age groups 0–14, 15–19, 20–24, 25–29, 30–34, 35–39, 40–49, 50–64, 65 and older, all – with the exception of the first age group – for married and unmarried people.²⁵ These numbers are transformed into age, marital status and sex specific numbers, using the distribution of emigrants by marital status and sex for the distinguished age groups in 1960 and the localization coefficients within these groups in the same year.

For the period 1960–85 we have age, sex and marital status specific numbers of emigrants and from 1986 to 2010 we use the NCBS 1984 forecast.²⁶ After 2010 the number of emigrants has been set equal to zero, in accordance with the NCBS 1984 forecast.

The number of return emigrants is calculated on the basis of the return emigration rates by number of years passed since the year of immigration and the number of immigrants in those years, still living in the

²⁴ Table received from NCBS.

²⁵ See NCBS [34].

²⁶ From tables from NCBS.

Netherlands. We used the return emigration rates for the years 1955 (covering the period 1947–57), 1960 (1958–62), 1965 (1963–67) and 1970 (covering the period from 1968 on).²⁷ We limited ourselves to people who had immigrated between the current year and five years earlier. However, we do not know the people who immigrated before 31 May 1947 in the sense that we do not know their personal attributes. They are in our sample, but our initial population does not contain information on this subject. For that reason, we determine – for 1947 (between 1 June and 31 December) – only the number of return emigrants, who had immigrated in the same year. Then, a proportion (based on the return emigration rates) of the new emigrants were called return emigrants. These are assumed to have immigrated between 1 January 1942 and 31 May 1947. In 1948 this procedure was applied for the determination of return emigrants who had immigrated between 1 January 1943 and 31 May 1947 and so on. This procedure must be applied to determine the real number of new emigrants, while return immigration only holds for new emigrants and does not apply to reemigrated persons.

Subtracting the resulting number of return emigrants from the total number of emigrants, we get the number of new emigrants (by age, sex and marital status). On the basis of the populations at risk, the probability of return emigration and the new emigration rates are determined. As for return emigration, it is assumed that the probability is independent of age and household composition. Return emigration takes place as a family, with the exception of children older than 17 years. The latter group decide for themselves whether they want to emigrate or not. Given the limited data, it is also assumed that return emigration differs from the emigration pattern of new emigrants only during the first five years of their residence in the Netherlands. The new emigration probability is dependent on age, sex and marital status. Cohabiting people are treated in the same way as married people. The whole emigration process is regulated by Monte Carlo decisions.

The process for new emigrating persons can be divided into four subprocesses:

- (i) Married and cohabiting men, who emigrate on their own ie without their families.²⁸ For the years 1981 and later, we assume that this probability does not change. If it is decided (using a Monte Carlo process) that a married man emigrates without his family, a boolean variable assigned to the wife (called *empartner*), which

stands for the question of whether her partner has emigrated alone, is put true. The variable year of emigration is put equal to the current year.

- (ii) Married and cohabiting men, who emigrate with their families. Again, it is assumed that children older than 17 years decide whether or not they want to emigrate, independent of their parents' choice. If a child, still living with his or her parent(s), decides not to emigrate, the consequence is that the child decides to live alone. So the year of leaving the parental home becomes the current year.
- (iii) Not cohabiting men and women older than 17 years. Minor children, if present, emigrate together with their father or mother.
- (iv) Cohabiting women, whose husbands have already emigrated in the past (the variable *empartner* is true, see under (i)) and who emigrate for reasons of family reunification.²⁹ If minor children are present, these will accompany their mother.

If a person has emigrated (and he or she had not immigrated before) he or she is only subjected to return immigration and mortality. All other (demographic) transitions are ignored.

Old people's homes

From the NCBS we know the probability of staying in an old people's home for the years 1950 and 1965 for the age groups younger than 65, 65–69, 70–74, 75–79, 80–84 and 85 years and older for both males and females. From 1965 we have yearly data on the number of people staying in an old people's home, the number of people who died in an old people's home, the number of discharges and the number of admissions for the same age groups with the exception of the oldest one, which has been split up into the age groups 85–89 and older than 89 years. Using these data, occurrence exposure rates were calculated for entering and leaving old people's homes for the years 1965, 1970, 1977 and 1984.³⁰ Between these years the rates were interpolated. From 1984 the rates were held constant, just as the departure rate before 1965 is held constant at the 1965 level. Between 1950 and 1965 we interpolated between the probability of remaining in an old people's home in 1950 and the entrance rate in 1965. For the years 1949 and earlier, the figures for the year 1950 were used. In all these rates and probabilities, a distinction was made between married and unmarried males and females.

Each year, a Monte Carlo process decides whether

²⁷ See NCBS [67], [69], [72] and [75].

²⁸ The probability is derived from NCBS [64], [67], [69], [72], [75], [78] and [83].

²⁹ The probability is also derived from NCBS [64], [67], [69], [72], [75], [76] and [83].

³⁰ See NCBS [61], [71], [74], [80] and [91].

a married woman or a unmarried person enters an old people's home or not, if the person in question is 60 years or older and still living outside such a home. If the person decides to enter an old people's home, the person's attribute 'living in an old people's home' becomes true instead of false. If this person is a married woman, her partner – if he has not emigrated alone – also enters the old people's home and his attribute 'living in an old people's home' also becomes true. Analogously, it is determined whether people living in an old people's home will leave this home. If this occurs their attribute 'living in an old people's home' becomes false.

Mortality

Mortality is simulated by subjecting the individuals in the sample to age, sex and marital status specific death occurrence exposure rates. These rates are derived from NCBS data and forecasts. The number of deceased people by age, sex and marital status for the period 1947–85 are known from NCBS.³¹ The numbers for the period 1986–2035 are derived from the NCBS 1984 forecast. The death rates were calculated using the population at risk. In this process we distinguish between persons remaining in an old people's home and persons living on their own. The death rates for people living in an old people's home and not living there respectively by age as a proportion of the death rate for the whole population by age, are derived from NCBS and the death rates as calculated before.³²

Each year a Monte Carlo process decides for each person whether he or she will die. If this occurs, the individual's year of death is put equal to the current year. If the person in question is married, his or her partner becomes widowed and the year of widowhood is put equal to the current year. If the individual cohabits, the partner's marital status changes back into the status he or she had before the start of the cohabitation with the deceased person. Also in this case, the attribute 'year of widowhood' is put equal to the current year. If a married man, who has emigrated without his family has died, then the attribute *empartner* is set to 'false'.

If the deceased person has children, then the pointer to the father or mother (depending on who died) of these children becomes zero, whereas for each child the last (last but one) child's pointer of the identification numbers of children at the former partner is put equal to the pointer to the child's father (mother). If the other parent had already died before, the child has become an orphan.

Marriage

From 1948 to 1985 we have the number of marrying persons by age, sex and (former) marital status for each year.³³ The same (forecast) data are available from the NCBS 1984 forecast for the period 1986–2035. For the year 1947 we have the total number of marrying persons by sex and (former) marital status for eight age groups.³⁴ Using the localization coefficients in the year 1948, these numbers are transformed to age, sex and (former) marital status specific numbers. Marriage rates were calculated by dividing these numbers by the appropriate population at risk.

For each person in our database who is older than 14 years and younger than 85 years and not married, a Monte Carlo process decides each year whether or not he or she will marry. If the decision is to marry, the person is stored in a file of potential marriage candidates. When all persons have passed through the marriage submodule, the matching process takes place (see Figure 4).

First, the number of marriages to be contracted is determined. It is put at half the sum of the number of male and female candidates, in which a Monte Carlo process decides in case of rounding off. The resulting number is indicated by *N*.

The next step in the matching process is contracting the marriages of cohabiting persons. The female is the starting point in this. Therefore, we take all women in our stock of candidates who live in a consensual union. They are said to marry their partner (even if this partner is not included in our stock of male candidates). These people get the marital status 'married' and their year of marriage becomes the current year. It is possible that the wife or husband has died in the current year. In that case the marital status of the person still alive becomes widowed. If both partners have died during the current year, nothing has changed.

If all women in our stock of candidates who lived in a consensual union, have married their partners, then all men and women in this stock who cohabit are deleted. To get the intended total number of *N* marriages, we determine whether we have to draw extra candidates (not married, not cohabiting, aged 15 to 84 years old), or delete candidates from the stock. In both cases, the localization coefficients for marriages are used as weights. In this way, we get an equal number of men and women who have to be coupled to each other.

Using the distribution of marriages by ages of both partners, the number of marriages still to be contracted

³¹ Tables provided by the NCBS.

³² See NCBS [71], [74], [80] and [91].

³³ Tables provided by the NCBS.

³⁴ Table provided by the NCBS.

after the aforementioned marriage process of cohabiting persons and the RAS method (see Steger [107], pp 147ff), we determined the number of marriages in our sample by age of the partners. This matrix is called *M*. The distribution of marriages by partners' age is available for the year 1960 and 1976 from unpublished data of the NCBS. We used eight age groups: 19 years and younger, 20–21 years, 22–23 years, 24–25 years, 26–29 years, 30–34 years, 35–44 years and 45 years and older. For the years 1959 and before we used the 1960 distribution, whereas for the years 1977 and later the 1976 distribution was the starting point for the RAS procedure. The matching process of the males and females, now left in our stock of candidates, is carried out in two steps. Again, the woman is the starting point. In the first step, a female candidate is taken from the stock and we look at whether the number of male candidates is still more than ten. If so, we draw ten male candidates using the distribution by age of the male partner, given the female's age, and under the condition that the respective cell in the matrix *M* is positive.

In the second step the most suitable candidate out of these ten is determined on the basis of the female's level of education and the distribution of husbands by level of education conditional on the woman's level of education. This last distribution is known for the marriages existing on 28 February 1971 by period of contraction.³⁵ In this, the following periods are distinguished: 1949 and before, 1950–54, 1955–59, 1960–64, 1965–69 and 1970. For the years 1971 and after, we use the 1970 distribution.

If the remaining number of women is ten or less, then these women are coupled to a husband on the basis of distribution by age only. When a wife has chosen her husband, the corresponding cell in the matrix *M* is diminished by one. Both the male's and female's year of marriage becomes the current year, their year of divorce (if positive) becomes zero, their marital status is now married. Each gets a pointer to his or her partner, which equals the partner's identification number. If one of the partners was still living with his or her parents, the year of leaving the parental home of this partner becomes the current year. If one of the partners already had children, these children are also assigned to the other partner. This means that the number of children of this last partner becomes positive, that he or she gets pointers to these children and that the children concerned get a pointer to their new parent. If both partners already had children, these children are mutually assigned to both parents.

It is also possible that the wife or husband has died

in the current year. In that case the marital status of the person still alive becomes 'widowed'. If both partners died during the current year, nothing has changed.

Divorce

Divorce is carried out by subjecting the married women in our sample population to an age specific divorce occurrence exposure rate. These rates are calculated from NCBS data and forecasts. For the years 1948–85 we have the numbers of divorces by age of the women at our disposal, whereas for the years 1986 and later, we use the numbers of divorces by age of the women as forecasted by the NCBS.³⁶ For the year 1947 we know the number of divorces for seven age groups.³⁷ Using the localization coefficients within these age groups in 1948, the number of divorces by age groups is transformed into the number of divorces by individual ages. A Monte Carlo process decides whether or not a wife's marriage ends in divorce in the current year.

If divorce occurs, the marital status of the woman and her husband are changed to 'divorced'. The pointer to the partner becomes zero, whereas the pointer referring to the former partner is set equal to the corresponding personal identification numbers. For both individuals the year of divorce becomes positive (and equals the current year). If the husband has emigrated, the number of emigrated people without a partner is diminished by one.

If children live with their parents, we determine for each child whether the woman or her husband is granted custody of the child. Up to 1975 we assumed that the father had a 5% chance of being granted custody. For the years 1980 to 2000 the probability amounts to 10% and 20%, respectively. Between these years an interpolation is made for the probability. If a child is allocated to the father (mother), the number of children allocated to the partner is increased by one for the mother (father) and the corresponding pointer of children allocated to the partner is set equal to the child's identification number.

Dehabitation

Dehabitation is carried out by subjecting the married women in our sample population to an age specific dehabitation occurrence exposure rate. These rates are derived from the age specific divorce rates, while data for the dehabitation process hardly exist. Here, we have to fall back on the scarce information from surveys held in the 1980s, especially as reported by the NCBS [89]. From this source we have drawn information on the dehabitation of persons who

³⁵ NCBS [81], pp 58 and 166–169.

³⁶ Tables from the NCBS. See also NCBS [68] and [77].

³⁷ Data provided by the NCBS.

cohabit for the first time between 1974 and 1979 and who were younger than 30 years old when they started cohabiting. The data refer to the duration of their cohabitation. If we compare the resulting rates with the divorce rates by duration of marriage in the early 1980s, we can deduce that the probability of moving out in the first six years of cohabitation is 3.13 times higher, on average, than the divorce probability of married people in the same period of six years. After about six years, however, the rate of splitting up approximates the divorce rate. Moreover, the older the persons in question are, the smaller the difference between both rates. Therefore, we take the dehabitation rate as being 3.13 times the divorce rate, corresponding to the wife's age, if the duration of the cohabitation amounts to six years or less, and the wife is 35 years old or younger. In all the other cases, the moving out rate is assumed to be equal to the divorce probability. A Monte Carlo process decides whether or not a wife's cohabitation ends in a dehabitation in the current year.

If moving out occurs, the marital status of the woman and her husband are changed into their marital status prior to the last cohabitation. The pointer to the partner becomes zero, whereas the pointer referring to the former partner is set equal to the corresponding personal identification numbers. For both individuals the year of dehabitation becomes positive (and equals the current year). If the husband has emigrated, the number of emigrated people without a partner is diminished by one. If children live with their parents, we determined for each child whether the woman or her husband was granted custody of the child. This was exactly the same as in the case of divorce.

Cohabitation

The process of (unmarried) cohabitation is largely analogous to the marriage process. The most important difference is the way in which we determine the transition rate for a first cohabitation. We assume that no cohabitations took place before 1965. From van 't Klooster-van Wingerden ([29], pp 35–37) we find 18 291 male headed non-family households with female non-relatives, 16 671 female headed non-family households with male non-relatives, 3126 male headed one-parent family households with other female non-relatives and 8294 female headed one-parent family households with other male non-relatives for the year 1971. This makes 46 382 unmarried cohabitations in 1971. It is assumed that 5000 cohabitations existed in 1965 and that there was an exponential growth between 1965 and 1971.

For the year 1982, we find about 205 000 cohabitations on the basis of the Netherlands Fertility

Survey³⁸ and additional estimates for persons older than 37 years. Between 1971 and 1982 we also interpolated exponentially. For the years 1983 and later, we assume that the percentage of people living together, married or unmarried, at every age, does not change.

Given the (expected) number of married people by age and sex, we can calculate the number of people cohabiting by age and sex. Using this number of cohabiting people, the number of people by age and sex cohabiting in the model population at the beginning of the year, the expected number of non-first cohabitations started in the current year (see below), and the (expected) number of those moving out in the current year yields us the number of first cohabitations in the current year. Coupling these numbers to the population at risk gives the first cohabitation rate. On the basis of the Netherlands Fertility Survey it is also assumed that the non-first cohabitation rate equals the remarriage rate for divorced people for those 33 years or older and it equals twice this remarriage rate for those younger than 25 years old. Between the two ages the multiplication rate is interpolated.

Using these cohabitation rates, a Monte Carlo process decides whether or not a non-cohabiting person will be a cohabitation candidate. If so, he or she is temporarily stored. After all persons have gone through the cohabitation submodule, the candidates are matched. This process is analogous to the marriage process for non-cohabiting marriage candidates. Therefore we can refer to the marriage process.

Splitting off

The last submodule but one in the demographic part of our model is the splitting off of children for other reasons than marriage or cohabitation. Only limited information is available on this subject. We use results from the 1971 Census and the 1977 and 1981 Housing Surveys (*Woningbehoefstenonderzoek*).³⁹ From these data we know the percentages of never married men and women by age, living with their parents in the years in question. Between 1971 and 1977 and between 1977 and 1981 we interpolated these percentages. For the years 1982 and later, we assume that these percentages do not change, whereas for the years before 1971 we take those of 1971, corrected for the relative participation rate in university education by persons between 18 and 25 years old (1971 = 100) squared.⁴⁰ This last factor is considered as a proxy

³⁸ Personal communication with van der Giessen of the NCBS and own calculations.

³⁹ See van Leeuwen [30], Ploegmakers and van Leeuwen [103] and Heida and Gordijn [17].

⁴⁰ Derived from NCBS [86].

for the developments in splitting off of children for other reasons than marriage or cohabitation in the 1950s and 1960s.

When marriage and cohabitation takes place in our model, we can count the number of never married persons still living with their parents. Decreasing these numbers by the numbers resulting when applying the calculated percentages and dividing these differences by the counted numbers (being the population at risk), gives us the desired probability for splitting off for other reasons than marriage or cohabitation.

Via a Monte Carlo procedure we determined whether a never married or never cohabiting person living with his or her parents, will leave the parental home. If so, he or she starts a one-person household.

This procedure is only executed for persons younger than 40 years. If someone reaches the age of 40 and is still living with his or her parents, we decided that the person in question leaves the parental home in the current year.

Fertility

Births are based on marital status specific parity progression rates for the ages 15 to 50 years. In this parity, one, two, three and four and over, are distinguished. The first step is the determination of the number of births within and without marriage by the mother's age and parity.

The total number of births by the mother's age is known for the years 1947–49 for the five-year age groups 15–19 up to and including 45–49 years.⁴¹ Using the localization coefficients by individual age and parity for the year 1950, these numbers are transformed into number of births by mother's age and parity. The total number of births by mother's age and parity are available for the period 1950–85 from NCBS.⁴² For the period 1986–2035, we use the resulting numbers from the NCBS 1984 forecast.

The number of extramarital births in 1947 and 1948 is known for five age groups: 15–19, 20–24, 25–29, 30–34 and 35–49.⁴³ Using the distribution by parity 1 and 2⁺ in 1950, the distribution of parity 2⁺ by parity 2, 3 and 4⁺ in 1968 and the localization coefficients by age in 1968, these numbers are transformed into numbers of extramarital births by mother's age and parity. For the period 1950–67 the total number of extramarital births is known for the same age groups as for 1948 and 1949, but also by parity (1 and 2⁺).⁴⁴ Using the distribution of parity 2⁺ by parity 2, 3 and 4⁺ in 1968 and the localization

coefficients by age in 1968, these numbers are transformed into numbers of extramarital births by mother's age and parity. The period 1968–79 readily furnishes us the desired information: the number of extramarital births by mother's age and parity.⁴⁵ For the period 1980–84 we have the number of extramarital births for the age groups 15–24, 25–29, 30–34 and 35–49.⁴⁶ Using the localization coefficients for 1979 and the distribution by parity in the same year, these numbers are also transformed into numbers of extramarital births by mother's age and parity. For the period 1985–2026 we use the total number of extramarital births from the low variant of the NCBS 1980 forecast and extrapolate this number for the period 2027–2035.⁴⁷ Using the localization coefficients for 1979 and the distribution by parity in the same year, these numbers are again transformed into numbers of extramarital births by mother's age and parity.

With the aid of the distribution of extramarital births by mother's marital status (never married, widowed and divorced) for the age groups 15–19, 20–24, 25–29, 30–34, 35–39 and 40⁺ in the years 1975 up to 1985 (NCBS [90]), these numbers are transformed into numbers of extramarital births by mother's age, marital status and parity. For the years before 1975 we use the 1975 distribution and for the years 1985 and later, we use the 1984 distribution. Subtracting these numbers of extramarital births from the total number of births, we get the number of births for married women by mother's age, marital status and parity.

The next step is to transpose these numbers into age, marital status and parity specific birth rates. The population at risk is not known from vital statistics. Therefore, we use the simulated female population by age, marital status and parity, as population at risk. For females with four or more children, it is assumed that the parity progression rate for a fifth, a sixth and so on child equals the parity progression rate for a fourth child.

Within the category of never married persons, we distinguish between cohabiting and non-cohabiting women. It appears that the first group's probability of having an additional child is about twice the second group's probability.⁴⁸ We assume that this also holds in the past and the future. Therefore, we adjust the parity progression rates for never married women in the following way.

For non-cohabiting never married women (with *k* children) the calculated parity progression rate (of age

⁴¹ See NCBS ([66], p 47).

⁴² Tables provided by the NCBS.

⁴³ See NCBS ([76], p 22).

⁴⁴ NCBS ([76], p 22).

⁴⁵ Source: tables provided by the NCBS.

⁴⁶ NCBS ([90], p 13).

⁴⁷ NCBS ([82], p 48).

⁴⁸ NCBS ([87], p 34).

a) is multiplied by the quotient of the number of never married women of age *a* with *k* children on the one hand, and the sum of the number of non-cohabiting never married women of age *a* and *k* children and twice the number of cohabiting never married women of the same age and the same number of children on the other hand.

For cohabiting never married women the multiplication factor is twice that for non-cohabiting ones. The parity progression rates for widowed and divorced women are not adjusted for cohabitation because too little information is available for this group.

Using these parity progression rates we determined for each year for every female aged 15–49 years old, whether a child is born or not. If this occurs, the number of children of the women is raised by one, the number of personal records is also increased by one and this number becomes the child's identification number. The same number is used as the pointer of the mother for her newly born child. The child's pointer for his or her mother is set equal to the mother's identification number. The child's sex is assigned by lot and his or her birth year becomes the current year. The number of children minus one is assigned to the newly born one as his or her number of brothers and sisters and the same is done with his or her brothers and sisters. The same adjustments have been made for, and with reference to the partner, if present.

It is possible that the mother has died in the remainder of the year. In that case, the pointer to the mother becomes zero and if no father was present, this means that the child becomes an orphan. Next, the child is subjected to the probability of death during the rest of the year. This procedure is analogous to the one for mortality.

The last step in the fertility submodule is the treatment of the possibility of being early disabled. From data on the *Algemene Arbeidsongeschiktheidswet* (AWW), we know that approximately 0.90% of the persons aged 20–24 years old are early disabled. Their death rate is considerably higher than for non-disabled persons, especially up to the age of 35 years. However, the inclusion of disabled specific death rates requires an extra consistency module for the determination of the death rates for non-disabled persons. In view of the relatively limited number of persons in question (about 0.54% of the total population) we will not use disabled specific death rates. Instead, the inflow probability of becoming early disabled is set equal to the proportion of early disabled persons in the total population. As a consequence, the number of early disabled persons is underestimated up to the age of about 40 years and, to a lesser extent, overestimated from the age of 40 on.

The construction of a representative initial population

To start the simulation process we need a representative sample in the starting year 1947. The available data of the 1947 Census are, in part, suitable for the construction of our initial population.⁴⁹ However, the data are not available at the microlevel. Consequently, we must derive our microdata set from some macro tables. We have, among others the following tables at our disposal:

- (i) The distribution of all marriages, still existing on 31 May 1947 by eight age groups for both partners.
- (ii) The distribution of all marriages still existing on 31 May 1947 by year of contraction of the marriage.
- (iii) The distribution of all marriages, still existing on 31 May 1947 by eight age groups for both partners at the moment of contraction of the marriage.
- (iv) The population by age, sex and marital status on 31 May 1947.
- (v) The number of marriages by year of contraction of the marriage and the number of live-born children in these marriages.
- (vi) The number of marriages with and without children by year of contraction of the marriage and woman's age at marriage.
- (vii) The number of households with children present at composition and number of children.
- (viii) The number of children born out of existing marriages by period of contraction of the marriage and woman's age at marriage.

The simulation of the initial population starts with the simulation of married persons. It appears that the distribution of all marriages existing on 31 May 1947, by age group of both partners fits very well with the marriages contracted in the period 1935–40. This implies that this distribution has remained rather constant during the 1930s and 1940s and that in the period between about 1900 and 1930 the number of marriages at higher ages (50+) must have decreased. This decrease must be comparable with the loss of marriages as a consequence of death, which have occurred between the year of contraction and 31 May 1947. If this was not the case, the correspondence between all existing marriages and those contracted in the late 1930s would be almost impossible. Assuming this, we can state that the distribution of marriages by age of both partners on 31 May 1947 can be used for all existing marriages, independent of the year of contraction of the marriage.

⁴⁹ See NCBS [59], [60] and [65].

Table 4. Family parity progression rates for the marriage generations 1914–46 (parity 1, a_1 , up to and including parity 10, a_9).

Marriage generation	a_0	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9
1944–46	0.47	0.17	0.05	0.00	0.00	0.00	0.00	0.00	0.00	0.00
1939–43	0.86	0.68	0.42	0.34	0.27	0.21	0.18	0.16	0.33	0.00
1934–38	0.89	0.82	0.64	0.57	0.54	0.50	0.45	0.38	0.32	0.26
1929–33	0.89	0.84	0.71	0.67	0.66	0.65	0.64	0.61	0.57	0.52
1924–28	0.89	0.84	0.72	0.70	0.70	0.71	0.70	0.69	0.67	0.64
1919–23	0.90	0.85	0.74	0.72	0.72	0.72	0.72	0.72	0.70	0.67
1914–18	0.92	0.89	0.78	0.74	0.74	0.74	0.74	0.73	0.71	0.68

Put the sampling ratio in our initial population at $1/n$. We now draw $1/n$ part of the number of married men and women out of the table as mentioned under a(iv). Rounding off is executed by a Monte Carlo process. We look for an equal number of males and females, as described above. Suppose that m men and m women result.

For each of the m women, for whom we have the exact age, we continue as follows:

- (i) With the aid of the distribution of all marriages still existing on 31 May 1947 by age groups for both partners at the moment of contraction of the marriage (see under (iii) above), we determine the probability for the age group in the year of her marriage. Thereafter the age group is assigned by lot.
- (ii) The husband's age group at marriage can also be determined using the distribution under (iii) above. The exact husband's age at marriage within this age group is fixed by lot.
- (iii) The woman's age in 1947 and the age group at marriage gives us the marriage period. Using the distribution under (ii) in the previous list the exact year of marriage has been determined. Now, also the husband's age in 1947 is known. It is assumed that both partners left their parental home at the moment of marriage.
- (iv) We now control whether a married man of that age is available in the group of m drawn married men. If a husband with that age is available, the man is coupled to the woman. If not, a man is looked for in the nearest age group (one year or older), and so on.
- (v) Next, we determine the number of children born in this marriage. The family parity progression rates for the different marriage generations can be derived from the data under (viii) above. These are given in Table 4. The family parity progression rates for the marriage generations 1913 and before are set equal to these for the generation 1914–18. Of course, the woman's age at marriage is also an important determinant for the number

Table 5. Correction factors for the mother's age at marriage for the parity progression rates for families with no children (a_0) and at least one child (a^*).

Age at marriage	a_0	a^*
<20	1.09	1.35
20–24	1.05	1.12
25–29	0.99	0.92
30–34	0.88	0.69
35–39	0.68	0.35
40–44	0.33	0.09
>44	0.11	0.01

of children born from that marriage. Correction factors for the age at marriage can also be deduced from the data under (viii) above. We assume that these are independent of the year of marriage. We only distinguish between the parity progression rate for families with no children and that for families with at least one child. These correction factors are to be found in Table 5. The number of children born out of the marriage is assigned by lot, using a Monte Carlo procedure. We assume that ten children at maximum are born. We also do not distinguish between first and non-first marriages. Mortality has been neglected. In their totality, these three factors keep each other in equilibrium. The limitation to ten children at maximum underestimates the number of children by about 2–3%, the neglect of the difference between first and non-first marriage overestimates the number of children by about 1% and the neglect of mortality results in an overestimate of the number of children by about 2%.

- (vi) Hereafter, the woman's age at which the children are born is determined. The age is fixed by a Monte Carlo procedure. The localization coefficients for births by duration of the marriage in the year 1937 is our starting point. So the duration since the contraction of the marriage is assigned by lot for the year of birth of each of

the children. The year of marriage is known and thus we know the year of birth of each of the children. In this, we limit ourselves to children who are younger than 26 years old in 1947. If the child is between 15 and 26 years old, it is assigned by lot whether the child has ever been married. These children are considered as not living with their parents and are also not matched to their parents. All other children who are younger than 26 years old are assumed to live with their parents. Next, the child's sex is assigned by lot. Now we control whether a boy or girl is available from our set under (iv) above, adjusted for the raise factor. If no child of the appropriate age is available, a child is looked for in the nearest ages. If no children are present at all, the procedure for children is stopped.

- (vii) All characteristics of the family are now assigned to the attributes of the family members: year of birth, year of marriage (which equals year of leaving parental home), number of children, identification numbers, pointers from parents to children and vice versa. And these are put into our file of personal records.

Now family households are determined. Thereafter the households headed by divorced and widowed women are simulated. Given our limited data, it is impossible to find the former husband of divorced women. For this reason steps (i) up to and including step (iv) for married women are not executed. We know the number of children younger than 21 years old and living in female headed one-parent households. The number of divorced women with at least one child is also known, and the same holds for widowed women. We assume that only women younger than 65 years old have children below the age of 21 years old in their household. Dividing the number of widowed women (divorced women) younger than 65 years old by the number of widowed women (divorced women) with at least one child in their household, we get the probability of having at least one child in a household headed by a widowed (divorced) woman. We also assume that only one or two children are present in such households. Now it is possible to determine the probability that two children are present in the household. The procedure is now analogous to that for married people. Instead of the localization coefficients for births by duration of the marriage, we use the localization coefficients by age. Because we cannot determine the father, the children do not get a pointer to their father.

The same process is used for divorced and widowed men, and the same localization coefficients by age are used. However, it is assumed that the father

Table 6. The average probability that the mother and father, respectively, are still alive for some ages of a child, based on the life table for the period 1931-40.

Age child	Probability mother lives	Probability father lives
15	0.9598	0.9560
20	0.9327	0.9247
25	0.8917	0.8817
30	0.8301	0.8005
35	0.7407	0.6920
40	0.6193	0.5516
45	0.4718	0.3851
50	0.3167	0.2248
55	0.1796	0.0994
60	0.0819	0.0311
65	0.0283	0.0071
70	0.0068	0.0005
75	0.0010	0.0000
80	0.0001	0.0000
85 ⁺	0.0000	0.0000

is two years older than the mother, which is about the mean age difference between husbands and wives. We assume that never married people do not have children. Our data are too limited to reconstruct this element. As a consequence, children born out of wedlock and whose mother did not marry between the child's birth and 31 May 1947, have got the 'wrong' mother in the simulation model. It is unknown how many children are involved.

If after all these assignments, some children have not got a mother or father, these children are assigned by lot to a mother. We implicitly assumed that the mother or father is alive for all persons younger than 26 years old. So, we have no orphans younger than 26 years old in our initial population. Another interpretation could be that if a child becomes an orphan, the child is adopted by another family as a foster child. Using the model of Goodman *et al* [13], we can calculate the probability of being an orphan (see Bartlema [3]). For a person aged 15 years, the probability amounts to 0.17%, for a 22 year old 0.46% and for a person aged 25 years 1.18%. So the error involved is rather limited.

Of course, we also want to know who are whose parents, for those who are not living with their parents. These are never married persons older than 26 years, and all persons ever married. We use Bartlema's model (Bartlema [3]) to determine whether no parent, or at least one parent is alive. The average probability of having a mother and a father still alive, respectively, is given in Table 6.

Now, for each of the persons concerned, we assign by lot the mother's age at birth of the person in question (using the localization coefficients for births in 1937 by mother's age). The person's age and the mother's age at birth of that person gives the mother's

age in 1947. Using Bartlema's data it is determined by chance whether the mother is still alive or not. If she is still alive, her marital status is assigned by lot. So, we have determined the mother, and a woman with these characteristics is drawn from our database. The woman's partner, if present, is considered to be the father of the person in question. Next, pointers are interchanged. If the mother was dead, it is determined whether the father is still alive. This father then has to be either divorced or widowed. If the father is still alive, we apply an analogous procedure to find the father. If it is decided that he has also died, the person in question is an orphan and does not have pointers to his or her parents.

The education module

When an individual has passed through the demographic module and is 6 years or older, he or she enters the education submodule. This submodule aims to simulate each individual's passage through the grades (year by year) and through the different fields of education. The input of the required data and the determination of the enrolment probabilities occur at the same moment as the demographic module (see Figure 4). If the concerned person is 6 years old, a decision is made whether or not he or she will enter into primary (from 1986 the third group of the basic education) or special education. If the decision is negative, the person will enter the education process when he or she reaches the age of 7. Then, he or she gets the economic activity 'student'. The choice between special and primary education is also subject to chance. If one enters primary education, one's level of education becomes 1 and if one enters special education this level of education becomes 11.

For each year that the person is in the education process, it is assigned by lot whether the individual is moved up (one's level of education is increased by 1), has to repeat (the level of education does not change) or has left the type of school in question (with or without a certificate, in which case the level of education ends on 8 and 7 respectively). When he or she leaves a particular type of school, a decision is made whether the pupil leaves education entirely or not, and if not, then it is decided which type of education the person will enter. After a pupil has left school he gets another economic activity and he may return to school after some time, and thus re-enter the education submodule. In that case, the probability of part-time education is also considered. If a student returns to school full time, that person again gets the economic activity 'student'. If part-time education is entered, the economic activity does not change, but the attribute 'part-time education' is set

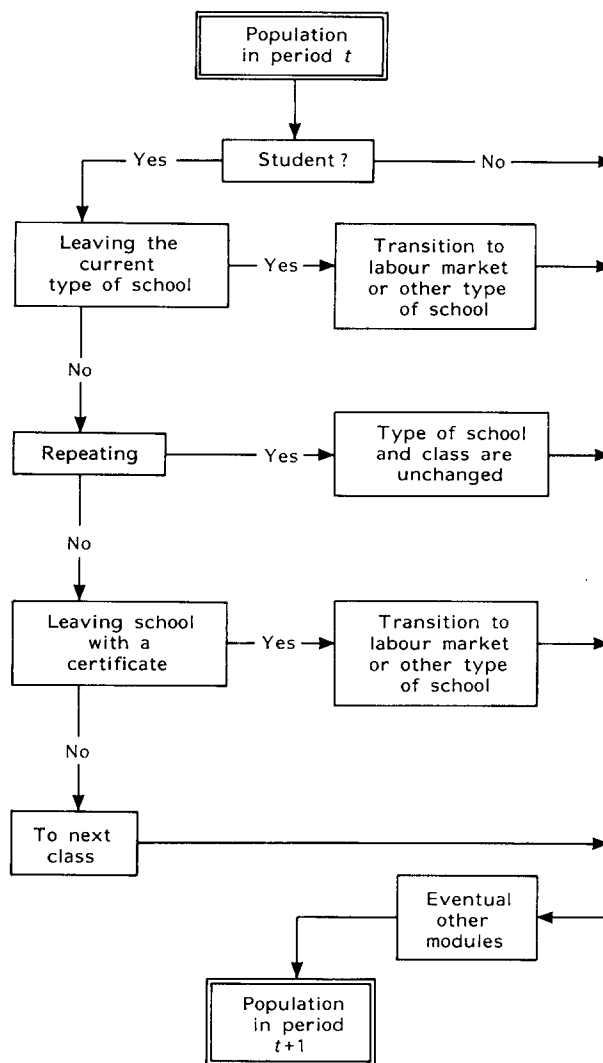


Figure 5. Enrolment through the education submodule.

equal to true. The enrolment and flow probabilities are assumed to be equal for part-time and full-time students. The process is sketched in Figure 5.

The education submodule contains the following routines:

- (i) entering school at age 6 or 7 (primary or special education)
- (ii) repeating, moving up and leaving primary and special education
- (iii) repeating, moving up and leaving secondary general education
- (iv) repeating, moving up and leaving junior vocational training
- (v) repeating, moving up and leaving senior vocational training
- (vi) repeating, moving up and leaving vocational colleges
- (vii) repeating, moving up and leaving university education: and

Table 7. The types of education distinguished in the education submodule and the corresponding levels of education.^a

Type of education	Level of education
Primary education	1-6
Continued primary education (<i>vglo</i>)	111-112
Special education	11-16
Secondary general education (<i>avo</i>)	
Elementary secondary education (<i>lavo</i>)	51-53
First form of secondary school (<i>brugklas</i>)	61-63
Junior secondary schools (<i>mavo/mulo</i>)	141-144
Senior secondary schools (<i>havo</i>)	151-155
Gymnasium/lyceum (<i>vhmo</i> , old style)	131-136
<i>vwo</i>	161-166
Junior vocational training (<i>lbo</i>)	
Junior technical training (<i>lto</i>)	21-24
Junior agricultural training (<i>lao</i>)	31-34
Junior domestic science training (<i>ihno</i>)	41-44
Junior economic and administrative training (<i>leao</i>)	51-54
Junior nautical training (<i>lno</i>)	71-74
Junior training for the retail trade (<i>lmo</i>)	81-84
Senior vocational training (<i>mbo</i>)	
Senior general/vocational training (<i>havo/mbo</i>)	151-155
Senior technical training (<i>mto</i>)	221-224
Senior agricultural training (<i>mao</i>)	231-233
Senior service and health care training (<i>mdgo</i>)	241-242
Senior domestic science training (<i>mhno</i>)	241-242
Senior economic and administrative training (<i>meao</i>)	251-253
Nursery school teachers' training (<i>klos</i>)	261-263
Senior nautical training (<i>mno</i>)	271-272
Senior training for the retail trade (<i>mmo</i>)	281-283
Short senior vocational training (<i>kmbo</i>)	291-292
Vocational colleges	
Technical and nautical colleges (<i>htno</i>)	321-324
Agricultural colleges (<i>hao</i>)	331-334
Health care colleges (<i>hgzo</i>)	341-344
Economic colleges (<i>heo</i>)	351-354
Socio-agogic colleges (<i>hsao</i>)	361-364
Fine art colleges (<i>hko</i>)	371-374
Teachers training colleges (<i>hpo</i>)	381-384
Domestic science colleges (<i>hhno</i>)	391-394
University education	401-406

^a The letters in brackets are acronyms for the Dutch nomenclature.

(viii) re-entering the education module (part time or full time).

Within the above mentioned fields of education different types of school are distinguished. These are given in Table 7. The possible transitions between the different fields of education are given in Figure 6. Children in primary education may only move up, leave the education process, repeat or switch to special education in the first four classes.

The following possibilities exist in the fifth class (with between parentheses the level of education): to repeat (5); to move up into class six (6); to leave the education process (7); or to go over to special education (15), to class one of junior technical training

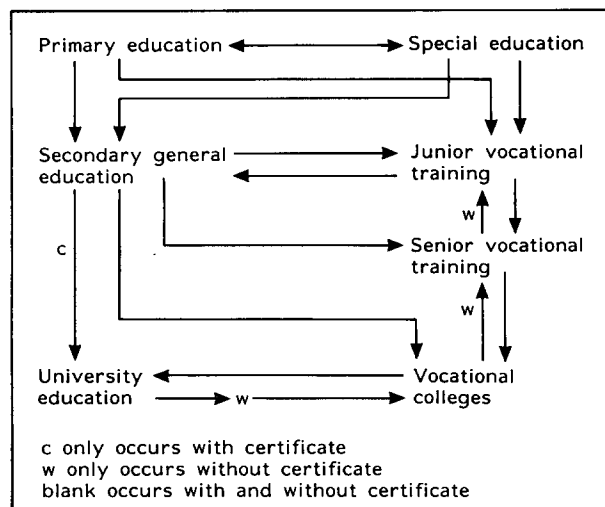


Figure 6. The fields of education and possible transitions.

(21), to junior agricultural training (31), to junior domestic science training (41), to elementary secondary education (51), to junior economic and administrative training (also 51) or to junior secondary school (121).

In class six of primary education, one can repeat (6), leave the educational process (7) or go over to special education (16) or to class one of the continued primary education (111), junior secondary school (121), pre-university education (161 or 131), senior secondary school (151), first form of secondary school (61), junior technical training (21), junior agricultural training (31), junior domestic science training (41), elementary secondary education (51), junior economic and administrative training (51), junior nautical training (71) or junior training for the retail trade (81).⁵⁰

Special education offers the following possibilities during the first five classes: to repeat the same class, to move up to the next class, to go up to primary education and to leave full-time education. In class six of this school type the pupil may repeat, leave full-time education or go up to primary education, continued primary education, junior secondary school, junior technical training (21), junior agricultural training (31), junior domestic science training (41), elementary secondary education (51), junior economic and administrative training (51) or junior training for the retail trade (81).⁵¹

A pupil in junior vocational training is allowed to repeat the same class, to move up to the next class (with exception of the last class), to leave full-time education or to choose another school type. If another

⁵⁰ The transition probabilities are derived from NCBS [33]-[36].

⁵¹ Data and transition probabilities are derived from NCBS [33], [34], [37] and [38].

school type is selected, the possibilities and transition probabilities differ for people with and without a certificate from the school left. The choices are between special education, another type of junior vocational training, senior vocational training and secondary general education. Whether or not a certificate of the junior vocational training has been obtained determines the class into which one moves. The type of junior or senior vocational training entered, depends on the type of junior vocational training left.⁵²

Analogously, a student in secondary vocational training is allowed to repeat the same class, to move up to the next class (with exception of the last class), to leave full-time education or to choose another school type. If another school type is selected, the possibilities and transition probabilities differ for people with and without certificate from the school type left. A student can enter junior vocational training, another type of senior vocational training and vocational colleges. The class into which a student goes is determined by whether or not a certificate of senior vocational training has been obtained. The type of junior or senior vocational training or vocational college entered depends on the type of junior vocational training left.⁵³

In vocational colleges the process is analogous to that in senior vocational training, on the understanding that in case of leaving vocational colleges, the possibilities are senior vocational training, another type of vocational college and university education.⁵⁴

In secondary general education it is again determined whether the student repeats the class in question, moves up to the next class or leaves (with or without certificate) the type of school. After this type of school, a student may go into junior and senior vocational training, vocational colleges, university and other types of general secondary education.⁵⁵

If a student enters university education, then it is assigned by lot whether a degree is obtained, and according to this decision it is determined how many years a student remains at the university. A student obtaining a degree is assumed to leave full-time education or else to go on to new university study or to a vocational college.⁵⁶

The enrolment and outflow probabilities in the model differ between classes in primary and special education since 1950 and for secondary general

education since 1968. In other cases we do not distinguish between classes. The same holds for the outflow probabilities and the distribution of people leaving a type of school without a certificate. The class in which a student goes is dependent on the class left.

The transition probabilities are influenced by social class. Children of low-class birth enter higher education to a much lesser extent than those of higher social class. In the Netherlands several studies have investigated the relationship between social class and school career.⁵⁷ The effect of social class is particularly noticeable at the point when a decision has to be made about post-primary education. We therefore adjusted these 'general' transition probabilities for social class with respect to this choice.⁵⁸ Before the Law on Secondary Schools became operative in 1967, the pupils in class 6 of primary education chose a certain type of school (continuing primary education, junior secondary school, gymnasium/lyceum or some kind of junior vocational training). Since then, a pupil generally chooses in the first instance between a type of vocational training and the *brugklas* (first form of secondary school). If he or she has entered the *brugklas* the pupil later (after one, two or three years) opts for junior secondary school, senior secondary school or gymnasium/lyceum. This implies a correction to the transition probabilities for:

- (i) The transition from the last class of primary education to continued primary education and the first class of some kind of secondary general education and the first class of junior vocational training of various types up to 1968, the year in which the Law on Secondary Schools became operative;
- (ii) the transition from the last class of primary education to *brugklas* 1 or the first class of junior vocational training and the transition from *brugklas* 1 or *brugklas* 2 or *brugklas* 3 to one of the types of secondary school after 1967.

Social class has been defined as the level of education of the father and in absence of a father, the mother's level of education. If both are deceased, then social class is put equal to the lowest level.⁵⁹ We distinguish seven social classes, which correspond with seven levels of education:

- i) The parent has at best finished primary education.
- ii) First stage of secondary general education: the

⁵² Transition probabilities and so on are based on NCBS [33], [34], [44], [45], [49], [50], [53]–[55].

⁵³ Transition probabilities and so on are based on NCBS [33], [34], [44]–[50], [53]–[55].

⁵⁴ The data are derived from NCBS [33], [34], [42]–[45], [49], [50], [52]–[55] and [57].

⁵⁵ The enrolment and flow probabilities are derived from NCBS [33], [34] and [39]–[41].

⁵⁶ Data are derived from NCBS [33], [34], [56] and [57].

⁵⁷ See Boon van Ostade [6], Faasse *et al* [11], Peschar [102] and Tesser [108].

⁵⁸ This part of the research was executed in collaboration with Jan-Bart Waterman; see Waterman [109].

⁵⁹ In other studies, social class has often been made operative as a function of parents' income or occupation or both. Recent research, however, concludes that the parental level of education is the explanatory factor.

Table 8. The first round social class specific transition probabilities by sex and social class for the year 1967(1968).

Social class	from 6 → 1		from <i>brugklas</i> → 2		
	<i>lbo</i>	<i>brugklas</i>	<i>mavo</i>	<i>havo</i>	<i>vwo</i>
Boys					
1	0.586	0.414	0.718	0.162	0.120
2	0.250	0.750	0.444	0.278	0.278
3	0.457	0.543	0.655	0.220	0.125
4	0.135	0.865	0.407	0.187	0.406
5	0.163	0.837	0.407	0.254	0.339
6	0.114	0.886	0.388	0.225	0.387
7	0.050	0.950	0.211	0.316	0.473
Girls					
1	0.611	0.389	0.855	0.113	0.032
2	0.200	0.800	0.583	0.301	0.116
3	0.409	0.591	0.764	0.140	0.096
4	0.001	0.999	0.260	0.318	0.422
5	0.190	0.810	0.559	0.206	0.235
6	0.067	0.933	0.464	0.143	0.393
7	0.041	0.959	0.217	0.174	0.609

Source: Waterman [109], p 34.

parent has entered secondary general education and has reached at minimum a junior secondary certificate or grade 3 of gymnasium or senior secondary school.

- (iii) Junior vocational training: the parent has left the education process with a certificate for this type of school.
- (iv) Second stage of secondary general education: the parent has left gymnasium or senior secondary school with a certificate.
- (v) Senior vocational training: the parent has left the education process with a certificate for this type of school.
- (vi) Vocational colleges or BA examination: the parent has left the education process with a certificate for a vocational college or the BA examination at a university.
- (vii) University: the parent has finished university.

The adaptation of the transition rates for social class is based on the effect of the level of the father's education, also in those cases in which no father is present. The adaptation is based on three cohort surveys which refer to pupils leaving primary school in 1967, 1977 and 1983 respectively.⁶⁰ For the years 1947 up to 1967 we use the 1967 cohort, for the years 1967(1968)⁶¹ up to 1977(1979) we interpolate between the 1967 and 1977 cohort,⁶² and for the years 1977(1979) up to 1983(1985) we interpolate between

⁶⁰ See NCBS [84], [85] and [93].

⁶¹ The first year refers to the transition from primary school to general secondary education and junior vocational training and the year between parentheses refers to the transition from *brugklas* to junior and senior secondary schools or gymnasium.

⁶² It is assumed that the distribution between the types of school available in 1968 corresponded to the distribution in the equivalent types in 1967 (before the 1967 Law on Secondary Schools).

Table 9. The first round social class specific transition probabilities by sex and social class for the year 1983(1985).

Social class	from 6 → 1		from <i>brugklas</i> → 2		
	<i>lbo</i>	<i>brugklas</i>	<i>mavo</i>	<i>havo</i>	<i>vwo</i>
Boys					
1	0.519	0.481	0.661	0.161	0.178
2	0.177	0.823	0.548	0.230	0.222
3	0.435	0.565	0.613	0.215	0.172
4	0.083	0.917	0.380	0.225	0.395
5	0.229	0.771	0.532	0.231	0.237
6	0.091	0.909	0.337	0.233	0.430
7	0.026	0.974	0.221	0.148	0.631
Girls					
1	0.390	0.610	0.695	0.170	0.135
2	0.108	0.892	0.521	0.220	0.259
3	0.283	0.717	0.665	0.187	0.148
4	0.061	0.939	0.322	0.270	0.408
5	0.164	0.836	0.521	0.229	0.250
6	0.066	0.934	0.350	0.286	0.364
7	0.011	0.989	0.202	0.202	0.596

Source: Waterman [109], p 36.

the 1977 and 1983 cohort, whereas from 1983(1985) on, the figures for the 1983 cohort are used. These resulting transition probabilities are called first round social class transition probabilities.

To show the impact of social class on the school choice of pupils, Tables 8 and 9 give the first round social class specific transition probabilities by sex and social class for the years 1967(1968) and 1983(1985). The procedure for the transition from the last class of primary education to *brugklas* and junior vocational training is as follows. In the first step it is decided whether a pupil repeats this class or not. If this is not the case, the level of education is adjusted and a teller, which depends on the pupil's sex and social class, is raised by one. If all potential leavers of the primary school have gone through this process, it is known how many pupils (by sex and social class) will enter junior vocational training or *brugklas*.

The social class specific transition probabilities, used in our calculations, are calculated in the second step. From the general transition probabilities we can derive how many pupils – distinguished by sex – will enter the junior vocational training and how many pupils will enter the *brugklas*. The concerning ratio has to be maintained. The first round social class specific transition probabilities can be used to calculate the expected ratio between numbers of pupils entering vocational training and *brugklas* respectively, which will generally differ from that previously mentioned. The first round social class specific transition probabilities are now adjusted in such a way that the concerning ratio equals the one which is based on the general probabilities. This process is executed by means of the RAS method.⁶³

⁶³ See Bulmer-Thomas [7].

Table 10. Number of persons (1000) by age for the years 1960, 1971, 1981 and 2000; M are model results and O are observations or forecasts (2000).

	1960		1971		1981		2000	
	M	O	M	O	M	O	M	O
0-4	1154	1150	1162	1186	859	888	855	873
5-9	1105	1106	1158	1218	1063	1023	863	891
10-14	1177	1168	1174	1161	1182	1224	855	883
15-19	904	907	1104	1113	1173	1262	883	880
20-24	803	800	1211	1204	1209	1211	928	901
25-29	782	775	939	927	1157	1156	1124	1098
30-34	745	753	821	825	1245	1225	1229	1236
35-39	788	771	786	774	949	933	1204	1257
40-44	662	657	774	759	832	825	1194	1182
45-49	668	662	759	751	784	764	1130	1120
50-54	615	623	669	662	757	737	1167	1141
55-59	561	555	639	627	711	714	874	857
60-64	464	471	565	571	591	605	723	723
65-69	370	381	482	485	542	542	659	646
70-74	283	299	374	370	435	448	559	552
75-79	199	193	251	257	314	329	457	456
80+	167	154	285	228	405	323	536	516
Total	11448	11417	13154	13119	14209	14209	15240	15213

The third step contains the use of the resulting social class specific transition probabilities for the pupils who left primary school. The transition from primary school to the next school-type before 1968 and the transition from *brugklas* to other kinds of secondary schooling is carried out in the same way.

Educational attributes are adjusted for each year in schooling (part time or full time), unless a class is repeated. The level of education is adjusted and if a student leaves school the boolean for part-time education is set equal to false and end of education is set equal to the current year.

If a student has left the education process, then he or she can be considered as a candidate for a re-entrance to education, full time or part time. Using the observed numbers of re-entrances in the past, and assuming that the re-entrance probabilities do not change in the future, it is assigned by lot who re-enters the education process.⁶⁴ The probability of being drawn is dependent of the level of education, age and sex. If a person is assigned by lot to re-enter full-time education, the 'end of education' is set equal to zero, the level of education becomes the new type of school entered and the attribute 'economic activity' becomes student. If one school is re-entered part time, the last attribute is not adjusted, but the boolean standing for part-time education becomes true.

Unfortunately, we do not know the level of education at the start of our simulation (1947). The first available information on the level of education of the Dutch population was provided by the 1971

Census.⁶⁵ The transition and enrolment probabilities are generally known from about 1967. For the period 1947–66 we have some information from specific publications of the Central Bureau of Statistics.⁶⁶ The required transition and enrolment probabilities for the other years are determined by interpolation and extrapolation. In that way it was possible to derive the level of education of the initial population in 1947 by means of simulation. For the period 1985–2000 we here use the probabilities for the year 1984. The flow probabilities (out of education) are calculated taking account of mortality and emigration.

If a person has immigrated in the current year and he or she has already reached school age, his or her level of education is simulated up to the concerning year. If we have to deal with a new immigrant, he or she flows twice through the education process and the lowest level of education is assigned to this person. This procedure results in a level of education that is quite near to the real level.

Simulation results

In this section we will give some simulation results and compare them, if possible, with observations from official statistical sources.

Table 10 gives the number of people by age for the years considered, starting the simulation in 1947. We can see that the total Dutch population is simulated very well. The error in the total number amounts to less than 0.4% for the years up to and including 1981.

⁶⁴ See NCBS [33]–[57].

⁶⁵ See NCBS [79].

⁶⁶ See NCBS [70].

For the year 2000 the deviation with respect to the official forecast of the Netherlands Central Bureau of Statistics (NCBS) is also very limited, being less than 0.2%.

However, the deviations are larger for the different age groups. These deviations originate from demographic processes that form the basis of the simulation. Births are simulated quite well. There is a slight overestimation in the 1950s and 1960s. On the other hand, mortality has been overestimated, especially in the period 1947–70. This can probably be attributed to the assumptions concerning the distribution of people in the oldest age groups. The distribution has to be adjusted. The marriage and divorce process are correctly simulated. The deviations are very small. The total number of emigrants and immigrants during the whole period has been simulated accurately, but there are periods with an overestimation that are counterbalanced by periods with an underestimation. This is caused by the way in which the probabilities of remigration are considered. For some years we use observations for these probabilities, for other years interpolations are used.

Table 11. Mean number of persons (1000) by age for the year 1981; M are model results and O are observations. The standard deviations are in parenthesis.

	M	O	σ
0-4	889	888	(27)
5-9	1027	1023	(37)
10-14	1232	1224	(18)
15-19	1251	1262	(13)
20-24	1215	1211	(25)
25-29	1171	1156	(21)
30-34	1254	1225	(11)
35-39	954	933	(11)
40-44	838	825	(9)
45-49	772	764	(7)
50-54	748	737	(12)
55-59	708	714	(19)
60-64	601	605	(11)
65-69	539	542	(10)
70-74	445	448	(9)
75-79	333	329	(16)
80+	337	323	(13)
Total	14315	14209	(99)

Another problem with migration is that we use a remigration probability which is the same for each age and independent of marital status. This assumption is too crude and can cause important deviations. This holds especially for the years after 1975. From that year family reunion had played an increasingly important role. This is also influenced by the large migration flows of people to and from Surinam in the late 1970s.

To give an indication of the variance of the simulation process, we made a simulation for the period 1971–81, which was repeated ten times. The results are given in Table 11. The mean values for the different ages is always near to the observation. Generally, the difference amounts to less than 1%. The largest differences hold for the youngest and oldest age group. The standard deviation, σ , is relatively low.

Table 12 shows us the population by marital status. The deviations for never married and married people are very limited up to and including 1981: less than 1%. In 2000 the deviation with respect to the forecast of the NCBS is about 4%. This is a consequence of the problems discussed above concerning migration. The number of divorced people before 1981 has been somewhat underestimated, which is probably caused by the size of the initial population, which numbered only 10 000 persons. The deviation for widowed people is due to an overestimation of mortality in the higher age groups.

Table 13 contains the results for the 1971–81 simulation. The mean values are generally better than the one-off simulation result in Table 13, especially for those groups with relatively few people (divorced and widowed persons).

Table 14 gives some information on the number of

Table 13. Mean simulated (M) and observed (O) number of persons and standard deviation (1000) by marital status 1971–81.

	M	O	σ
Never married	6282	6235	(69)
Married	6873	6866	(68)
Divorced	334	343	(16)
Widowed	789	765	(36)

Table 12. Number of persons (1000) by marital status; M are model results and O observations or forecasts (2000).

	1960		1971		1981		2000	
	M	O	M	O	M	O	M	O
Never married	5772	5731	6105	6114	6232	6235	6841	6678
Married	5076	5103	6279	6253	6765	6866	6202	6434
Divorced	89	88	142	119	364	343	1188	1162
Widowed	511	495	628	631	848	765	1009	939

people cohabiting. We see a rapid growth between on the one hand the years 1971 and 1981 and on the other hand 1981 and 2000. The real number of cohabiting persons is not known exactly for these years, but the simulated number for the year 1981 is close to the result of a survey held in 1982. For the years 1981 and later on we assumed that the probability of starting a first cohabitation equals the 1981 level plus the difference in first marriage rates in 1981 and the year under investigation. This means that the decrease in first marriage rates, as expected by the Netherlands Central Bureau of Statistics, are compensated for by higher rates for the start of a cohabitation. The table also gives the standard deviation for the 1971–81 simulations. This standard deviation is relatively larger than in the foregoing tables. It is clear that it rises relatively when the number of persons involved declines.

Table 15 gives the distribution of households by size. The distributions for the years 1971 and 1981 are close to reality (see Nollen-Dijcks [95], p 65). From a report drawn up by the Ministry of Housing [32] we know that the population of one- and two-person households was 0.53 in 1982; our simulation yielded 0.56 for the year 1981. For three- and four-person households the proportions were 0.36 and 0.33 respectively. For households with five or more

persons the proportions were 0.11 and 0.10 respectively. The difference is mainly caused by the fact that in our model only one-family households are considered and institutional households are excluded. When we assume that the mean multifamily household and mean institutional household counts for two and two-and-a-half one-family households respectively, then the deviation between the simulated number of households and the observed number amounts to less than 0.5% in 1981, the simulated number being the larger of the two.

For the year 1960, however, we have an overestimation of the simulated number of single-person households. In 1971 the overestimation is still present, but to a lesser extent. This problem is, of course, also caused by the fact that we ignored multifamily and institutional households. When we adapt the distribution for this kind of household, then the simulated distribution is very close to reality. The exclusion concerned leads to an underestimate of the average household size by about 12% in the period 1947–81. For the year 1981 the standard deviation for the 1971–81 simulation is added again.

During the entire period we establish a decline in the percentage of households with more than two persons; this holds especially for households with six or more persons. The proportion of single-person households is rather constant between 1971 and 1981, which also holds for two-person. The growth of both one- and two-person households between 1981 and 2000, however, is considerable. We also remark a continuous decrease in the average household size. The decline amounted to 16.3% between 1960 and 1981 and we may expect a further decline by about 19% between 1981 and 2000.

The mean number of persons per household by age of the head of household is given in Figure 7. The simulated distributions for the years 1960, 1971 and 1981 deviate only a little from the observed curves (see Nollen-Dijcks [95], p 66). The differences are once again a consequence of the exclusion of

Table 14. Number of unmarried cohabiting persons (1000) by marital status (1981: the mean). Standard deviations (1981 only) are in parentheses.

	1971	1981	2000
Never married	53	309 (22)	1327
Divorced	35	98 (8)	260
Widowed	2	22 (3)	41
Total	90	429 (26)	1628

Table 15. The simulated distribution of households by size (1981: the mean). The standard deviation (1981 only) is in parentheses.

Number of persons in household								Mean household size	
	1	2	3	4	5	6	7+	Simulated	Observed
1960	0.29	0.21	0.15	0.12	0.08	0.06	0.08	3.06	3.50
1971	0.29	0.24	0.16	0.15	0.09	0.05	0.03	2.80	3.15
1981	0.30 (0.007)	0.26 (0.009)	0.15 (0.006)	0.18 (0.007)	0.07 (0.002)	0.02 (0.002)	0.01 (0.001)	2.56 (0.013)	2.89
2000	0.41	0.32	0.12	0.11	0.03	0.01	0.00	2.07	

Source: Ministry of Housing, [32], p 42.

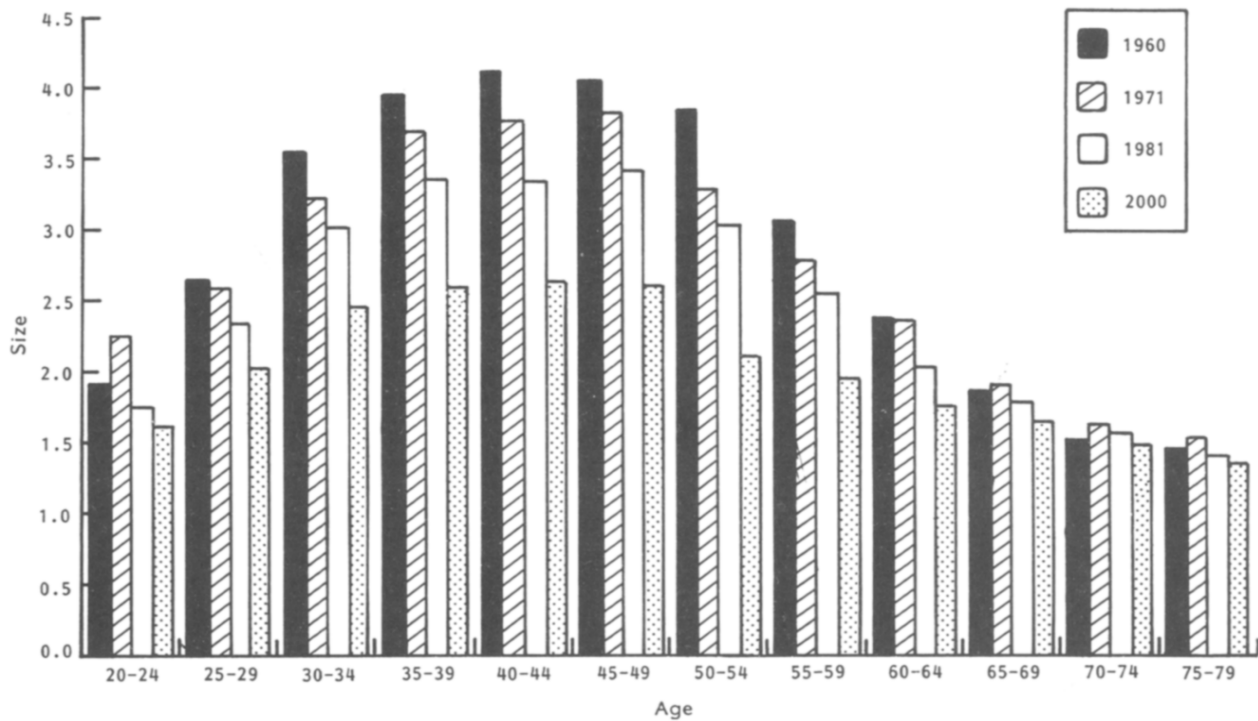


Figure 7. Average household size by age of head.

multifamily and institutional households. The trend which we saw in Table 14 can be seen here, too: in each age group, excluding the very aged, the average number of household members declines between 1971 and 2000.

The average household size by age of the head of the household and its standard deviation for the year 1981, based on ten repeats for the period 1971–81, is given in Figure 8.

Of course, the picture for the mean is comparable with the concerning size for the year 1981 in Figure 7. Most interesting is the relatively small standard deviation.

Summarizing the demographic module, we can conclude that the microsimulation approach is a useful tool for the simulation of household structures of the Dutch population, both for the past and the future. The simulation results, covering a 34-year period (1947–81) and an initial population of only 10 000 persons are very good. A relatively small number of repeats (ten times) leads already to acceptable standard deviations during the period 1971–81, so the method seems very appropriate. Because micro-units (individuals and households) form the point of departure for the model, individuals can be followed through their life. From the demographic point of view, this is a very important issue. It creates the possibility of deriving the distribution of many

processes. For example, it is possible to find out in which household position and for what expected period of time children whose parents divorced in the period 1965–69 will remain (and also its variance). And this can provide an answer to the question of what percentage of these children will stay in a one-parent family for less than one year, what percentage from one to two years, and so on. So demographic analysis offers more scope than current techniques.

Because information on the education level of the Dutch population is limited, and the available sources differ in definition of the levels, we present only the results of the simulation of the outflow for the period 1971–79. Since detailed information is only available for full-time education, we made our simulations without part-time flows. Because of the very small number of children leaving the education process during their primary or special education, these have been left out of consideration. We have also omitted the outflows from university without a certificate. Data on this group are not available during a number of years as a consequence of the tuition fee boycott in the Netherlands during this period, so that a part of this outflow stream was not registered. The comparison between observation and simulation results (based on 10 repeats) can be found in Table 16. In this table, the outflow has been given as a proportion of the nine

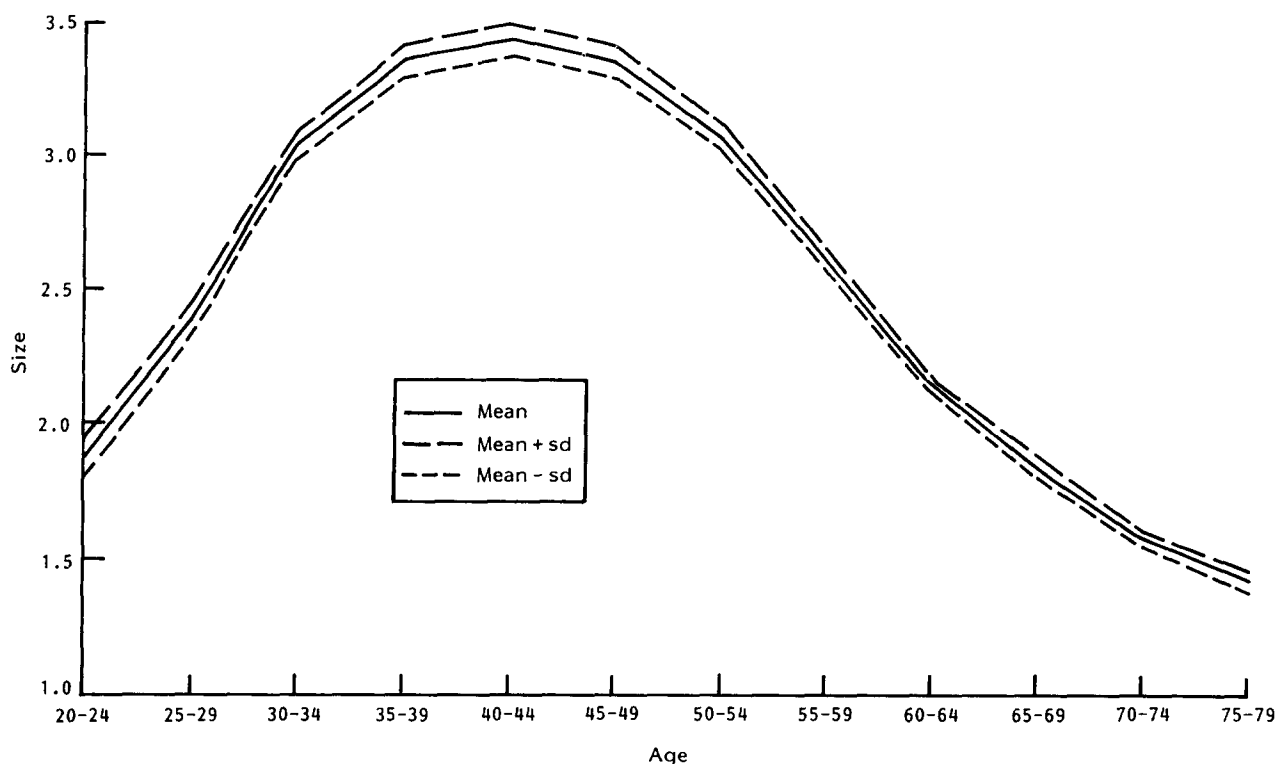


Figure 8. Average household size and standard deviation, 1981.

Table 16. Full-time outflow (relative proportions) from school by type of education and with certificate (c) or without (w) for the period 1971-79, observed (O) and simulated mean (S) and standard deviation (σ).

Type education		Males			Females		
		O	S	σ	O	S	σ
Junior vocational training	w	0.099	0.099	0.008	0.094	0.104	0.012
	c	0.268	0.289	0.014	0.178	0.190	0.009
Senior vocational training	w	0.066	0.064	0.010	0.069	0.079	0.007
	c	0.088	0.088	0.010	0.181	0.166	0.010
Vocational colleges	w	0.058	0.058	0.007	0.031	0.030	0.007
	c	0.095	0.081	0.007	0.079	0.069	0.007
Secondary general education	w	0.082	0.097	0.010	0.097	0.100	0.008
	c	0.177	0.158	0.009	0.254	0.242	0.011
University	c	0.067	0.066	0.003	0.018	0.018	0.002

types of education under investigation. The mean absolute deviation amounts to 0.007 for males and 0.008 for females respectively, whereas the mean absolute percentage deviation amounts to 6.2% and 7.1% respectively. Given the relatively low number of persons involved in the simulation during the period concerned (on average 921 men and 867 women in our sample) these results are satisfactory. The relatively largest deviations are found for vocational colleges and secondary general education, both with certificate and for males. Also the deviation in the case of junior vocational training with certificate is considerable, both for men and women.

Apart from the relatively small size of the simulation

sample, the deviation is also a consequence of the fact that our initial population in 1971 does not exactly cover the real population in that year. Our simulation actually starts in 1947 and not in 1971. We did not make a correction for this and the simulated population in 1971 was treated as the initial population for our simulation of the educational outflow. As a result, the number of persons aged about 10 to 15 years was overrepresented at the start of our simulation in comparison with the other age groups relevant to us. This leads to a larger outflow out of junior vocational training with certificate and explains partly also the other relatively large deviations for some types of education, as mentioned before. So, the deviations

are not only the result of the chosen (random) microsimulation approach, but also a consequence of an error in the initial population.

Conclusions

In this paper, two elements of a microsimulation model are presented: the demographic module and the education submodule. Here we use them to generate Dutch household structures and the level of education of the Dutch population. Until now household composition has only been known for a number of specific years. The model presented simulates household composition for the period from 1947 onwards. This enables us to study problems which hitherto could only be analysed incompletely or not at all. We should mention in this respect the issue of family histories and the distribution by events. In this way microsimulation expands the state space of the orthodox increment-decrement life table modelling technique, methods applied hitherto to generate family distributions. In this way, we provide data which can be used as exogenous inputs for the studies mentioned in the introduction.

The simulation results presented above are encouraging. The established drawbacks of the demographic module can be solved in the next version of the model. In that version a number of other factors will also be included (transition to and from institutions, adoption and multifamily households). The education module is also satisfactory.

In comparison with other microsimulation models, our model has the advantage that migration and unmarried cohabitation are explicitly modelled and that standard deviations are given (for the year 1981). In none of the demographic microsimulation models known to us have these events been modelled, or standard deviations given. Migration is usually accounted for by weighting the microdata, whereas unmarried cohabitation is not considered.

Another important advantage stems from the fact that the model is integrated into a comprehensive socioeconomic model. In time it will be possible to interact with other modules in the model (eg income, social security, labour market), so that very detailed subpopulations can be derived. The extension of the demographic model by the addition of socioeconomic variables can then be carried out easily by introducing submodules into the demographic (main) module. Using the example in the foregoing section, this means that we can look at the income position of single-parent families, which could affect the school career of their children, and so on. In the long run, demographic and educational decisions will also be endogenized.

The organization of the model guarantees consistency. It is not necessary to introduce limiting conditions, as for example is the case in macrosimulation models (see eg Keilman and van Dam [25], who have to use a consistency algorithm in their demographic model).

A last important advantage of the model is that it is also relatively simple to apply it to other countries or to apply it to regions or cities; the most important modification has to take place in the input module, whereas in the demographic module changes will be minor.

References

- 1 R.J.M. Alessie, F.A. von Dewall, A. Kapteyn and E.R. Koster, 'De invloed van demografie op bestedingen en besparingen', in Koninklijke Vereniging voor de Staathuishoudkunde, ed, *Demografische veranderingen en economische ontwikkelingen*, Stenfert Kroese, Leiden/Antwerp, 1987.
- 2 K.J. Arrow, 'Microdata simulation: current status, problems, prospects', in R.H. Haveman and K. Hollenbeck, eds, *Microeconomic Simulation Models for Public Policy Analysis*, Vol 2, Academic Press, New York, 1980.
- 3 J. Bartlema, *Developments in Kinship Networks for the Aged in the Netherlands*, PhD dissertation, Tilburg University, Tilburg, 1987.
- 4 J.M. Bekkering, Y.K. Grift and J.J. Siegers, *Belastingen premieheffing en de arbeidsmarktparticipatie door gehuwde vrouwen: een econometrische analyse*, Department of Social Affairs and Employment, The Hague, 1986.
- 5 R.L. Bennett and B.R. Bergmann, *A Microsimulated Transactions Model of the United States Economy*, Johns Hopkins University Press, Baltimore, MD/London, 1986.
- 6 A.H. Boon van Ostade, 'Het milieu en de overgang van het basis- naar het voortgezet onderwijs in Nederland', *Sociologische Gids*, Vol 19, 1972, pp 211-219.
- 7 V. Bulmer-Thomas, *Input-output Analysis in Developing Countries: Sources, Methods, and Applications*, Wiley, Chichester, 1982.
- 8 S.B. Caldwell, 'How can microsimulation help develop better theories?', in D.A. Wolf, ed, *Recent Developments in Microsimulation: An International Perspective*, Urban Institute Press, Washington, DC, forthcoming.
- 9 M. Clarke, 'Demographic processes and household dynamics: a microsimulation approach', in R. Woods and P. Rees, eds, *Population Structures and Models: Developments in Spatial Demography*, Allen and Unwin, London, 1986.
- 10 J. Devine and R. Wertheimer, *Aging Techniques Used by the Major Microsimulation Models*, Working Paper 1453-01, The Urban Institute, Washington, DC, 1981.
- 11 J.H. Faasse, B. Bakker, J. Dronkers and H. Schijf, 'Voor en na de Mammoet, vergelijking van individuele schoolloopbanen in het voortgezet onderwijs in Noord-Brabant voor en na de invoering van de Mammoetwet', *Mens en Maatschappij*, No 1, 1986, pp 75-94.
- 12 H.P. Galler, 'Microsimulation of household formation and dissolution', in N.W. Keilman, A.C. Kuijsten and A.P. Vossen, eds, *Modelling Household Formation and Dissolution*, Oxford University Press, Oxford, 1988, pp 139-159.
- 13 L. Goodman, N. Keyfitz and T. Pullum, 'Family formation and the frequency of various kinship relationships', *Theoretical Population Biology*, Vol 5, 1974, pp 1-27.
- 14 M. Greenberger, M. Crenson and B. Crissey, *Models in the Policy Process: Public Decision Making in the Computer Era*, Russel Sage Foundation, New York, 1976.
- 15 H.-J. Hansen, E. Klein and D. Mannel, *Das makroökonomische*

- Modell des Sfb3; Version 5.0, Sonderforschungsbereich 3 Working Paper No 238, J.W. Goethe University, Frankfurt and Mannheim University.
- 16 B. Hayes, *Alternative Methodologies for Life-Cycle Simulations of the Incidence of Taxation and Benefit Systems*, Discussion Paper TIDI/137, London School of Economics, London, 1989.
 - 17 H. Heida and H. Gordijn, *PRIMOS Huishoudenmodel; analyse en prognose van de huishoudensontwikkeling in Nederland*, Ministry of Housing, The Hague, 1985.
 - 18 O. Hellwig, *Micromodelling the Australian Household Sector: A Proposal*, Department of Statistics and Econometrics, Darmstadt Technical University, Darmstadt, 1988.
 - 19 O. Hellwig, *Probleme und Lösungsansätze bei der Datenbeschaffung und Parameterschätzung für Mikro-simulationsmodelle der privaten Haushalte*, PhD dissertation, Darmstadt University, Darmstadt, 1988.
 - 20 O. Hellwig, *Modelling and Software Aspects of Demographic Microsimulation*, Darmstadt University, Darmstadt, 1988.
 - 21 G. Hendricks, R. Holden and J. Johnson, *A File of Simulated Family and Earnings Histories Through the Year 2000*, Working Paper No 985-3, The Urban Institute, Washington, DC, 1977.
 - 22 P. Hoschka, 'Requisite research on methods and tools for microanalytic simulation models,' in G. Orcutt, J. Merz and H. Quinke, eds, *Microanalytic Simulation Models to Support Social and Financial Policy*, North Holland, Amsterdam, 1986.
 - 23 A. Kapteyn and I. Woittiez, *Preference Interdependence and Habit Formation in Family Labor Supply*, Research Memorandum FEW 262, Tilburg University, Tilburg, 1987.
 - 24 N.W. Keilman, 'Why microsimulation - why not? A discussion of simulation approaches in the field of family and household demography', in D.A. Wolf, ed, *Recent Developments in Microsimulation: An International Perspective*, Urban Institute Press, Washington, DC, forthcoming.
 - 25 N.W. Keilman and J. van Dam, 'LIPRO - een dynamisch model ter vooruitberekening van huishoudens', *Bevolking en Gezin*, No 2, 1986, pp 35-61.
 - 26 A. Kirman, 'The intrinsic limits of modern economic theory', *Economic Journal*, Vol 99, 1989, pp 126-139.
 - 27 A. Klevmarcken, 'On estimation and other problems of statistical inference in the micro simulation approach', in B. Bergmann, G. Eliasson and G. Orcutt, eds, *Micro Simulation: Models, Methods and Applications*, IUI Conference Reports No 1, Industrial Institute for Economic and Social Research, Stockholm, 1980.
 - 28 A. Klevmarcken, 'There is a need for applied micro-econometrics: survey research about the household', in G. Eliasson and H. Fries, eds, *Microeconomics: IUI Yearbook 1982-1983*, Industrial Institute for Economic and Social Research, Stockholm, 1983.
 - 29 C.M. van't Klooster-van Wingerden, *Huishoudensvorming en samenlevingsvormen*, Census Reports 1971, No 11, Staatsuitgeverij, The Hague, 1979.
 - 30 L.Th. van Leeuwen, *Ontwikkelingsfasen van het gezin*, Census Reports 1971, No 9, Staatsuitgeverij, The Hague, 1980.
 - 31 J. Merz, *Microsimulation: A Survey of Principles, Developments and Applications with Focus on the Static Case and the Static Sfb3-Microsimulation Model as an Example*, Working Paper No 268, J.W. Goethe University, Frankfurt, 1988.
 - 32 Ministry of Housing, *Toetsing en actualisering trendrapport volkshuisvesting*, The Hague, 1985.
 - 33 NCBS, *Overgangen binnen het onderwijs en intrede in de maatschappij: onderwijsmatrix*, Staatsuitgeverij, The Hague, several years.
 - 34 NCBS, *Statistiek van de buitenlandse migratie*, De Haan NV, Utrecht, De Haan NV, Zeist and Staatsuitgeverij, The Hague, several years.
 - 35 NCBS, *Statistiek van het gewoon- en voortgezet lager onderwijs*, Staatsuitgeverij, The Hague, several years.
 - 36 NCBS, *Statistiek van het gewoon lager onderwijs*, Staatsuitgeverij, The Hague, several years.
 - 37 NCBS, *Statistiek van het buitengewoon lager onderwijs*, Staatsuitgeverij, The Hague, several years.
 - 38 NCBS, *Statistiek van het buitengewoon onderwijs*, Staatsuitgeverij, The Hague, several years.
 - 39 NCBS, *Statistiek van het vwo, havo en mavo; in-, door- en uitstroom van de leerlingen*, Staatsuitgeverij, The Hague, several years.
 - 40 NCBS, *Statistiek van het voorbereidend hoger en middelbaar onderwijs*, Staatsuitgeverij, The Hague, several years.
 - 41 NCBS, *Statistiek van het uitgebreid lager onderwijs*, Staatsuitgeverij, The Hague, several years.
 - 42 NCBS, *Statistiek van het beroepsonderwijs; sociaal-pedagogisch onderwijs*, Staatsuitgeverij, The Hague, several years.
 - 43 NCBS, *Statistiek van het beroepsonderwijs; kunstonderwijs*, Staatsuitgeverij, The Hague, several years.
 - 44 NCBS, *Statistiek van het beroepsonderwijs; economisch en administratief onderwijs, middenstandsonderwijs en lager algemeen voortgezet onderwijs*, Staatsuitgeverij, The Hague, several years.
 - 45 NCBS, *Statistiek van het beroepsonderwijs; technisch onderwijs, nautisch onderwijs*, Staatsuitgeverij, The Hague, several years.
 - 46 NCBS, *Statistiek van het beroepsonderwijs; opleidings-scholen voor kleuterleidsters, pedagogische academies*, Staatsuitgeverij, The Hague, several years.
 - 47 NCBS, *Statistiek van het beroepsonderwijs; opleidings-scholen voor onderwijzers*, Staatsuitgeverij, The Hague, several years.
 - 48 NCBS, *Statistiek van het beroepsonderwijs; opleidings-scholen voor kleuterleidsters, opleidingsscholen voor onderwijzers*, Staatsuitgeverij, The Hague, several years.
 - 49 NCBS, *Statistiek van het beroepsonderwijs; agrarisch onderwijs*, Staatsuitgeverij, The Hague, several years.
 - 50 NCBS, *Statistiek van het beroepsonderwijs; landbouwonderwijs*, Staatsuitgeverij, The Hague, several years.
 - 51 NCBS, *Statistiek van het beroepsonderwijs; beroepsbegeleidend onderwijs, leerlingenwezen*, Staatsuitgeverij, The Hague, several years.
 - 52 NCBS, *Statistiek van het beroepsonderwijs; nieuwe lerarenopleidingen*, Staatsuitgeverij, The Hague, several years.
 - 53 NCBS, *Statistiek van het beroepsonderwijs; huishouden niverheidsonderwijs*, Staatsuitgeverij, The Hague, several years.
 - 54 NCBS, *Statistiek van het beroepsonderwijs; land- en tuinbouwonderwijs*, Staatsuitgeverij, The Hague, several years.
 - 55 NCBS, *Statistiek van het niverheidsonderwijs*, Staatsuitgeverij, The Hague, several years.
 - 56 NCBS, *Statistiek van het wetenschappelijk onderwijs*, Staatsuitgeverij, The Hague, several years.
 - 57 NCBS, *Statistiek van het hoger onderwijs*, Staatsuitgeverij, The Hague, several years.
 - 58 NCBS, *Statistiek van de buitenlandse migratie 1948 en 1949*, Staatsdrukkerij- en uitgeverijbedrijf, The Hague, 1950.
 - 59 NCBS, *12e Volkstelling, annex woningtelling 31 Mei 1947. Serie A Deel 4 Statistiek der bestaande huwelijken en van de vruchtbaarheid dezer huwelijken*, De Haan NV, Utrecht, 1951.
 - 60 NCBS, *12e Volkstelling, annex woningtelling 31 Mei 1947. Serie A Deel 3 Woning- en gezinstelling*, De Haan NV, Utrecht, 1951.
 - 61 NCBS, *Verzorging van bejaarden in tehuizen naar de toestand op 1 oktober 1950*, De Haan NV, Utrecht, 1951.
 - 62 NCBS, *12e Volkstelling, annex Woningtelling 31 Mei 1947, Serie A Rijks- en provinciale cijfers, Deel 2 Beroepstelling*, De Haan NV, Utrecht, 1952.
 - 63 NCBS, *Bevolking van Nederland. Leeftijd en geslacht, 1900-1952. Leeftijd, geslacht en burgerlijke staat, 1947-1952*, De Haan NV, Utrecht, 1953.
 - 64 NCBS, *Statistiek van de buitenlandse migratie 1950-1952*, De Haan NV, Utrecht, 1953.
 - 65 NCBS, *12e Volkstelling, annex woningtelling 31 Mei 1947. Serie A Deel 1 Belangrijkste uitkomsten der eigenlijke Volkstelling*, De Haan NV, Utrecht, 1954.
 - 66 NCBS, *Statistiek van de loop der bevolking 1938-1954*, De Haan NV, Utrecht, 1955.

- 67 NCBS, *Statistiek van de buitenlandse migratie 1955-1956*, De Haan NV, Zeist, 1958.
- 68 NCBS, *Echtscheidingen in Nederland, 1900-1957*, De Haan NV, Zeist, 1958.
- 69 NCBS, *Statistiek van de buitenlandse migratie 1959-1960*, De Haan NV, Zeist, 1962.
- 70 NCBS, *Overgangen binnen het onderwijs en intrede in de maatschappij; 1936, 1956 en 1961-1966 'Matrix'*, Staatsuitgeverij, The Hague, 1967.
- 71 NCBS, *Statistiek van de bejaardenoorden 1965*, Staatsuitgeverij, The Hague, 1968.
- 72 NCBS, *Statistiek van de buitenlandse migratie 1965-1966*, Staatsuitgeverij, The Hague, 1968.
- 73 NCBS, *Bevolking van Nederland naar geslacht, leeftijd en burgerlijke staat 1930-1969*, Staatsuitgeverij, The Hague, 1970.
- 74 NCBS, *Statistiek van de bejaardenoorden 1970*, Staatsuitgeverij, The Hague, 1973.
- 75 NCBS, *Statistiek van de buitenlandse migratie 1970-1971*, Staatsuitgeverij, The Hague, 1974.
- 76 NCBS, *Buitenechtelijke geboorte 1840-1973*, Staatsuitgeverij, The Hague, 1975.
- 77 NCBS, *Echtscheidingen in Nederland 1900-1974*, Staatsuitgeverij, The Hague, 1976.
- 78 NCBS, *Statistiek van de buitenlandse migratie 1975*, Staatsuitgeverij, The Hague, 1977.
- 79 NCBS, *14e Algemene Volkstelling annex woningtelling 28 Februari 1971, Deel 5A Het genoten onderwijs en het volgen van onderwijs*, Staatsuitgeverij, The Hague, 1980.
- 80 NCBS, *Statistiek van de bejaardenoorden 1977*, Staatsuitgeverij, The Hague, 1980.
- 81 NCBS, *14e Algemene Volkstelling 28 februari 1971. Serie A, deel 2, Bestaande huwelijken en vruchtbaarheid van gehuwde en gehuwd geweest zijnde vrouwen*, Staatsuitgeverij, The Hague, 1981.
- 82 NCBS, *Prognose van de bevolking van Nederland na 1980. Deel 1, Uitkomsten en enkele achtergronden*, Staatsuitgeverij, The Hague, 1982.
- 83 NCBS, *Statistiek van de buitenlandse migratie 1980*, Staatsuitgeverij, The Hague, 1982.
- 84 NCBS, *Schoolloopbaan en herkomst van leerlingen bij het voortgezet onderwijs, deel 2; cohort 1977, schoolkeuze*, Staatsuitgeverij, The Hague, 1982.
- 85 NCBS, *Schoolkeuze en schoolloopbaan bij het voortgezet onderwijs, cohort 1964/65*, Staatsuitgeverij, The Hague, 1982.
- 86 NCBS, *Vijftientig jaren statistiek in tijdreeksen*, Staatsuitgeverij, The Hague, 1984.
- 87 NCBS, *Onderzoek gezinsvorming*, Staatsuitgeverij, The Hague, 1984.
- 88 NCBS, *Bevolkingsprognose voor Nederland 1984-2035*, Staatsuitgeverij, The Hague, 1985.
- 89 NCBS, 'Hoe duurzaam in samenwonen', *Maandstatistiek van de Bevolking*, Vol 33, No 1, 1985, pp 7-8.
- 90 NCBS, 'Buitenechtelijke levendgeborenen naar burgerlijke staat van de moeder, 1975-1984', *Maandstatistiek van de Bevolking*, Vol 34, No 2, 1986, pp 11-13.
- 91 NCBS, *Statistiek van de bejaardenoorden 1984*, Staatsuitgeverij, The Hague, 1987.
- 92 NCBS, *Schoolloopbaan en herkomst van leerlingen bij het voortgezet onderwijs, SLVO-cohort 1982, deel 1: schoolkeuze*, Staatsuitgeverij, The Hague, 1988.
- 93 J.H.M. Nelissen, *Microsimulation of Household and Education Dynamics*, Discussion Paper No 89/3, Growth Dynamics University Institute, Erasmus University, Rotterdam, 1989.
- 94 J.H.M. Nelissen and A.P. Vossen, *Demografische ontwikkelingen en het sociale zekerheidssysteem*, Department of Social Affairs and Employment, The Hague, 1985.
- 95 J.B. Nollen-Dijcks, 'Huishoudens en personen in huishoudens, 1960-1981', *Supplement Sociaal Economische Maandstatistiek*, No 10, 1985, pp 63-85.
- 96 G.H. Orcutt, 'A new type of socioeconomic system', *Review of Economics and Statistics*, Vol 39, 1957, pp 116-123.
- 97 G.H. Orcutt, 'Simulation of economic systems', *American Economic Review*, Vol 50, 1960, pp 893-907.
- 98 G.H. Orcutt, S. Caldwell and R. Wertheimer II, *Policy Exploration Through Microanalytic Simulation*, Urban Institute, Washington, DC, 1976.
- 99 G. Orcutt and A. Glazer, 'Microanalytic modelling and simulation', in B. Bergmann, G. Eliasson and G. Orcutt, eds, *Micro Simulation: Models, Methods and Applications*, IUI Conference Reports No 1, Industrial Institute for Economic and Social Research, Stockholm, 1980, pp 121-176.
- 100 G. Orcutt, A. Glazer, R. Harris and R. Wertheimer II, 'Microanalytic modelling and the analysis of public transfer policies', in R.H. Haveman and K. Hollenbeck, eds, *Microeconomic Simulation Models for Public Policy Analysis*, Vol 1, Academic Press, New York, 1980, pp 81-106.
- 101 G.H. Orcutt, M. Greenberger, J. Korbel and A. Rivlin, *Microanalysis of Socioeconomic System: A Simulation Study*, Harper and Row, New York, 1961.
- 102 J.L. Peschar, *Van achteren naar voren, achterstandssituaties in het onderwijs*, Bijlagen tot de Onderwijsdagen 1979, Stichting voor Onderzoek van het Onderwijs, Staatsuitgeverij, The Hague, 1979.
- 103 M.J.H. Ploegmakers and L.Th. van Leeuwen, *Veranderende huishoudens in Nederland, Statistische analyse over de periode 1960-1977*, Ministry of Housing, The Hague, 1983.
- 104 Social and Cultural Planning Agency, *Berekend Beleid*, Staatsuitgeverij, The Hague, 1985.
- 105 H. van de Stadt, *The Dynamics of Income and Welfare*, PhD dissertation, University of Amsterdam, Amsterdam, 1988.
- 106 H. van de Stadt, R.D. Huigen and C. Zeelenberg, 'Inkomenseffecten van de Oort-voorstellen', *Economisch Statistische Berichten*, Vol 74, 1989, pp 524-528 and 537.
- 107 A. Steger, *Haushalte und Familien bis zum Jahr 2000. Eine mikroanalytische Untersuchung für die Bundesrepublik Deutschland*, Campus Verlag, Frankfurt/ New York, 1980.
- 108 P.T.M. Tesser, *Sociale herkomst en schoolloopbanen in het voortgezet onderwijs*, Instituut voor Toegepaste Sociologie, Nijmegen, 1986.
- 109 J.-B. Waterman, *Sociale herkomst, bereikt opleidingsniveau en demografische transitie in een microsimulatiemodel*, Erasmus University, Rotterdam, mimeo, 1989.