Working Paper 2.1a

Demographic modelling: the state of the art

Elisabeth Morand
Laurent Toulemon
Sophie Pennec
Rodolfo Baggio
Francesco Billari

Ined, Paris

Revision: 1

FP7-244557

SustainCity

EUROPEAN COMMISSION
European Research Area
SEVENTH FRAMEWORK PROGRAMME
Funded under Socio-economic Sciences & Humanities

Contents

Demographic modelling: the state of the art

Elisabeth Morand
Laurent Toulemon
Sophie Pennec

INED
133 boulevard Davout
75020, Paris

Teleph.: +33 (0) 1 56 06 20 00
e-Mail-Address elisabeth.morand@ined.fr

Rodolfo Baggio
Francesco Billari

Bocconi University
Milan, Italy

e-Mail-Address : rodolfo.baggio@unibocconi.it

Abstract

Demographic microsimulation models differ from traditional macrosimulation on two ways. First, they operate at individual level rather than use aggregate data; second, they produce results based on repeat random experiments rather than producing average indicators. During the 1990s, many microsimulation models have been developed in OECD countries in order to study population changes and the demographic and economic consequences of specific policy measures, not only at the aggregate levels but also at the individual level and for subgroups of the population. The main models which simulate individual behaviour are Standard microsimulation models (MSMs) and Agent based models (ABMs). The logic of both types of dynamic individual simulation models is presented and some micromodels used in social sciences are described. Based on a literature review and this description of existing population microsimulation models, concrete proposals are made for the SustainCity demographic module.

Keywords

demographic; microsimulation; agent based

Preferred citation style

Morand, E., Toulemon, L., Pennec S., Baggio R., and Billari F. (2010) Demographic modelling: the state of the art , *SustainCity Working Paper*, **2.1a**, Ined, Paris.

1 Introduction

This paper presents a literature review of the state of the art in demographic modelling. The review was done in the context of the project: *SustainCity: Microsimulation, land use and transportation models for more sustainable cities in Europe*[1].

This project considers the development of a land use modelling platform for policy evaluation in European cities. This platform will be based in the software UrbanSim (Waddell, 2002), an urban microsimulator that models the behaviour of the agents in the land use system with an agent-based approach. UrbanSim will be updated and adapted to the context of European cities, resulting in an Europan urban simulator called UrbanSim-E.

Population growth (and the consequent formation of new households through matching of individuals) is the main determinant of demand in the residential real estate market. The current version of UrbanSim does not include any internal demographic model but is flexible to be adapted to work with any demographic model that delivers consistent and complete demographic inputs.

This present paper analyses the state of the art in demographic modelling, including microsimulation, dynamics models and agent-based models. The paper finalizes with a proposal of the modelling approaches and conditions that should be considered for the development of a customized demographic model for UrbanSim-E.

[1] www.sustaincity.eu

2 Microsimulation

In the 1950s, to avoid the errors derived from simulation of aggregate to capture the real situation, the Ortcutt (Orcutt G., 1957) proposes to create a model ‘of various sorts of interacting units which receive inputs and generate outputs’, inputs being ‘anything which enters into, acts upon, or is taken into account of, by the unit’ and outputs ‘anything which stems from, or is generated by, the unit’. Its principal aim was to better simulate variants of policies and in particular distribution of individuals. With this kind of model it is possible to check whether a policy affect the individuals targeted by the rule and look whether there are no counter productive effect of the policies (i.e. a benefit that makes the individual paying more taxes than without the benefit).

It is possible to identify two main streams of research that deals with Ortcutt idea. The well established and worldwide used *microsimulation models* (MSMs) and in particular demographic dynamic microsimulation models and kinship models, while a more recent attempt applies to demographical issues models born in the realm of mathematics and computer science, *Agent based models* (ABMs) *or multi agent systems* (MAS).

The subsequent implementation of this idea, first carried out by the same author with the creation of DYNASIM (Orcutt 1961) and DYNASIM2 (Orcutt et al. 1976) and later followed by many researchers and institution (see Sec. 3 for some examples) defined a set of features which today keep characterizing MSMs. First of all, these models strongly rely on data, very often on a large amount of them; their main aims have often been predictions and analysis of the impact of policies. As in multi-state models, MSMs usually allow individuals to be in a predefined number of *states* which act as ‘containers’ for the single units: a switch between states occurs with a transition probability, which usually depends on a set of individual characteristics. Based on these individual characteristics, MSMs thus allow taking into account the heterogeneity of the population but, contrarily to multi-state models, they also allow for individual heterogeneity, each actual transition (events experienced by the individuals) being chosen by chance, based on the transition probabilities. More important, the events experienced by individual are not restricted to the transitions from one ‘container’ to the other; several events can be deduced from the whole population situation and behaviour. For instance, a simple MSM including fertility, union formation and mortality will allow describing events such as becoming a grandparent, having stepchildren, etc., without any a priori estimate of the these events probabilities. Very often changes /transition probabilities are derived from empirical evidence through statistical analysis (e.g. logistic regressions) but some microsimulation models considers markets based approach to determine some events (housing). Mate matching is in closed model a market based approach. As far as demographic modelling is concerned, MSMs have been extensively used to model nation-specific demographic evolution which acts as a starting point for more complex modules (especially tax-benefit, pension, health and transport policies). It has also been used to study kinship (Wachter, 1997; Le Bras, 1973; Pennec, 1997; Tomassini and Wolf, 2000; Ruggles, 1987). For a review of MSM benefit over aggregate simulations see Van Imhoff and Post (1998).

Some microsimulation models are static, which means that they allow estimating the elasticity of some behaviour as a reaction to any specific change in the parameters in the model. These are the "morning after" type of model (Cassels et al, 2006). They are much easier to built and very useful for very short term simulation. To simulate a complete lifepath or on a longer term, dynamic microsimulation models are needed. Their aim is to move micro-units through time by updating the situation and behaviour of all micro units for each time interval. Dynamic MSMs also allow the individual events to depend on the characteristics of related individuals or the whole population, thus describing interactions between individuals, which is the basic aim of ABMs.

Firstly we will describe some features of both microsimulation models (MSMs and ABMs) and present a typology of such models. Some definition about microsimulation model will be given. Secondly nine Empirically-based microsimulation models, and more specifically their demographic modules, will be presented. Thirdly, we will present Agent-Based Models (ABMs) and their specific advantages and shortcomings compared to MSMs, before making a concrete proposal for SustainCity.

3 Some important features of Microsimulation models

3.1 Continuous versus discrete time

Similar to the difference between statistical models of regression and duration models, the time framework distinguished two types of micro simulation models.

For continuous time models, a random process generates the durations up to all possible events considered for each member of the population. The first event occurring, that is to say the event occurring closer to the beginning point, is executed. The time of occurrence of this first event is the new starting point and the whole procedure is repeated until the event of exiting the population (migration or death) occurs.

For discrete time models, time periods are treated one after the other. For each time period, for all individuals, all possible transitions are determined (without taking into account the exact time of the event within the time period). Each event is considered to occur only once during the period. When different events may occur during the same period of time, the simultaneous events are considered explicitly by the order chosen for the different events.

Continuous time models have certain theoretical advantages (Willekens, 2006), but their implementation is more complex and model operations are less transparent for the user than for discrete time model (Scott et al., 2003, p.5).

If the time period unit is very small, discrete-time models can be consider as “pseudo continuous” models, and in practice they lead to very similar results.

3.2 Open vs. closed models

Models can be open or closed. A close model is a model that considers as a starting population, a sample representative of the population and usually the characteristics of units of the population include links between the individuals when appropriate (most often family ties).

In open models, individual ageing (ageing meaning here the events that an individual can face during each time period) are simulated independently from one another: the individuals enter a union, have children, etc. without any explicit relationship with other individuals. For instance, having a partner is an individual feature of this population member; the partner and children are not explicit individuals, they are created for the sake of simulating our individual lifepath and are not accounted as such in the simulated population.

In Closed model, individuals can enter the population by “birth” or immigration and exit by “death” or emigration. These individuals that enter are new members of the population under study and “full” members in the sense that they will be simulated like the other ones for the future time period. Moreover, as closed models keep track of the ties between the individuals and in particular family ties, some event occurring to one individual may affect some other individuals of the population. For example, when a married person dies, another person he/her spouse become a widow or a widower and the children become orphan by this parent. In practice each individual is related to his/her father and mother, successive partners, children and so on. This allows describing all family bonds at each period of time, such as parity (number of children ever born), or the fact whether father and mother are still alive, etc.

An example of the difference between open and close models is the simulation of partner matching in case of union formation.

In an open model, when the individual enter in union, a partner is created with suitable characteristics. This partner is not a full member of the population, he/she only exists by his/her relationship to a full member. In a closed model, a marriage market is created to match those entitled to partner. This procedure matches the partners according to some characteristics (i.e. age, education level, socio-economic status…). Maximisation techniques are used and some behavioural rules are applied for those who can’t find a perfect partner, they may relax their wishes in terms of partner’s characteristics or wait another time period in the marriage market.

Closed models can include migrations, assuming that the “rest of the world” is a tank of individuals where immigrants are taken from and where emigrants are sent, or where immigrants are “cloned” from the present population.

3.3 Microsimulation versus agent-based models

ABMs’ origins can be traced back to von Neumann (Von Neumann and Burks 1966). Later, well-known pioneering examples of ABMs are John Conway’s Game of Life (Gardner 1970), a grid of cells which can be dead or alive depending on the local alive cell density (they can die both of overcrowding or of isolation) and the famous segregation model by Schelling (Schelling 1971), which demonstrated that urban segregation can arise even in a rather tolerant society. ABMs basically simulate interactions of autonomous *agents* in order to study the emergence of global features from the bottom up, the idea being that simple behavioural rules are able to produce complex global phenomena. Each single agent has built-in rules on the basis of which (s)he acts and reacts through interactions with other agents and with the simulated environment. ABMs are first and foremost used to develop and explore different models and theories, while MSMs usually intend to empirically evaluate the consequences of given estimated probabilities (Billari and Prskawetz 2003). Moreover, MSMs are often criticized for their weak modelling capabilities and their limits in simulation scope: they mainly model macro-to-micro interactions (i.e. the impact of policies on individuals) while usually overlook the micro-to-macro direction (i.e. the effect of individuals on the policy) (Krupp 1986). ABMs' core is a behavioural modelling and they can easily take into account feedback effects between micro and macro levels. More importantly still, MSMs rely on transition rates which are determined once and for all by the background conditions at the time they are captured, while in ABMs agents dynamically adapt and react to the ever-changing simulation environment. For these reasons, ABM techniques appear to be an ideal complement to MSMs. Indeed, a completely agent-based simulation of large-scale demographic dynamics which doesn't make use of MSM tools (such as transition probabilities) hasn't yet appeared. On the contrary, as we'll see in Sec. 5, some hybrid models are beginning to emerge, while 'hard-core' ABMs still focus on more specific and tractable demographic issues (e.g. marriage market, residential pressure, urban development etc.).

In the introduction of a rather recent collection of ABMs applied to demography (Billari and Prskawetz 2003), the authors list the possible added gains for demographic insights coming from this new way of modelling. They identify seven "pros" of agent-based demographic models (ABDMs):

- they are easy to implement, as already mentioned, feedback mechanisms can be included and they integrate micro-based demographic behavioural theories with macro-level demographic outcomes;
- they make it easy to simulate heterogeneous and not-fully-rational agents;
- ABMs adhere to the KISS principle (i.e. Keep It Simple, Stupid), thus calling for simple formulations of theoretical statements, which still are usually capable of generating complex and novel global outcomes;
- they allow to analyse the out-of-equilibrium dynamics of demographical processes;
- the bottom-up approach incorporated in ABMs lets the researchers identify the (minimal) key hypotheses on individual decisional processes needed in order to obtain realistic global patterns: the aim is to 'grow' social phenomena from simple individual assumptions.
- they allow dealing with situations for which no analytical solutions exist (e.g. complex non-linear models): ABMs come in when mathematical specifications become intractable;

- they make it possible to build artificial societies which can work as computational laboratories and even reproduce past global events starting from individual interactions.

Indeed, the authors go as far as arguing that:

> *'[...] it would be useful – particularly in the study of micro-macro relationships and dynamic feed-backs – to build Agent-Based Computational Demography as a computational approach to demography, with emphasis placed on simulations based on agents, including empirically-based microsimulation as a "special" case. [our emphasis] (Billari and Prskawetz 2003, p. 4)*

Of course, ABDMs also have drawbacks: one of the papers featured in the aforementioned collection lists five of them (Chattoe-Brown 2001):

- Social sciences usually model very specific aspects of social behaviour: the standard statistical approach involves a static conception of cause and effects and only identifies some direct causes of the behaviour of interest, while treating all other variables as exogenous. This approach can't be transposed on ABMs, as a more complete representation of social structures (e.g. social networks) is needed in order to 'grow' phenomena from the bottom up.
- ABMs, in a way, are built upon a completely new perspective, which allows agents not to be aggregated (vs. common statistical models) and to be heterogeneous (vs. common mathematical theories). As a consequence, many existing social theories are to be discarded in the implementation of ABMs, if too restrictive or unrealistic: new 'agent-based' theories are needed in order to exploit the fullest potential of ABMs.
- Spatial and temporal structures are fundamental ingredients to simulate a realistic society: certain activities and interactions can only be carried out at appropriate times and places. So ABMs should become able of reproducing spatial and temporal constraints on social action. In most models, even if social structure is represented (and this is not always the case), these kinds of constraints are not implemented.
- Even if the assumption of full rationality is dropped, convergence to equilibrium can derive from the weaker assumption of a 'common model of the world'. It would be better to find ways of simulating in a more realistic way the cognitive complexity of agents, and this may prove a hard task (the agents' memory and how information is stored and turned into actions should be analysed).
- The relation between ABMs and data is rather problematic. Not much thought has been given to what data are needed in order to calibrate ABMs and how to collect them. Apparently, new kinds of data and collection techniques which inform us about the decision processes going on in the agents' minds have to be developed (interviews being one possible solution).

Overall, it appears that MSMs and ABMs share many theoretical foundations, while they differ in their practical applications. For sure, their boundaries, already blurred, are likely to fade away in the near future.

4 A overview of some dynamic microsimulation models

Nine dynamic microsimulation models are considered here. DYNASIM3 (US), DYNACAN (CANADA), MOSART (NORWAY), SESIM (SWEDEN), SAGE (UK), DYNAMOD (AUSTRALIA), APPSIM (AUSTRALIA), DESTINIE (FRANCE), SVERIGE (SWEDEN).

For each model, we focus on the following aspects:

- the initial population or sample
- the list of events and the modelling sequence
- the data used for estimations

All these models are built using a similar structure: an initial base population, a simulation cycle and population outputs. Within the simulation cycle a set of functions calculates the probability of events occurring. But the main aim of these models is not creating a demographic model and the aim can be different from a project to another.

DYNASIM3 (Favreault and Smith, 2004) is a dynamic microsimulation model designed to analyse retirement and ageing issues.

DYNACAN (Morrison, 2000) is designed to study the impact of changes in the formal process for assessing the Canada Pension Plan (CPP).

MOSART (Fredrikse and Stølen, 2007) is designed for the projection of population, human capital, labour Supply and public pension benefits. The model is used both by the Norwegian Ministry of Finance and the Ministry of Labour and Social Inclusion, as well as for research projects in Statistics Norway.

SESIM (Sundberg, 2007) model is developed at the Swedish Ministry of Finance to study the pension system and impact of ageing on public finances

SAGE (Evandrou et al., 2007) is the result of the work of ESRC Research Group 'Simulating Social Policy for an Ageing Society'. The aim is to study the impact of population ageing for pensions and issues regarding health and long term needs.

DYNAMOD (Kelly, 2007) is a dynamic microsimulation model of the Australian population which is designed to project characteristics of the population over a period of up to 50 years.

APPSIM (Pennec and Bacon, 2007) is a dynamic microsimulation model of the Australian population to assess the future distributional and revenue consequences of changes in tax and outlay programs.

DESTINIE (Blanchet and Le Minez, 2007) is designed for studying French pension scheme. The model computes social security contributions, benefits and taxes and simulates the socio-economic evolution of a micro population.

SVERIGE (Holm et al., 2002), "System for Visualizing Economic and Regional Influences Governing the Environment", is a spatial microsimulation model used to based on a database covering the whole Swedish population.

4.1 Starting population

Microsimulation models have all a starting point. The starting population can be a cross-sectional population or a birth cohort. In the first case, a sample of the whole population at time t_0 is taken as a starting point and all individuals are simulated from this time to the ending date (or their death, whatever first). In the second case, a sample of individuals belonging to the same birth cohorts are simulated from their birth to their death. The starting population can be based on survey data or based on a synthetic population i.e. created by imputing the needed variables gathered from different information sources.

Even though microsimulation models are based on Markov chains and the transitions are based on current situation, this potential drawback has been ruled out by using biographical information and time based information. This is the way these models take into account the individuals' past behaviour and some of the covariates of interest (e.g. in order to model the occurrence of a birth we need to know if the women have or not another child). Furthermore, if the survey investigate both at family and individuals levels then it is rather easy to determine the family ties between individual within the population. It is not necessary to know the complete individuals' histories before the starting time, but their rather detailed situation at the starting time must be known (or imputed if necessary). For instance, to simulate fertility varying by number of children ever born and couple situation, the information on women's parity and couple situation must be known and those variables must be on the starting file. If fertility is also varying with union duration, then the date of union (or at least union duration at t_0 must also be known or imputed. The parameters chosen for the transitions must take into account the variables available in the starting population and there is a trade off between imputation of variables in the population that weaken the population representativeness and the need for covariates for the events transitions.

Some authors consider that synthetic starting population (Wachter 1997) has its advantages. Firstly, it allows creating a population as large as needed and, secondly, there are no confidentiality conflicts or privacy problems, because everything is imputed and the information does not come from a survey sample when some individuals could eventually be identified.

As it has been underlined by Zaidi and Rake (2001) in their Lesson 8 “The choice of base data is an early, important decision in the model building process”. The representativeness of base data is a principal concern here, with the use matching techniques to be one way of compensating for the shortage of variables in such datasets”.

All the microsimulation models presented here start with a cross-sectional population. The starting populations consist in large samples of population coming from a survey or from census data. The initial "population" size is varying between 54,000 and 480,000 individuals (see Table 1) except for SVERIGE which cover all the Sweden population.

Table 1: Nature of initial database for seven dynamic microsimulation models

Model (Country)	Initial database	Sample size
DYNASIM3 (USA, 2004)	1990-1993 Survey of Income and Program Participation (SIPP) panels / LSIA a verifier	100,000 people
DYNACAN (Canada, 1994)	Heavily modified version of 1971 Census microdata file	200,000 individuals
MOSART (Norway, 1988)	12 per cent random sample of population administrative data from 1993	About 480,000 individuals
SESIM (Sweden, 1997)	LINDA longitudinal database, created from administrative data, typical sample about 1% of population	about 110,000 individuals
SAGE (UK, 1999)	0.1% of 1991 Census sample	about 54,000 individuals
DYNAMOD (Australia, 1992)	1/1000 sample from the 1986 Census	150,000 individuals
APPSIM (Australia)	1% household sample from the 2001 Census	75400 dwellings and 188,000 individuals
DESTINIE 2 (France,1999)	1999 Family History Survey and 2003 Assets Survey collected by INSEE.	65 000 individuals 20 000 households
SVERIGE (Sweden,199)	Administrative data from 1985 to 1995	9 million individuals

Source: Cassells and al., 2006

4.2 Events simulated and simulation order

In all the models under-study, demographic events can be described in 3 groups:

- population growth: migration, birth, death
- household formation: children leaving home, cohabitation and marriage, divorce and separation.
- Education (enrolment and level of education) ,health status, working status

Table 2: Events simulated and simulation order

(2-Emigration means that the event emigration is simulated subsequently)

Model	Family constitution	Population growth	education and disability
DYNASIM3	First marriage Remarriage Mate matching Divorce Leaving home Living arrangements	Births Deaths Immigration	Education Disability
DYNACAN	6-First marriage 6-Remarriage 6-Mate matching 8-Leaving family of origin 7-Divorce	1-Emigration 2-Interprovincial migration 3-Immigration 4-Deaths 5-Births	9-Education 10-disability
MOSART	4- Moves out of old age care institutions 4- Moves into old age care institutions 4- Children leaving home 4- Adults living in other households without family relations 4- Cohabitation and marriage, including matching of spouses 4- Divorce	1. Immigration 2. Emigration 3. Deaths 5. Births	6. Educational activities and accomplishments
SESIM	1-Mortality 2- Adoption 5-Children leaving home 6-Cohabitation 7-Separation	3-Migration (from/to Sweden) 4-Fertility 10-Regional mobility	8− Disability 9− Rehabilitation 11- Education: -Dropout from upper secondary education - From upper secondary to university - Dropout from university − From labour market to university − From labour market to adult education − From adult education to university
SAGE	Separation Cohabitation Marriage	Mortality Fertility	

Table 2 (Continued): Events simulated and simulation order

DYNAMOD	Couple formation Couple dissolution Young people living home	Births Deaths Overseas immigration Overseas emigration	Disability Education
APPSIM	5-Couple formation 6-Couple dissolution 7-Young people living home	1-half immigration and half emigration 2-Deaths 3-Births 4- half immigration and half emigration	9-Disability 8-Education
DESTINIE	2- Reservation of candidates for couple formation 4- Union dissolution 5- Young people living home 7-union (among reserve create in 2)	1-Deaths 3-Births 6-Migration	**Age at end of studies**
SVERIGE	5-mariage or union 4-living home 5-divorce	1- Deaths 2-emigration 3-Fertility 6-migration (inter or intra regional) 7-immigration	4-Education

All authors have underlined the necessity to subdivide a model into smaller components; "the subdivision of the model into smaller modules is important. This aspect of dynamic microsimulation […] helps to make the working of the model more systematic and makes it easier to check problems at different stages in the running of the model" (Zaidi and Rake, 2001, p. 8).

Even if the events studied are similar, differences between models exist firstly in the sequence chosen of the events within the time period simulation loop and secondly in the information used to simulate each event. Ageing (i.e. adding one year to the age of people who do not die during the considered time period) is in most cases performed at the end of a time period, before a new cycle begins or at the very beginning of the new time period. But in some models ageing is performed during the cycle. In Dynacan, for example, ageing is performed before marriage module and after fertility simulation.

All these models are closed, with migrations. To create the immigrants, most of the models use a duplicate part of last year immigrants. But other solutions can be found depending on available information. For example, in MOSART, an exogenous number of "net immigrants" by gender and age are added to the population each year. Second, net migrants are assigned characteristics like the average Norwegian population.

Furthermore to estimate and simulate each event for each people a lot of information has to be found. For example, in the case of the divorce in Dynamsim3 and Dynacan model, you need about 7 variables. More variables to simulate more precisely events need more microdata set.

Table 3: examples of covariates used in two models

Model	Characteristics used for estimation
DYNASIM3	**Death:** age, race, marital status, education, parity, sex **Birth**: age, race, marital status, education parity **Marriage**: age, race, sex, previous marital status, income, education, region, weeks worked, hourly wage, asset income, receipt of welfare, unemployment compensation **Mate matching:** Difference in age, difference in education **Divorce**: Distribution over time of expected divorces for this marriage cohort, age at marriage, education, previous marital status, presence of young children, weeks worked, wages **Leaving home** (beside marriage, giving birth, divorce and death): Age, race, sex **Education:** Race, sex, age, years at current school level, parents' education

Table 3 (Continued): examples of covariates used in two models

DYNACAN	**Death:** age, sex, marital status, labour force participation, employment, earnings, family income, education, disability status, region (Quebec, RoC), place of birth (Canada, other), year (DYNACAN will not permit both spouses in a family to die in same year.) Equations originally are from CORSIM, adapted to Canadian data. **Birth:** Age, marital status, duration of marriage, birth(t-1), birth(t-2), number of children, education, school attendance, labour force participation, employment, living arrangements, family income, earnings, home ownership, work status (full/part-time), region (Quebec, RoC) Equations from CORSIM **First mar**riage: Age, sex, education, earnings, presence of children **Remarriage:** Age, sex, previous marital status (divorced/widowed), earnings, presence of children, education **Mate matching**: Age, sex, education, previous marital status **Leaving family of origin**: Age, sex, school attendance (lagged), number of parents, parents' education, employment, earnings (lagged), presence of children, family income, family structure, number of younger siblings (Uses CORSIM data) **Divorce**: Duration of marriage, presence, number, and ages of children, wages, does wife have earnings; earnings level, education attainment **Custody of children at divorce**: Biological ties, if equal ties random of parents assignment of 90 percent to mother **Emigration from Canada** Rates by age, sex, marital status from 1991 Statistics Canada data; aligned by age, sex, source region (Quebec, RoC) Aligned to CPPAVM. **Immigration (cloning of existing family)** Rates by age, sex, marital status from 1991 Statistics Canada data; aligned by age, sex, destination region (Quebec, RoC) Aligned to CPPAVM

4.3 Alignment and calibration

In order to make the results compatible with external constraints (macro-level projections), microsimulation models like most models might need some alterations. This "alignment" is made ex post on the total population (i.e. adding immigrants or emigrants) in order to increase or decrease the estimated population size. A better approach is to align by flow of events (births, deaths, net migrations). Another way to ensure consistency is to use “calibration” with changing parameters value (i.e. changing fertility rates) in order that the outputs remain consistent with the macro-level trend.

Calibration can also be made on the number of households. For instance, if the simulated number of households is too large (and the number of persons per household too low), it is possible to delay the age at leaving the parental home, or to increase union rates, or to decrease divorce rates. Some changes can be ambiguous, such as changing the rates from ordinary households to communal establishments (depending on the household size of people going to or coming from educational establishments or care homes).

4.4 Microdata used

Microsimulation model need a base year dataset but also information to model all events. Most models use longitudinal survey to estimate transitions. In some cases, the information is not available in the country and an estimation based on another country is used. For example, in French model Destinie the USA survey Longitudinal Study on Aging (about 12,000 individuals studied in 1984, 1986, 1988 and 1990) was used to estimate transition to disability. The same happens for Dynacan which used US estimation for items that were not available for Canada.

Table 4: Microdata used in the Microsimulation models

Model	Microdata
DYNASIM3	Current Population Survey (CPS); Health and Retirement Survey (HRS); National Longitudinal Mortality Study (NLMS); National Longitudinal Survey of Youth (NLSY); Vital Statistics
DYNACAN	1971, 1981, 1991 Canadian Census microdata file from the full census tabulations, supplemented by large sample versions of the Canadian Survey of Consumer Finances, Labour Market Activities Survey, CPP administrative data, data from CPP Actuarial Valuation Model, Canadian vital statistics, and other data. In some instances where Canadian microdata are not available, U.S. data have been used, aligned to Canadian totals.
MOSART	1988 Family and Occupation Survey Population administrative data
SESIM	Initial data base (LINDA) Longitudinal Individual Data for Sweden HEK is based on an interview, merged with register information, and therefore have a much better possibility of creating a household according to an economic definition. This is in contrast to LINDA, which essentially is based on a household according to a tax definition. Regional mobility and tenure choice is based on GEOSWEDE5. This database is constructed from Louise and RTB: GEOSWEDE have been used is in models for regional mobility and health.
SAGE	1991 Household SAR ++ Mortality: ONS Longitudinal Study + GAD Fertility, partnership changes: BHPS w1-9 Partner selection: GHS 1992-8Labour
DYNAMOD	Longitudinal Study of Immigrants 1986 national survey Retrospective longitudinal data from the 1986 Australian Family Project Australian Longitudinal Survey
APPSIM	1% household census sample Longitudinal surveys on immigrants to Australia (LSIA) Household, Income, Labour force Dynamics in Australia survey (HILDA) Health survey

Table 4 (Continued): Microdata used in the Microsimulation models

DESTINIE	1999 Family History Survey 2004 French Census 2003 Assets Survey
SVERIGE	ASTRID longitudinal individual database Administrative registers

4.5 Software

Most of models use object-oriented languages to perform the simulation and a statistical module for evaluating parameters used and presenting the results.

Table 5: software used in the Microsimulation models

Model	Software	Extraction and presentation of results
DYNASIM3	the Family and Earnings History (FEH) model is in FORTRAN	FEH model output is a set of longitudinal demographic and labour force histories
DYNACAN	Most of the model's components are built in the **C language**, and run under the **Linux operating system**.	Other components, e.g. programs like SAS and TPL used for post-processing, secondary analysis, and parameter estimation
MOSART	the simulation part of the model is written in **Simula,** an object-oriented language which run from a **Unix platform** Translated from Simula to **C # language** in 2006	An ASCII-file with one record per selected person per selected year with selected variables. Special tables may be produced from this file with a suitable table production programme, usually SAS.
SESIM	Visual Basic 6.0, with Microsoft Excel and Access	'Excel report generator'. More extensive analysis of the results is performed in standard statistical packages like SAS or R.
SAGE	**Borland C++** compiler, command-line version 5, for **Windows 32** operating systems.	Results are stored in tab-delimited files. A variety of other software, particularly Excel and STATA, contribute to the extraction and presentation of the model's results.

Table 5 (Continued): software used in the Microsimulation models

DYNAMOD	Writing in **C++ language**	DYNAMOD-3 was shifted to a personal computer (PC) platform
APPSIM	Writing in C# language	Parameters are stored and aggregate results are in Excel spreadsheets.
DESTINIE	Translated into Pearl for Destinie 2 in 2005	Final outputs: Excel and text files SAS is used for parameters estimation
SVERIGE	simulation engine (Turbo) and SVERIGE model are programmed in C++,	Input and output are stored in Excel spreadsheets.

4.6 A specific example: the fertility module

The following examples present the differences in the hypotheses and the simulation rules used to simulate a birth process.

In the SESIM model marital status is taken into account but a birth is a woman's characteristic. In DESTINIE, birth is a couple's behaviour and not an individual one.

In SESIM, simulation is stochastic while DESTINIE combines behavioural and stochastic rules.

DESTINIE Model

It assumes that children can be born only within a couple where the woman is aged between 15 and 49 and that a woman can have at most 6 children.

The probability that a woman gives birth depends on several covariates such as age of the mother, number of children ever born, and union duration for having a first birth and time elapsed since previous birth for children of parity 2 and above.

To simulate first birth, a proportion of women to become a mother, 90% as estimated from the 1999 Family History Survey, is used to generate a distribution over the 10 first years of women's first union, the mode being set at 3 years. The probability of a first birth depends explicitly on union duration and age at union, graduate women enter in their first union later, and have their first child at an older age than non graduate women. The model therefore generates the postponement at childbearing observed in recent years with the increasing proportion of graduates among the successive cohorts.

For birth of children of higher parity the transition takes into account whether the woman is still in the same union or if she is in a new union.

- If the woman did not separate from her partner since the birth of her last child, the probability of birth of a new child is estimated by dividing the parity of the offspring derived from the Family Survey of 10 years from the birth of previous child with a mode equal to 4 years.
- If the woman has separated from her husband since the birth of her last child and subsequently entered a new union, the fertility rates are higher, for comparable age of the latest child. This approach help to avoid having a mechanical fertility decline due to union dissolutions.

SESIM Model

Two models are evaluated, one for the women who remain childless, another for women who have at least one child. Children can be born from women aged 18 to 49, who have moved out from the parental home.

To estimate whether a woman has or not any children, a logistic regression is run with the following covariates: age, marital status, pensionable income (quartiles), working status, highest education.

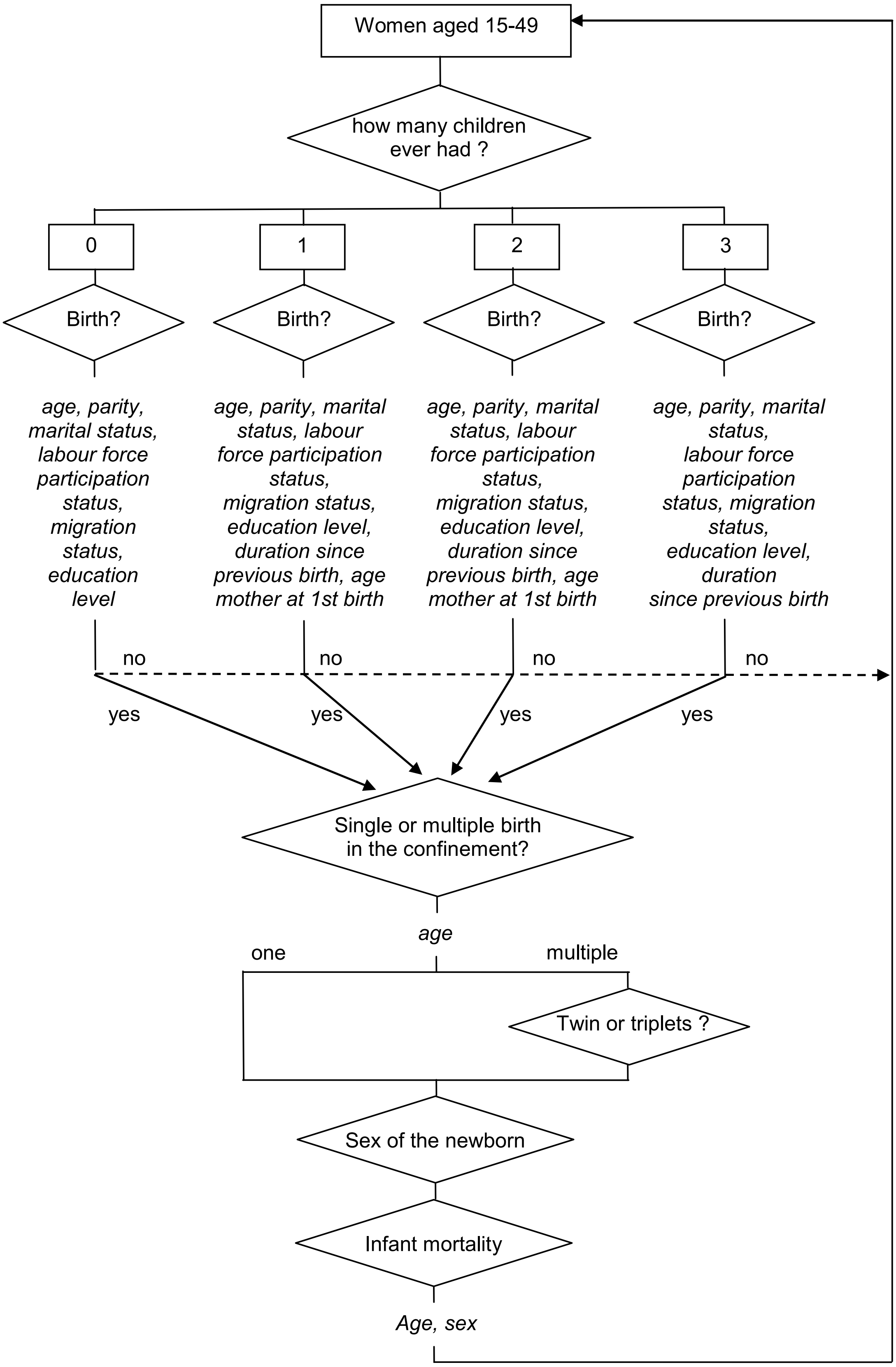
Among mothers, another logistic regression is used with the following covariates: age, marital status, pensionable income (quartiles), working status, highest education, age of youngest child. Separate models have been estimated for women who have given birth to one, two and three children respectively, births of order five and more are not modelled.

Each model gives for each woman a probability to give birth to a child each year. The event “giving birth to a child” is then simulated by comparing the probability estimated from the regression model with a uniform random number. If the random number is lower than the probability, the event takes place and the woman gives birth to a child.

APPSIM Model

Each woman aged 15 to 49 enters into the fertility loop. Her probability of having a child each year is related to a set of covariates — the number of children she has ever had (parity), her age, education level, labour force participation status, migration status, and for parity 2 and 3 her age at the first birth (figure 1).

Figure 1. Childbearing in APPSIM



Source: From Pennec and Bacon, 2007.

Once a woman is set to have a child within the year, some characteristics of the outcome of the confinement have to determined, such as whether the woman gives birth to one or more children. Next, some characteristics of the newborn(s) have to be determined; their birth cohort is deduced from the year the birth arises, but the sex of the newborn and whether he/she will survive from birth to the beginning of the next period of simulation (Pennec and Bacon, 2007).

5 Agent-based models

As already stated, full-fledged ABDMs which take into account all main demographic processes have not yet been developed. This is probably due to the insufficient expertise gained in the scientific community: ABDMs are still in their infancy. Yet, they have indeed been useful to gain insight into some more specific demographic issues. Moreover, they are increasingly used in combination with standard MSMs in order to enhance the models' realism and add a behavioural dimension to the analyses. In this section we present some instances of pure ABDMs along with some innovative hybrid simulations.

5.1 Spatial demography

The Schelling model has inspired many simulations of spatial demography. For example, Benenson and co-authors have used ABDM to 'grow' residential migration in the Yaffo area of Tel Aviv (Benenson et al. 2003). Individuals can belong to three different cultural groups (Arab Christians, Arab Muslims and Jews) and recognize the cultural type which prevails in their neighbourhood along with the architectural style of their own building. These two variables can make the agents move to another area. The timing is as follows: agents decide whether to move; they randomly select houses with vacant dwelling; they move either to a randomly chosen one or to the most attractive, according to their preferences. The results of the simulations fit very well the real dynamics of the Yaffo area, especially when agents are allowed to rank vacant dwellings by attractiveness.

Another interesting application of ABMs to spatial demography has been devised in order to simulate state-level migration patterns from East to West Germany following the collapse of the Berlin wall (Heiland 2003). The decision to migrate is assumed to be triggered by unemployment, income differentials, moving costs and job search costs. Agents are heterogeneous with respect to their location (i.e. which of the five eastern German states they live in) and mobility (i.e. individual moving costs vary). Heterogeneity appears to be crucial in producing a realistic pattern of east-to-west migration from 1989 on. Stylized facts of migration dynamics are well captured by the model: e.g. unemployed agents are more likely to migrate and geographical proximity along with the demand of technically skilled labour explains most of the observed distribution of migrants across western German states.

ABMs' topologies (i.e. how space is represented) span from simple 2-dimensional lattices to Geographic Information Systems (GIS). The latter space representation, in its essence, results from the merging of cartography and database technology. Recent contributions have tried to build methodologies to integrate ABMs with GIS (Castle and Crooks 2006; Crooks 2006; Crooks et al. 2008). In particular, GIS exhibit some problems in representing time and change and a purposely built modelling system is called for in order to best exploit the potential of these two tools. Some applications of this integration are also developed to show its strong and weak points: for example, a pedestrian model for emergency evacuation in central London; a Schelling-like model of residential segregation tuned to London geography; an AB residential location model for Greater London (Crooks et al. 2008).

Lately, hybrid MSM + ABM simulations are starting to emerge. Wu et al.(2008) is an excellent example. This paper is part of the MoSeS project (Modelling and Simulation of e-Social Science), which aims at producing a MSM for the entire UK population, while improving the behavioural dimension of the simulation by use of ABM. Specifically, this contribution shows the added gains of combining these two tools in an application to the study of migration in the Leeds area: pure spatial MSM is not able to capture the subtleties of local migration patterns, as it cannot differentiate higher education students from other migrants. Rather intuitively, students exhibit very specific migration patterns: they tend to cluster and live close to the university, they move away from Leeds once graduated and new 'immigrant' students replace them, thus keeping the university neighbourhoods 'younger' than the others. MSM cannot capture this different behaviour as it cannot easily handle heterogeneous agents: for example, it fails to reproduce the student population renewal, as it treats students as other migrants, thus making them settle down in Leeds (the peaks of young people in the university neighbourhood disappears after 10-30 years of simulation). To solve this problem, while keeping the standard transition-probability-driven MSM structure, which models six demographic processes (ageing, mortality, fertility, health change, marriage and migration), the authors implement an ABM specifically devised for students' migration. Higher education students are divided into four subgroups on the basis of age, as they have distinctive migration behaviours (e.g. freshmen tend to live on-campus etc.). Each group is allowed to live in a particular area for a set number of years and students stay as close as possible to their universities, while at the end of their studies they go back to the general migration process. Not surprisingly, the hybrid model results reflect reality much better than the pure MSM results.

5.2 Family demography

Decision-making processes (e.g. partnership formation) constitute one of the main objects of study in family demography: this makes this field an ideal playground for ABMs. Todd (1997) is probably one of the first to explore this subject. He makes use of a very simple ABM in order to shed light on best-mate-choice algorithms. He finds that a choice faster than the optimal algorithm gives very good results in finding the best mate. Surprisingly, on some dimensions it gives even better results. Additionally, it would be impossible to derive the results in an analytical way. This is a perfect example of the KISS principle applied to social simulations.

Todd and Billari (2003) build on this issue by combining the micro-level mate-choice algorithm with the macro-level emergence of global patterns of marriage. The authors simulate the behaviour of a cohort of agents looking for mates through both one-sided and mutual choice. They find that individual aspiration-based heuristics, which assume bounded rationality and are psychologically plausible (vs. full rationality mathematically-derived optima) need an additional ingredient in order to produce realistic global patterns: heterogeneity in strategies. These are easily implementable in the ABM, heterogeneous agents being probably the strongest added gain of this modelling technique.

Billari et al. (2003) use an ABM to study the cultural evolution of age-at-marriage norms. These norms are modelled as built-in agent constraints and they can be transmitted to the next generation by those individuals who can marry and thus have children. Agents can marry each other only if they share part of the socially acceptable age interval. The authors show that under particular assumptions about the intergenerational transmission of norms, these can persist in the long run. The model is able to take into account norm heterogeneity (i.e. subgroups following different norms), diversity of rationality (i.e. subgroups not following the norms) and also social influence, implemented as a conformity bias, which significantly modifies the speed of evolution of the process. Finally, the model exhibits strong path dependence, the initial situation having a long-term impact on the results.

An interesting agent-based marriage model grounded on social interaction has recently been proposed (Billari et al. 2007). This simulation considers both key factors which intervene between social structure and marriage patterns, i.e. availability of mates and desirability of marriage. The former variable is derived from the set of potential partners, modelled through age-dependent network size and two-sided (mutual) mate search, while the latter explicitly considers the dynamics of social pressure, proxied by the share of married peers, which directly influences the propensity to marry. The results show that the marriage hazard function emerges from the micro-level dynamics and even more importantly its qualitative form appears to be rather robust towards changes in parameters. The possible generalizations of this model include courtship and divorce.

Hills and Todd (2008) have developed the MADAM model (Marriage and Divorce Annealing Model), which aims at simulating assortative, non-competitive, mate selection processes along with divorce dynamics. Basically, agents search for mate similar to themselves, but relax expectations as they age. ‘Similarity’ is modelled through a number N of traits, from which each individual chooses a set of k traits to represent her/his identity. Divorce is possible whenever an agent finds a partner ‘more similar’ to her/himself than their current partner. The empirically observed increase in mean age at first marriage can then derive both from increasing population heterogeneity and from a reduction in the rate of relaxing expectations. Very interestingly, MADAM is also able to accurately predict marriage and divorce patterns across different cultures.

5.3 Historical demography

ABMs have been used to explore the dynamics emerging from small-scale human societies. These artificial societies are computational social laboratories where the main focus of analysis is on which changes in micro-level behaviours and in the environment are able to produce global-level regularities. The most famous contribution to this field is the Sugarscape model by Epstein and Axtell (1996). In this simulation, the authors set up an environment where agents are located on a landscape of a generalized resource (sugar). They have many individual attributes (e.g. vision, metabolism, speed etc.) and follow a simple survival rule: at each step they move to the visible spot with more sugar, and eat it. Sugar is required for surviving and is burned up in movements at an amount equal to the agents’ metabolic rates. By adding layers of simple behavioural rules and environment attributes, Epstein and Axtell are able to ‘grow’ complex collective behaviours, such as group formation, cultural transmission, combat, trade and migration. Indeed, they manage to simulate a stylized ‘proto-history’ of civilization, where emerging tribes compete for cultural dominance and produce complex social histories. Their approach has inspired an interesting stream of research which analyses ancient/historical societies, for which no standard approach is possible due to the lack of reliable data.

Two subsequent papers provide an empirical, ‘real world’ evaluation of the Sugarscape model, the Artificial Anasazi Project, along with exploring its potential in ‘explaining’ proto-histories of civilizations (Dean et al. 2000; Axtell et al. 2002). The Kayenta Anasazi agricultural civilization inhabited the Long House Valley in the Black Mesa area of north-eastern Arizona from about 1800 B.C. to about A.D. 1300. Long House Valley appears to be well suited for a Sugarsacape-like ABM as it exhibits a bounded topography along with a rich paleo-environmental record, which permits an accurate replication of the landscape and of the climate conditions from A.D. 400 to 1400. The basic environmental variable is annual potential maize production for each hectare of the study area. The smallest simulation units (i.e. the agents) represent single households, which are defined by certain individual attributes (e.g. age, size, composition, amount of maize storage etc.). Specific anthropologically plausible rules of behaviour define how the households select their planting and dwelling location. The success in meeting nutritional needs determines the migration pattern. The simulation closely reproduces the main features of the Anasazi demographic dynamics, such as ebb and flow, internal migration patterns and eventual decline.

Other contributions built upon the Sugarscape model include the papers by Ewert et al. (2003) and König et al. (2003). The former implements an ABM of a pre-modern town in JAMES (Java-based Agent Modelling Environment for Simulation) in order to analyse mortality crises in pre-modern European towns. Three types of agents populate the landscape (merchants, craftsmen and labourers) and maximize their utility and profit, while a planning agent which represents local authorities intervenes into market and social structures following a particular objective function. Different kinds of disaster are then introduced: realistic consequences on both demographic and economic variables emerge. In the paper by König et al., the Sugarscape model is augmented in some dimensions in order to study the effect of agent coordination and hierarchies on sustainability. Individuals are allowed, e.g., to subordinate and unsubordinate, start and end collaborating and rest. Coordinators act as information 'hubs': subordinates share their information about food location with them, and coordinators suggest the best location to find food in return for a contribution. The model shows that sustainability is more easily achieved when agents are allowed to coordinate.

6 Summary and proposals for SustainCity

This contribution has reviewed the main issues regarding the simulation of demographic characteristics of a population. The two main methods for simulations have been described and discussed, namely microsimulation and agent-based models.

Microsimulation models (MSMs) traditionally strongly rely on empirical data and, usually, they aim at predicting and analysing the impact of policies. MSMs simulate the events that individuals may face over time. These events can be deterministic based (age is an example), stochastic transition based or market based. These transitions usually depend on a set of individual characteristics but for market transitions, behavioural rules can be set also along with other individual characteristics, like in ABMs; these behavioural rules are a way to take into account the environmental and market situation.

Agent-based models (ABMs) simulate the interactions between autonomous agents in order to study the emergence of global features. The idea is that simple behavioural rules can produce complex global phenomena. Each agent has built-in characteristics and acts according to simple rules that govern its behaviour and the interactions with other agents and with the environment. ABMs are first and foremost used to develop and explore different models and theories, while MSMs usually intend to empirically evaluate the consequences of given estimated actions to be performed on the system under study or to perform elaborated population projections.

Agent based models had the ambition to cover properties of artificial agent such as ”achieve strategic goals by changing immediate and close constraints”. Agent based models would exploit more complexity than standard microsimulation models. Conversely, microsimulation is supported by substantial data collection, representation, estimation, and validation in an empirical setting.

For these reasons, ABM techniques appear to be an ideal complement to MSMs. Agent based models are particularly efficient when an individual’s behaviour mainly depends on the other individuals’ situation or behaviour. This is the case e.g. for the marriage market. For many demographic events anyway the actual transition probabilities depend on individual characteristics, such as age, marital situation, level of education, etc. Thus, hybrid models are beginning to emerge and look promising.

For what concerns SustainCity, the idea is to follow the suggestion by Ann Harding (NATSEM's Director during the construction of APPSIM): "Today, I would place a much greater importance on developing the simplest possible (but functioning) version of a model, on getting that well documented and on producing papers containing illustrative results within the project budget and timeframe." (Cassels et al, 2006).

In other words, the minimum requirements for a demographic model within SustainCity can be expressed as such:

- Closed, discrete time model (step = 1 year)
- Women driven (cohabitation, children, separation)
- Including a partner matching module following agent based logic
- Including household and individual levels (identification of father, mother, partner and household). A new household identification number is given to each man and each woman when he/she leaves parental home. Women keep this number all through the simulation.

APPSIM or SVERIGE appear to be good candidates for these requisites.

It must be noted here, however, that URBANSIM has no provisions for an internal demographic model. Using a ready external model is a relatively easy way of facing this issue, but data integration aspects must be addressed and evaluated. An exchange routine has to be built in, allowing updating the demographic information from the demographic model to the general model. If no feed back is foreseen from the global model to the demographic dynamics, the demographic features of the population may be considered as exogenous to the model, and the data have to be included only once; on the contrary, if the features of the population given by the global model must be considered as inputs for the demographic model, then there is a feed back from the general model to the demographic model, and the exchange of data must be included in the yearly loop. In this latter case, it would seem then reasonable to think about the possibility to implement at least a basic framework, internal to URBANSIM, for providing these functions. Even if probably more demanding from a development (resources) point of view, this solution would ease very much the integration of this aspect into the final product.

On the current stage of the project, it seems more careful to run separately an existing demographic model, which allows for more flexibility in the construction of the model and its elaboration in terms of needs for the other modules of URBANSIM. The integration could then be made when the demographic model has already been implemented; the feed back loops and the necessary calibrations will then be easier to foreseen. The demographic module will produce yearly estimates of the population in a “census-like” database: one line per individual, with a constant identification number, and a series of variables describing

- his/her characteristics (age, sex, marital status)
- the identification number of his/her household
- the identification numbers of related individuals (current and previous partners, children, parents

These estimates can be produced for the whole period altogether or one year after the other, if a feed-back from URBANSIM is needed. From these data, URBANSIM will locate the individuals in order to integrate them into the other modules.

7 References

Aaberge, R., Colombino, U., Holmøy, E., Strøm, B. and Wennemo,T. (2007), "Population ageing and fiscal sustainability: integrating detailed labour supply models with CGE models", in Harding, A. and Gupta, A. (eds), *Modelling Our Future: Population Ageing, Social Security and Taxation*, North-Holland, Amsterdam.

Anderson, JM. (2001), *Models for Retirement Policy Analysis*, Report to the Society of Actuaries, USA,

Antcliff, S. (1993), *An Introduction to DYNAMOD: A Dynamic Microsimulation Model*, DYNAMOD Technical Paper No.1, NATSEM, University of Canberra, Canberra.

Axtell, R., J. Epstein, Dean J., Gumerman G., Swedlund A., Harburger J., Chakravarty S., Hammond R., Parker J., and M. Parker (2002). "Population growth and collapse in a multiagent model of the Kayenta Anasazi in Long House Valley." *Proceedings of the National Academy of Sciences* 99(Suppl 3): 7275.

Baldini, M. (2001), "Inequality and Redistribution over the Life-Cycle in Italy: An Analysis with a Dynamic Cohort Microsimulation Model*", Brazilian Electronic Journal of Economics*, vol. 4, no. 2.

Benenson, I., Omer I., and E. Hatna. (2003). Agent-Based Modeling of Householders' Migration Behavior and Its Consequences. *Agent-Based Computational Demography*. F. Billari and A. Prskawetz, Physica-Verlag.

Billari, F. and A. Prskawetz, Eds. (2003). *Agent-based computational demography*. Contributions to Economics, Physica-Verlag.

Billari, F., A. Prskawetz, and J. Fürnkranz. (2003). On the cultural evolution of age-at-marriage norms. *Agent-based computational demography*. F. Billari and A. Prskawetz, Physica-Verlag.

Billari, F., A. Prskawetz, et al. (2007). "The "Wedding Ring": An Agent-Based Marriage Model Based on Social Interaction." *Demographic Research,* 17: 59.

Blanchet, D. and Le Minez, S., (2007). *The Destinie II Model: A Progress Report*, 1st General Conference of the International Microsimulation Association, Vienna, Austria

Bonnet, C. and Mahieu, R. (2000), "Public pensions in a dynamic microanalytic framework: the case of France", in Mitton, L., Sutherland, H. and Weeks, M. (eds), *Microsimulation Modelling for Policy Analysis*, Cambridge University Press, Cambridge, pp 175-199.

Booth, H. (2004), "On the importance of being uncertain: forecasting population futures for Australia", *People and Place*, vo1. 12, no. 2, pp 1-12.

Bourguignon, F. and Spadaro, A. (2006), *Microsimulation as a Tool for Evaluating Redistribution Policies*, Working Paper 2006-20, Society for the Study of Economic Inequality, (available from www.ecineq.org)

Buddelmeyer, H. Freebairn, J. and Kalb, G. (2006), Evaluation of policy options to encourage welfare to work, Paper presented at the 8th Nordic Seminar on Microsimulation Models, Oslo, Norway, 8-9June,

Caldwell, S. (1990), *Static, Dynamic and Mixed Microsimulation,* mimeo, Department of Sociology, Cornell University, Ithaca, New York

Caldwell, S. (1996), "Health, welfare, pensions and life paths: the CORSIM dynamic microsimulation model", in Harding, A. (ed), *Microsimulation and Public Policy*, North- Holland, Amsterdam, pp. 505-522

Caldwell, S. (2006), *Dynamic microsimulation: directions for improvement*, Paper presented to the 8th Nordic Seminar on Microsimulation Models, Oslo, Norway, 9 June

Castle, C. and Crooks, A. (2006). "Principles and concepts of agent-based modelling for developing geospatial simulations." *Centre for Advanced Spatial Analysis (University College London): Working Paper # 110.*

Chattoe-Brown, E. (2001). "The Role of Agent-Based Modelling in Demographic Explanation". *Agent-based computational demography*. F. Billari and A. Prskawetz, Physica-Verlag.

Chenard, D. (2000), "Individual alignment and group processing: an application to migration processes in DYNACAN", in Mitton, L., Sutherland, H. and Weeks, M. (eds), *Microsimulation Modelling for Policy Analysis*, Cambridge University Press, Cambridge.

Cassels, R. Harding, A. and Kelly S. (2006). " Problems and Prospect for Dynamic Microsimulation. A Review and Lessons for APPSIM. Online Discussion Paper-DP63, NATSEM, University of Canberra, Canberra

Citro, CF. and Hanushek, EA. (1991), *The Uses of Microsimulation Modelling, Vol. 1: Review and Recommendations*, National Academy Press, Washington. 46 *Discussion Paper no.* 63

Cotis, J. (2003), "Population ageing: facing the challenge', *OECD Observer*, No 239, September, p. 1.

Crooks, A. (2006). "Exploring cities using agent-based models and GIS." *Centre for Advanced Spatial Analysis (University College London): Working Paper # 109*.

Crooks, A., C. Castle, and M. Batty. (2008). "Key challenges in agent-based modelling for geo-spatial simulation." *Computers, Environment and Urban Systems* 32(6): 417-430.

Dean, J., G. Gumerman, Epstein J., Axtell R., Swedlund A., Parker M., and S. McCarroll (2000). "Understanding Anasazi culture change through Agent-Based Modeling." Kohler T. and Gumerman G. (eds), *Dynamics in human and primate societies: Agent-based modeling of social and spatial processes:* 179-205.

Epstein, J. and R. Axtell (1996). *Growing artificial societies. Social science from the bottom up*. Washington, DC, Brooking Institutions Press.

Evandrou, M., Falkingham, J., Johnson, P., Scott, A. and Zaidi, A. (2007), "SAGE dynamic microsimulation model for Britain", in Gupta, A. and Harding, A. (eds), *Modelling our Future: Population Ageing, Health and Aged Care,* North-Holland, Amsterdam.

Ewert, U., Roehl M., and A. Uhrmacher. (2003). "Consequences of mortality crises in pre-modern european towns: a multiagent-based simulation approach.". F. Billari and A. Prskawetz, *Agent-based computational demography* Physica-Verlag.

Falkingham, J. and Harding, A. (1996), "Poverty alleviation vs social insurance systems: a comparison of lifetime redistribution", in Harding, A. (ed), *Microsimulation and Public Policy,* North Holland, Amsterdam, pp. 233-265

Falkingham, J. and Hills, J. (1995), *The Dynamic of Welfare: The Welfare State and the Life Cycle,*Prentice-Hall, New York.

Favreault, M. and Caldwell, S. (2000), "Assessing distributional impacts of social security using microsimulation", in Gupta, A. and Kapur, V. (eds), *Microsimulation in Government Policy and Forecasting*, North-Holland, Amsterdam, pp. 395-426.

Favreault, M. and Sammartino, FJ. (2002), *Impact of Social Security Reform on Low-Income and Older Women*, AARP Public Policy Institute Report No. 2002-11, AARP, Washington, DC.

Favreault, M. and Smith, K. (2004), *A Primer on the Dynamic Simulation of Income Model (DYNASIM3)*, The Retirement Project, The Urban Institute, Washington.

Favreault, M. et al. (2002), "Introduction", in Favreault, M., Frank, M., Sammartino, J. and Eugene Steuerle, C. (eds), *In Social Security and the Family: Addressing Unmet Needs in an Underfunded System*, Urban Institute Press, Washington DC, pp. 1-17.

Favreault, M. et al. (2002), "Social security benefits for spouses and survivors: options for change", in Favreault, M., Frank, M., Sammartino, J. and Eugene Steuerle, C. *In Social Security and the Family: Addressing Unmet Needs in an Underfunded System*, Urban Institute Press, Washington DC, pp. 117-227.

Flood, L., Jansson,F. Pettersson, T, Sundberg, O and Westerberg, A, (2003), *SESIM III- a Swedish dynamic micro simulation model* Handbook of SESIM, http://www.sesim.org/.

Flood,L. (2007), '"Can we afford the future? an evaluation of the new Swedish pension system", in Harding, A. and Gupta, A. (eds), *Modelling our Future: Population Ageing, Social Security and Taxation,* North-Holland, Amsterdam.

Fredriksen, D. and Stølen, N. (2007), "MOSART", in Gupta, A. and Harding, A. (eds), *Modelling our Future: Population, Ageing, Health and Aged Care*, North-Holland, Amsterdam.

Fredriksen,D., Heide, KM., Holmøy, E. and Solli, IF. (2007), "Macroeconomic effects of proposal pension reforms in Norway", in Harding, A. and Gupta, A. (eds), *Modelling our Future: Population Ageing, Social Security and Taxation,* North-Holland, Amsterdam.

Gardner, M. (1970). "Mathematical games: The fantastic combinations of John Conway's new solitaire game 'Life'." *Scientific American* 223(4): 120-123.

Gribble, S., Hicks, C. and Rowe, G. (2003), The LifePaths microsimulation model, Paper presented at the International Microsimulation Conference on Population Ageing and Health: Modelling Our Future, Canberra, 7-12 December, (available from www.natsem.canberra.edu.au).*Problems and Prospects for Dynamic Microsimulation: A Review and Lessons for APPSIM* 47

Gruskin, A. (1994), "Dynamic Microsimulation: A programming approach called DYNAMOD", in Antcliff et al, *Development of DYNAMOD: 1993 and 1994*, Dynamic Modelling Working Paper no. 1, National Centre for Social and Economic Modelling, University of Canberra, pp 1-36.

Hain, W. and Helberger, C. (1986), "Longitudinal simulation of lifetime income", in Orcutt, G.H., Merz, J. and Quinke, H. (eds), *Microanalytic Simulation Models to Support Social and Financial Policy*, North-Holland, Amsterdam, pp. 251-270.

Hancock, R. (2000), "Charging for care in later life: an exercise in dynamic microsimulation", in Mitton, L., Sutherland, H. and Weeks, M. (eds). (2000), *Microsimulation Modelling for Policy Analysis*, Cambridge University Press, Cambridge, pp. 226-237.

Harding, A. (1993a), *Lifetime Income Distribution and Redistribution: Applications of aMicrosimulation Model,* North-Holland, Amsterdam.

Harding, A. (1993b), "Lifetime vs annual tax-transfer incidence: how much less progressive?" *Economic Record*, vol. 69, no. 205, pp. 179-191.

Harding, A. (1995), "Financing higher education: an assessment of income-contingent loan options and repayment patterns over the life cycle", *Education Economics*, vol. 3, no. 2, pp.173-203.

Harding, A. (1996), *Microsimulation and Public Policy*, North-Holland, Amsterdam.

Harding, A. and Gupta, A. (2007a), *Modelling our Future: Population Ageing, Social Security and Taxation,* North-Holland, Amsterdam.

Harding, A. and Gupta, A. (2007b), "Introduction and overview", in Harding, A. and Gupta, A. (eds), *Modelling our Future: Population Ageing, Social Security and Taxation,* North-Holland, Amsterdam.

Harding, A., Payne, A., Vu QN. and Percival, P. (2006), *Trends in Effective Marginal Tax Rates, 1996-97 to 2006-07,* AMP NATSEM Income and Wealth Report Issue 14, Sydney, September, (available from www.amp.org.au/ampnatsemreports).

Heiland, F. (2003). "The collapse of the Berlin wall: Simulating state-level east to west German migration patterns". *Agent-based computational demography*. F. Billari and A. Prskawetz, Physica-Verlag: 73-96.

Hills, T. and P. Todd (2008). "Population heterogeneity and individual differences in an assortative agent-based marriage and divorce model (MADAM) using search with relaxing expectations." *Journal of Artificial Societies and Social Simulation* 11(4): 5.

Holm E., Holme K., Mäkilä K., Mattsson-Kauppi M. and Mörtvik, G. (2002) *The SVERIGE spatial microsimulation model. Content, validation, and example applications*. GERUM.

Holm E, Lindgren U, Lundevaller E and Strömgren M. (2006) The SVERIGE spatial microsimulation model, Paper presented at the 8th Nordic Seminar on Microsimulation Models, Oslo, 8-9 June 2006. (http://www.ssb.no/misi/papers/holm_einar_svesim.doc)

Immervoll, H., Levy, H., Lietz, C., Mantovani, D., O'Donoghue, C., Sutherland, H. and Verbist, G. (2006), "Household incomes and redistribution in the European Union: quantifying the equalizing properties of taxes and benefits", in Papadimitriou, DB. (ed), *The Distributional Effects of Government Spending and Taxation,* Palgrave Macmillan, Basingstoke, Hampshire, pp. 135-165.

Johnson, RW., Favreault, M. and Goldwyn, J. (2003), *Employment, Social Security, and Future Retirement Outcomes for Single Mothers*, Working Paper CRR WP 2003-14, Center for Retirement Research, Boston College.

Kelly, S. (2001), Trends in Australian wealth – new estimates for the 1990s, Paper presented at the 30th Annual Conference of Economists, University of Western Australia, 23-26 September.

Kelly, S. (2002), Simulating future trends in wealth inequality, Paper presented at the 31st Annual Conference of Economists, Adelaide, South Australia, 1-3 October.

Kelly, S. (2003), *Estimating the Wealth of Australians: A New Approach Using Microsimulation*, PhD thesis, University of Canberra, Canberra. 48 *Discussion Paper no.* 63

Kelly, S. (2007), "The Australian DYNAMOD model", in Gupta, A and Harding, A. (eds), *Modelling our Future: Population Ageing, Health and Aged Care,* North-Holland, Amsterdam.

Kelly, S. and King, A. (2001), "Australians over the coming 50 years: providing useful projections", *Brazilian Electronic Journal of Economics,* vol. 4, no. 2.

Kelly, S., Percival, R. and Harding, A. (2002), "Women and superannuation in the 21st century: poverty or plenty?", in Eardley, A. and Bradbury, B. (eds), in *Competing Visions: Proceedings of the National Social Policy Conference*, SPRC Report 1/02, University of New South Wales.

King, A. and Bækgaard, H. (1997), DYNAMOD*:* the experience and lessons, Note prepared for Uppsala meeting on DYNACAN, National Centre for Social and Economic Modelling, University of Canberra, May.

Klevmarken, NA. (2005), Dynamic microsimulation for policy analysis – problems and solutions, Paper presented at the 2005 Australian Conference of Economists, University of Melbourne, 26-28 September.

Knudsen, P. (2006), Lessons from the experience of reprogramming MOSART from Simula to C#, Paper presented at the Dynamic Microsimulation Modelling and Public Policy lnternational Conference, London School of Economics, 29 September.

König, A., Möhring M., and K. Troitzsch (2003). “Agents, hierarchies and sustainability”. *Agent-based computational demography.* F. Billari and A. Prskawetz, Physica-Verlag.

Krupp, H. (1986). “Potential and limitations of microsimulation models”. *Microanalytic Simulation Models to Support Social and Financial Policy*. G. Orcutt. Amsterdam, Elsevier Science Publishers: 31-43.

Le Bras H. (1973). « Parents, grand-parents, bisaïeux ». *Population* 28 (1): 9-38.

Legendre F., Lorgnet J. P.and Thibault F. (2001), « Les modèles socio-économiques de microsimulation ; panorama et état des lieux pour la France ». *Recherches et Prévisions*, n 66, p. 11-32.

Lloyd, R. (2007), “STINMOD: use of a static microsimulation model in the policy process in Australia”, in Harding, A, and Gupta, A. (eds), *Modelling our Future: Population Ageing, Social Security and Taxation,* North-Holland, Amsterdam.

McKay, S. (2003), Dynamic microsimulation at the U.S. Social Security administration, Paper presented at the International Microsimulation Conference on Population Ageing and Health, Canberra, 11 December.

Morrison, R. (2007a)*, “*DYNACAN”, in Gupta, A and Harding, A. (eds), *Modelling our Future: Population Ageing, Health and Aged Care,* North-Holland, Amsterdam. *Problems and Prospects for Dynamic Microsimulation: A Review and Lessons for APPSIM* 49

Morrison, R. (2007b)*,* “Pension analysis using dynamic microsimulation: adding private pensions to the Canadian DYNACAN model”, in Harding, A. and Gupta, A. (eds), *Modelling our Future: Population Ageing, Social Security and Taxation*, North-Holland, Amsterdam.

Morrison, RJ. (2000), “DYNACAN”, in Gupta, A. and Kapur, V. (eds), *Microsimulation in Government Policy and Forecasting*, North-Holland, Amsterdam, pp. 341-360.

Morrison, RJ. (2006a), DYNACAN’s methods for gauging the after-tax impacts of Canada pension plan, Paper presented at the Dynamic Microsimulation Modelling and Public Policy lnternational Conference, London School of Economics, 29 September.

Morrison, RJ. (2006b), Demonstration of new ‘Results Browser’ software that facilitates interactive access to DYNACAN model results’, Paper presented at the Dynamic Microsimulation Modelling and Public Policy lnternational Conference, London School of Economics, 29 September.

Murphy, B. (2000), ‘SPSD/M’, in Gupta, A. and Kapur, V. (eds), *Microsimulation in Government Policy and Forecasting*, North-Holland, Amsterdam, pp 587-592.

Nelissen, JHM. (1996), “Social security and lifetime income redistribution”, in Harding, A. (ed), *Microsimulation and Public Policy,* North-Holland, Amsterdam, pp. 267-292.

Neufeld, C. (2000), “Alignment and variance reduction in DYNACAN”, in Gupta, A. and Kapur, V. (eds), *Microsimulation in Government Policy and Forecasting*, North-Holland, Amsterdam, pp. 361-382

NOU (2004), *Modernisert Folketrygd. Bærekraftig Pensjon for Framtida,* The Ministry of Finance and The Ministry of Social Affairs, Oslo. (In Norwegian).

O'Donoghue C. (2002), "Redistribution over the lifetime in the Irish tax-benefit system: an application of a prototype dynamic microsimulation model for Ireland", *Economic and Social Review*, vol. 32, no. 3, pp. 191-216.

O'Donoghue, C. (2001), "Dynamic microsimulation: a methodological survey", *Brazilian Electronic Journal of Economics,* vol. 4, no. 2.

O'Hare, J. (2000), "TRIM3", in Gupta, A. and Kapur, V. (eds), *Microsimulation in Government Policy and Forecasting*, North-Holland, Amsterdam, pp. 581-586.

Orcutt, G. (1957). "A new type of socio-economic system." *The Review of Economics and Statistics* 39(2): 116-123.

Orcutt, G. (1961). *Microanalysis of socioeconomic systems: A simulation study*. New York, Harper and Row.

Orcutt, G. (1957), "A new type of socio-economic system"*, Review of Economics and Statistics,* vol. 58, no. 2, pp. 773-797.

Orcutt, G., Caldwell, S., Wertheimer, R., Franklin, S., Hendricks, G., Peabody, G., Smith, J. and Zedlewski, S. (1976), *Policy Exploration through Microanalytic Simulation*, The Urban Institute, Washington DC.

Orcutt, G., Greenberger, M., Korbel, J. and Wertheimer, R. (1961)*, Microanalysis of Socioeconomic Systems: A Simulation Study*, Harper and Row, New York.

Orcutt, G., S. Caldwell, et al. (1976). *Policy exploration through microanalytic simulation*, University Press of America.

Pancs, R. and N. Vriend (2007). "Schelling's spatial proximity model of segregation revisited." *Journal of Public Economics* 91(1-2): 1-24.

Pennec S. (1997). "Four-Generation Families in France." *Population an English Selection*, 9: 75-100.

Pennec, S; and Bacon, B ,2007, *APPSIM - Modelling Fertility and Mortality*, APPSIM Online Working Paper WP7, NATSEM, University of Canberra, Canberra (available from http://www.canberra.edu.au/centres/natsem/publications)

Productivity Commission 2005, *Economic Implications of an Ageing Australia, Overview*, Productivity Commission, Canberra (available from www.pc.gov.au).

Ruggles S. (1987). *Prolonged connections: The rise of the extended family in nineteenth century England and America*. Madison; London, The University of Wisconsin Press, 283 p.

Tomassini C., Wolf D. (2000). "Shrinking Kin Networks in Italy Due to Sustained Low Fertility". *European Journal of Population* 16 (4): 353-372.

Schelling, T. (1971). "Dynamic Models of Segregation." *Journal of Mathematical Sociology* 1(2): 143-186.

Scott, A., Evandrou, M., Falkingham, J., Johnson, P. and Zaidi, A. (2003), Beyond 2020: Lessons from the Development of the SAGE Dynamic Microsimulation Model for Britain, ESRCSAGE Discussion Paper no.14, London School of Economics, (available from www.lse.ac.uk/collections/SAGE)

Smith, K. (2003), *How will Recent Patterns of Earnings Inequality affect Future Retirement Incomes?* Project report 2003-106, AARP Public Policy Institute, Washington, DC, (available from (http://research.aarp.org/econ/2003_06_retire.html). 50 *Discussion Paper no.* 63

Spielauer, M. (2002), *Dynamic Microsimulation of Health Care Demand, Health Care Finance and the Economic Impact of Health Behaviour, Part 2: Survey and Review*, Interim Report, IR-02- 036/May, Laxenburg, Austria (available from www.iiasa.ac.at).

Sundberg, O. (2007), "SESIM (longitudinal dynamic microsimulation model)", in Gupta, A. and Harding, A. (eds), *Modelling out Future: Population Ageing, Health and Aged Care*, North- Holland, Amsterdam.

Sutherland, H. (1997), *The EUROMOD Preparatory Study: A Summary Report*, DAE Working Papers Amalgamated Series No. 9725, Department of Applied Economics, University of Cambridge.

Sutherland, H. (2007), "EUROMOD – the tax-benefit microsimulation model for the European Union", in Gupta, A. and Harding, A. (eds), *Modelling out Future: Population Ageing, Health and Aged Care*, North-Holland, Amsterdam.

Todd, P. (1997). "Searching for the next best mate. Simulating Social Phenomena". R. Conte, R. Hegselmann and P. Terna. Berlim, Springer-Verlag: 419-436.

Todd, P. and F. Billari (2003). Population-wide marriage patterns produced by individual mate-search heuristics. *Agent-based computational demography*. F. Billari and A. Prskawetz, Physica-Verlag.

Toohey, M. and Beer, G. (2004), 'Financial incentives for working mothers under a new tax system', *Australian Journal of Labour Economics,* vol. 7, no. 1, pp. 53-69.

Treasury (2002), '*Intergenerational Report 2002-03*, 2002-03 Budget Paper no. 5, Commonwealth Treasury of Australia, Canberra.

Uccello, C. E., Favreault, M., Smith, K. E., and Thompson, L. H. , (2003), Simulating the distributional consequences of personal accounts: sensitivity to annuitization options, Working Paper CRR WP 2003-17, Center for Retirement Research, Boston College, October.

Van Imhoff, E. and W. Post (1998). "Microsimulation methods for population projection." *Population: An English Selection* 10(1): 97-138.

Von Neumann, J. and A. Burks (1966). *Theory of self-reproducing automata*. Urbana, IL, University of Illinois Press.

Wachter K., Blackwell D., and E. Hammel (1997). "Testing the Validity of Kinship Microsimulation" *Journal of Mathematical and Computer Modeling* 26: 89-104.

Waddell, P. (2002) UrbanSim: Modeling Urban Development for Land-Use, Transportation and Environmental Planning: Journal of the American Planning Association, 68 (3) 297–314.

Willekens, F. (2006), *Description of the micro-simulation model (Continuous-time micro-simulation)* Deliverable D8 (first part), MicMac Bridging the micro-macro gap in population forecasting, NIDI, The Netherlands, September

Wolfson, M. (1988), "Homemaker pensions and lifetime redistribution", *Review of Income and Wealth*, vol. 34, no. 3, pp. 221-250.

Wu, B., M. Birkin, and P.H. Rees. (2008). "A spatial microsimulation model with student agents." *Computers, Environment and Urban Systems* 32(6): 440-453.

Zaidi, A. and Rake, K. (2001), *Dynamic Microsimulation Models: A Review and Some Lessons for SAGE*, SAGE Discussion Paper no.2, London School of Economics

Zedlewski, SR. (1990), "The development of the dynamic simulation of income model (DYNASIM)", in Lewis, GH. and Michel, RC. (eds), *Microsimulation Techniques for Tax and Transfer Analysis*, Urban Institute Press, Washington, DC, pp. 109-171.