International Institute for
Applied Systems Analysis
Schlossplatz 1
A-2361 Laxenburg, Austria

Tel: +43 2236 807 342
Fax: +43 2236 71313
E-mail: publications@iiasa.ac.at
Web: www.iiasa.ac.at

Interim Report **IR-02-032/April**

Dynamic Microsimulation of Health Care Demand, Health Care Finance and the Economic Impact of Health Behavior

Part I: Background and a Comparison with Cell-Based Models

Martin Spielauer (spielaue@iiasa.ac.at)

Approved by

Landis MacKellar (mckellar@iiasa.ac.at)
Social Security Reform

April 2002

Contents

1. Introduction 1
2. What is dynamic microsimulation? 2
3. Classification 5
    The purpose: projection versus explanation 5
    General versus specialized models 6
    Cohort versus population models 7
    Steady-state versus forecasted projections 7
    Open versus closed population models 8
    Data-based versus agent-based models 8
4. Microsimulation modeling in the context of the life-course paradigm 9
    The human life course 11
5. Microsimulation versus cell-based approaches 14
    The representation of populations 14
    Modeling of population dynamics 15
    Linking microsimulation and cell-based macrosimulation models 16
    Areas of possible application of the microsimulation approach 17
    Strengths and limitations of microsimulation 21
Summary and conclusions 23
References 24

Abstract

Cell-based health care models, as well as macro-level projections of future population and economic trends used as input to health care models, are limited to a few variables, which makes microsimulation an interesting modeling option, especially as it allows for modeling of the interaction of demographic with social, environmental and economic variables. Micro-approaches can incorporate the wealth of substantive analysis gained from a large number of micro- and macro-level studies with regard to demographic, economic and health behavior. Compared to cell-based macro models, microsimulation can produce useful projections for the analysis of different health-related phenomena considering additional dimensions, i.e., detailed issues regarding health care finance (insurance schemes, individual accounts etc.) and individual risk exposure.

This paper constitutes the first part of an investigation of the potential of dynamic microsimulation for the modeling and projection of health care demand, health care finance and the economic impact of health behavior. The main purpose of this part is to provide a brief theoretical background with regard to the dynamic microsimulation approach and a comparison of the microsimulation approach with the cell-based macro approach. Starting with a definition of dynamic microsimulation and a classification of the types and approaches, microsimulation modeling is brought into the context of the life-course paradigm. This paradigm, meanwhile being the dominant paradigm in demography, can also be a useful organizational principle for the study and projection of health-related phenomena using microsimulation. Microsimulation is then compared with cell-based approaches, and the potential strengths as well as drawbacks of the microsimulation approach with regard to health care modeling are investigated. Dynamic microsimulation might turn out to be increasingly appropriate as a modeling approach in this field, which is currently dominated by cell-based macro-models.

About the Author

Martin Spielauer is a Research Scholar with the Social Security Project at IIASA and Leader of the Department for Socioeconomic Analysis, Simulation and Surveys at the Austrian Institute for Family Studies (ÖIF) in Vienna.

Dynamic Microsimulation of Health Care Deamnd, Health Care Finance and the Economic Impact of Health Behavior

Part I: Background and a Comparison with Cell-Based Models

Martin Spielauer

1. Introduction

Cell-based health care[1] models, as well as macro-level projections of future population and economic trends used as input to health care models, are limited to a few variables, which makes microsimulation an interesting modeling option, especially as it allows for modeling of the interaction of demographic with social, environmental and economic variables. Micro-approaches can incorporate the wealth of substantive analysis gained from a large number of micro- and macro-level studies with regard to demographic, economic and health behavior. Compared to cell-based macro models, microsimulation can produce useful projections for the analysis of different health-related phenomena considering additional dimensions, i.e., detailed issues regarding health care finance (insurance schemes, individual accounts etc.) and individual risk exposure (such as smoking). Additional examples on the demographic side are educational composition, rural/urban differentials, household structures and family networks, which become increasingly important in the context of the ongoing demographic change, that goes hand in hand with changes in health status, health care demand and availability of informal care supply. In the context of population forecasting, most of these new challenges have been discussed in detail in Lutz et al. (1999).

The use of microsimulation models in health studies is not limited to projections, and ultimately forecasts and policy recommendations. Models can also be designed and used to study the implications of certain assumptions and thereby to develop and test theories. Regardless of the main intention – prediction versus explanation – an important purpose of modeling is to clarify concepts. The process of modeling itself can produce valuable insights on the subject being modeled or may help to identify internal inconsistencies and gaps of theories, since, for the translation of a theory into a simulation model, everything needs to be quite explicit.

---

[1] The term "health care model" is used for the broad range of models addressed in this paper, including models of health care demand, health care finance and health behavior, and their economic impacts. The term "health studies" is also used in this broad context.

This paper constitutes the first part of an investigation of the potential of dynamic microsimulation for the modeling and projection of health care demand, health care finance and the economic impact of health behavior. The main purpose of this part is to provide a brief theoretical background with regard to the dynamic microsimulation approach and a comparison of the microsimulation approach with the cell-based macro approach. Starting with a definition of dynamic microsimulation and a classification of the types and approaches, microsimulation modeling is brought into the context of the life-course paradigm. The impact of massive social change on people's lives has become a vital area of research, and great progress has been made in the ways of studying how lives change over time. In this context, an important paradigm shift can be observed during the last decades that led to the integration of structural and dynamic approaches to the life-course paradigm. This paradigm, meanwhile being the dominant paradigm in demography, can also be a useful organizational principle for the study and projection of health-related phenomena using microsimulation. Microsimulation is then compared with cell-based approaches, and the potential strengths as well as drawbacks of the microsimulation approach with regard to health care modeling are investigated. Dynamic microsimulation might turn out to be increasingly appropriate as a modeling approach in this field, which is currently dominated by cell-based macro-models.

2. What is dynamic microsimulation?

Microsimulation is a rather confusing term, both due to the wide range of models it addresses and the very different concepts in which the terms simulation and modeling are used. Very generally, a microsimulation model can be defined as a model which uses simulation techniques and which takes micro-level units – in the social sciences usually individuals, families or firms - as the basic units of analysis (O'Donoghue, 2001). Following this broad definition, dynamic microsimulation would include a broad variety of models and modeling approaches ranging from data-based empiric dynamic microsimulation to concept-driven microsimulation based on the distributed artificial intelligence approach[2]. A detailed classification is given in the next chapter.

In the social sciences, dynamic microsimulation was introduced in the late 1950s dominantly in the form of "empirical" dynamic microsimulation models, which are models designed and used operatively for forecasting and policy recommendations (Klevmarken, 1997). This tradition can be traced back to a "direct" and an "indirect" source. The direct source of dynamic microsimulation can be found in Guy Orcutt's idea about mimicking natural experiments in economics, which led to the development of the

[2] As data-based dynamic microsimulation and agent-based simulation evolved 'in almost total ignorance of each other' (Troitzsch, 1996), the term dynamic microsimulation is also often used as a term for the first approach as distinguished from agent-based simulation. Inasmuch as data-based empirical models move from 'black box' models of behavior to models that incorporate theory and individual goal orientation, the distinction becomes more difficult. Both traditions increasingly use concepts of each other and a synthesis might be desirable. To support this view, a broad definition of dynamic microsimulation was chosen in the context of this paper, even though it clearly concentrates on data-based microsimulation.

DYNASIM model (Orcutt, 1957). The indirect source lies in the static tax benefit microsimulation models, resulting from the increased interest among policy-makers in distributional studies. As attempts are made to enlarge the initially static tax benefit models with behavioral models to capture the second-order effects of policies and to simulate behavior over time, these tax-benefit models approach Orcutts's DYNASIM and/or its various successors. This tradition is also labeled data-based microsimulation, as it is usually based on empirical micro-data and dominated by statistical and econometric behavioral models. In general, there are various additional ways of modeling the behavior of the micro-units, ranging from simple rules to economic optimization behavior to agent-based models. In data-based models, theory is often sacrificed in favor of a highly detailed model with a good fit to the data. Behavior is mostly modeled implicitly, and so are corresponding assumptions, which can make models difficult to understand. In contrast, "abstract models" incorporate behavior explicitly. These models are rather designed and used to test and develop theories, i.e., for explanation rather than prediction. This also holds true for context-driven agent-based microsimulation. Agents are defined by their behavior and act according to the environmental context they are placed in. Context-driven microsimulation goes back to the 1980s.

The term **micro** indicates the level of analysis, in the social sciences usually individuals or households. In contrast to static microsimulation models, in which these micro-units are only used as rather passive accounting units, the common element of all dynamic microsimulation approaches and traditions is that they analyze the behavior of a system by using the characteristics of micro-units that are changed – or autonomously change – according to a behavioral model. The main idea of microsimulation is that processes resulting from the actions and interactions of a large number of micro-units can be explained best by looking at the micro-units and their behavior. One expects to find more stable behavioral relationships on the micro-level than in aggregated data that are affected by structural changes when the number or size of the micro-units in the population changes, even if the behavior of the individual micro-units and their individual characteristics do not change. These micro-units might be particles moving in line with probability laws, e.g., in fluids or thermodynamics, a field in which microsimulation was first introduced. They might also represent artificial species of 'artificial societies' as is the case in most agent-based simulations. But they can also represent individuals, families or households of empirical populations, as it is the case in 'data-based' microsimulation.

Beside this 'direct' source of **dynamic** microsimulation, the modeling and simulation of dynamics over time and in response to context changes (i.e., policy response) is not introduced in static microsimulation models. Tax-benefit systems are the most typical application of static microsimulation, where individuals or households (represented in a micro-database) are only used as accounting units with the necessary characteristics to calculate taxes and benefits. Reduced to its bare essentials, a data-based static microsimulation model consists of two parts (Martini and Trivellato, 1997):

- a baseline database: a data set containing information on individual or family/household units, in particular socio-demographic characteristics and economic information that bears a relationship with a set of policies;

- a set of accounting rules: these are computer language instructions that produce, for each unit, the provisions of existing or alternative tax and transfer systems, or other relevant institutional features.

The construction of representative data sets containing all necessary variables, and modeling at least part of a complex tax-benefit system, absorbed all the resources in the early days of microsimulation. The work of Pechman and Okner (1974) analyzing the redistribution effects of the US tax system represents the most celebrated example of this type of research. Historically, microsimulation moved from the description of the distributional impact of the existing tax and transfer system to a second stage, in which it became a tool for understanding the differential impact of alternative proposals for reforming existing systems. A more recent example is the investigation of the treatment of the family in income tax systems across Europe by O'Donoghue and Sutherland (1999). In this study different European tax systems were examined for the UK, using the tax-benefit microsimulation model POLIMOD (Sutherland, 1995a, 1995b).

In static models, time – if introduced at all – has no effect on individual characteristics, such as to reflect the future composition of the population; the dataset is simply re-weighted at each time step (a process called static aging). Dynamic microsimulation includes behavior over time, a set of behavioral relationships which varies greatly in scope and importance across models. This can be of two types:

- behavior that produces events that take place over time such as demographic events, i.e. marriage, divorce, death, etc., and economic events such as leaving the labor force;

- behavior producing feedback reactions of individuals and/or families to changes in external circumstances, notably to changes in public policies.

In dynamic microsimulation, the behavior of the micro-units is modeled – most importantly, its behavior over time. Various approaches can be used in order to model behavior over time ranging from simple transition tables to elaborated econometrical models, neuronal networks or artificial intelligence. Typical behavioral models are statistical models that, for a given set of personal characteristics, determine probabilities for a defined set of possible transitions like marriage, pregnancy or death. Monte-Carlo simulation is then used to determine if a transition takes place in the simulation experiment. This allows to dynamically update personal characteristics over time and to add and remove micro-units to or from the population due to birth, death or migration. Dynamic microsimulation simultaneously addresses point-in-time “snapshot” distribution issues as well as longitudinal "life-path" issues, making it a powerful and flexible tool for policy analysis. Another type of dynamic behavior is policy response, which might be modeled using econometrical approaches or based on theory such as utility maximization. Again, there is a wide range of possibilities for modeling individual behavior, from the modeling of a rational forward-looking utility-optimizing "homo economicus" to more realistic human behavior including learning processes, etc., as is done in agent-based simulation based on the artificial intelligence approach.

Dynamic microsimulation **models** are the result of a synthesis of various models usually including a population database as model representation of an empirical or artificial society, model representations of alternative tax-benefit systems (like health-care finance) as well as behavioral models as outlined above.

The use of the term **simulation** can be quite misleading in labeling microsimulation models, as simulation (among other applications as a techniques) is a particular type of modeling in itself (Gilbert and Troitzsch, 1999), but not all microsimulation models are “simulation models” in this sense. Simulation modeling constitutes a research method that is quite different from the logic of statistical modeling. While agent-based models are typical simulation models in this sense, data-based models are usually statistical models and simulation does not “add” anything to these models but is used as a technique to “run” them into the future, i.e., in the form of a Monte-Carlo simulation. One main difference lies in the notation: statistical models are expressed in statistical equations, whereas simulation models are usually expressed in the form of computer programs. In other words, in data-based microsimulation it is possible to distinguish the model itself from the computer software used to “run” the model, a process that might be done by or include (Monte-Carlo) simulation, with the whole exercise also being called a simulation in the sense that it mimics a “real” experiment using a model. In contrast to statistical models and their notation, simulation models have a considerably extended scope as they are not restricted to a theory that can be formalized in mathematical notation, but underlie the much wider notation of computer languages. In this respect, computer simulation models represent a third domain, complementing both natural language and mathematical/statistical analysis.

Due to their complexity and the quantity of data to be processed in microsimulation models, these are inevitably run on **computers**. To the degree that computer programming itself can be seen as a modeling exercise, microsimulation modeling approaches also correspond to some degree to programming paradigms. Static microsimulation can be technically described as the manipulation of a population micro-database by computer procedures that produce, for each unit, the provisions of existing or alternative tax and transfer systems, or other relevant institutional features. There is a clear correspondence with the procedural programming paradigm that clearly distinguishes data from codes. To the extent that individual behavior is introduced, object oriented programming becomes the more adequate programming paradigm, as individuals can be described much better as objects that encapsulate both, the data structure holding all individual characteristics that describe the status of an object, as well as the methods that describe the behavior with changes these characteristics. Agents as modeled in agent-based simulation directly correspond to agents in the computer terminology that can be described as “extended objects”, being characterized by purposefulness, autonomy, and reactivity.

3. Classification

The purpose: projection versus explanation

A first distinction can be made with regard to the intended use of a microsimulation model which can either lie in projections (and consequently in producing forecasts and policy recommendations), or in the explanation of social phenomena. In this sense, microsimulation models can be empirical models or abstract models. If designed and used operatively for forecasting and policy recommendations, such models ‘need to be firmly based in an empirical reality and its relations should have been estimated from real data and carefully tested using well-established statistical and econometric methods. In this case the feasibility of an inference to a real world population or economic process is of great importance’ (Klevmarken, 1997). In contrast, abstract models are rather designed and used to study the implications of certain assumptions without an ambition to produce reliable forecasts.

Typical demographic applications of abstract models in demography and population studies are models for partner matching and geographical segregation with ABCD (Agent-based Computational Demography), which is currently becoming a vital area of research that is not primarily intended to forecast the behavior of actual populations, but to study dynamics and patterns of artificial societies resulting from the interactions of artificial species. By "growing" these societies, simulations serve as a tool to develop and test theories that might help to explain human behavior, on the assumption that artificial societies might show similar behavioral patterns as empirical ones. An example in which agent-based simulation was successfully used to reproduce observed residential patterns is a model developed for Israeli communities by Benenson and Omer (2001). Simulation models of this type might also be interesting in the field of health behavior, where agents might follow conflicting goals (e.g., health, pleasure) in different contexts, i.e., with regard to their own responsibilities, job situations and the observed behaviors of other agents.

Regardless of the intention – prediction versus explanation – an important purpose of modeling is to clarify concepts (Burch, 1999). The social sciences are over-rich in descriptive theories that have limited practical application (Lane, 1999). The process of modeling itself can produce valuable insight into the subject being modeled. The act of translating a theory into a simulation model requires that everything be made explicit and quickly exposes internal inconsistencies and gaps. In contrast to pure mathematical models and mathematical notation, simulation models have a considerably extended scope, as they are not restricted to a theory that can be formalized in mathematical notation, but underlie the much wider notation of computer languages. In this respect, computer simulation represents a third domain, complementing both natural language and mathematical/ statistical analysis.

General versus specialized models

In the data-based tradition, many microsimulation models were developed for a wide range of purposes and are therefore rather general models, typically covering the whole household sector of a country. Such general models exist for various countries and are reviewed in the second part of this paper series. Apart from these general models, there are also very specialized microsimulation models that typically concentrate on one specific behavior (most prominently, the labor market behavior) or population segment. An example of the latter is the British NCCSU Long-term Care Model (Hancock et al., 2002). This model simulates the incomes and assets of future cohorts of older people and their ability to contribute towards care-home fees. It thereby concentrates on the simulation of the means test of long-term care policies, with the results fed into a macro model of future demands and costs.

The same distinction can be made for context-driven models that would be constructed to model specific behaviors like matching processes or to model whole economies, like the American ASPEN model (Pryor et al., 1996) which was developed in order to study the consequences of various legal, regulatory and political changes.

Cohort versus population models

Cohort models age a single cohort over his entire lifetime, predicting each individual's major life-course events. In contrast, dynamic population microsimulation models age entire cross-sections. Studies typically done with single cohort models investigate lifetime income and interpersonal distributions. Examples of this kind of models include the HARDING and the LIFEMOD models developed in parallel for Australia (HARDING) and Great Britain (LIFEMOD) (Falkingham and Harding, 1996). This kind of models typically assumes a steady-state world, i.e., the HARDIG cohort is 'born' into and lives in a world that looks like Australia in 1986.

Several limitations of these cohort models are derestricted when simulating a whole cross-section population, including issues of demographic change and distributional issues between cohorts. Population models are usually far more complex and demanding with regard to data. Some models only focus on a certain age range, like women in their reproductive age, e.g. in FAMSIM (Spielauer, 2000), or the retired population, e.g. in the NCCSU Long-term Care Model.

In population models, the ability to isolate single cohorts with regard to financial accounting will support the study of the sustainability of a social security system and the study of intergenerational fairness issues.

Steady-state versus forecasted projections

Steady-state assumptions are common, especially in single-cohort models and as 'benchmark' scenarios in population models. In these models, individuals are aged in an unchanging world with regard to the environmental context such as economic growth and policies, and the individual behavior is "frozen" not allowing for cohort or period effects. As a today's population cross-section does not result from a steady-state world, "freezing" individual behavior and the socioeconomic context can also serve to isolate and study future dynamics and phenomena resulting from past changes, such as the population momentum.

For many models steady-state assumptions (if made) only serve as a benchmark; usually the models try to include (and produce) forecasts with regard to the future world. "Pure" steady-state models would be inappropriate for studying micro-macro interactions, such as the impact of demographic change on social security systems, but "freezing" some behaviors will help to isolate the contribution of single processes to the future dynamics of the whole system.

Open versus closed population models

In microsimulation the terms open and closed population usually corresponds to whether the matching of spouses is restricted to persons within the population or whether spouses are imputed. In open population models partners are usually attached as attributes to the "dominant" individuals of the population with characteristics synthetically generated or sampled from a host population. In contrast, closed models allow to track kinship networks and also enforce more consistency, given a large enough population to find appropriate matches. Major drawbacks of closed models are the computational demands associated with mate matching and sampling problems. In a starting population derived from a sample, the model may not be balanced with respect to kinship linkages other than spouses, as a person's parents and siblings are not included in the base population if not living in the same household. (Toder et al., 2000).

A related topic is how to model immigration. Approaches range from the cloning of existing ‘recent immigrants’ to sampling from a host population or even from different ‘pools’ of host populations representing different regions.

Data-based versus agent-based models

As mentioned above, in data-based microsimulation a clear distinction can be made between the data representing the population, the model that determines the behavior, the Monte-Carlo simulation typically used to run the model, and the software necessary for the whole exercise. Associated with this type of microsimulation are usually micro-econometric and statistical models, whose behavior is usually expressed in transition probabilities or duration times. Two main approaches can be distinguished according to the way of modeling time: (1) the continuous-time competing-risk approach to dynamic microsimulation modeling, and (2) approaches based on a discrete-time framework. These issues are explored in detail in Galler (1996).

Agent-based microsimulation, based on the distributed artificial intelligence approach, represents a very different modeling tradition. Agents are defined by their behavior and act according to the environmental context they are placed in. As stated before, today the “artificial society” approach is mainly used to explore theories. A good example is the Evolution of Organized Societies (EOS) project set out to explore theories accounting for the growth of social complexity among the human population in the Upper Paleolithic period in south-western France (Doran et al., 1994, quoted from Gilbert and Troitzsch, 1999). Micro-units are "intelligent" and acting agents, having goals and following rules. The following features characterize agents:

- agents have receptors, they get input from the environment;
- agents have cognitive abilities, beliefs and intentions;
- agents can follow different rules and make decisions which rules to follow;
- agents live in groups of other agents and interact;
- agents can act and act simultaneously;
- agents can learn.

A synthesis might be desirable and could be approached by combining or allowing various "rules of motion" and population types according to the respective research questions and goals. As an example, fertility might be modeled in a two-step process by combining a child-heaving decision model – a model that might incorporate theory and could be agent-based – with a (statistical) waiting-time model (Vencatasawmy, 2002). Similar approaches might also be appropriate in the modeling of health risks, where the choice for a given behavior (like smoking) can be distinguished from the modeling of the consequences of this behavior. To the extent that dynamic microsimulation incorporates concepts of goal orientation, planned behavior and strategic adaptation, the more attractive it might become as a tool in demographic research in the context of its dominant paradigm: the life-course paradigm.

4. Microsimulation modeling in the context of the life-course paradigm

The massive social and demographic change in the last decades went hand in hand with tremendous technological progress, with computers now being a powerful and indispensable tool in various fields of research. Their ability to process large amounts of data has boosted data collection, enabled new survey designs and ways of data analysis. In general, the impact of massive social change on people's lives has become a vital area of research, and great progress has been made in the ways of studying how lives change over time. In this context, an important paradigm shift can be observed in the last decades that led to the integration of structural and dynamic approaches to the life-course paradigm, which has meanwhile become the dominant paradigm in demography. It combines several major theoretical and empirical streams of research, connecting social change, social structure, and individual action (Giele and Elder, 1998). This chapter outlines the main recent paradigm shifts in social sciences and puts microsimulation into the context of the life-course paradigm that can be seen as useful organizational principle for the study and projection of population phenomena including health issues by microsimulation.

Demography and health studies involve a variety of research disciplines and are therefore not only influenced by general changes and shifts in the focus of attention but also benefit from their developments and innovations. The changes that can be observed can occur along four dimensions (Willekens, 1999):

- from structure to process
- from macro to micro
- from analysis to synthesis
- from certainty to uncertainty.

The change from structure to process shifts the focus of attention from a static view of the social systems to the dynamics of the systems over time and to the processes generating the dynamics. While this "transition from entity-oriented perception of reality to process-oriented perception" was made by nearly every social and natural science (Willekens, 1999; 4), its importance increases with the speed of the observed social and demographic changes and the various new questions raised by these changes. The focus on processes brings in various new concepts, with causality and time being among the most important. Various population phenomena are characterized by their rapid change over time, and a substantial research effort is required to identify and understand the underlying processes generating these phenomena. Good examples are low fertility, increasing divorce rates and changes in the distribution of income and wealth. An example with regard to health issues is the excess-mortality of adult males in various post-communist countries as a consequence of various reasons including economic stress, unemployment and alcohol abuse.

The importance of time is also increasingly recognized in the field of policy analysis, where the attention shifts to the long-term dynamics and the sustainability of tax-benefit and social security systems. In studying distribution effects of policies, time adds a new dimension to research, as distribution effects are not only analyzed in a cross-sectional view for a given time, but also over time, between cohorts and over generations. Regarding health care costs, such studies can also include the distributional impacts of policies between different risk groups, such as, e.g., smokers and non-smokers, and might furthermore include excise taxes on certain products, etc. This shift in focus is mirrored in the development of microsimulation models, both by the increasing efforts undertaken to extend static models to dynamic models, and generally by their increasing importance as a research tool that can handle dynamic processes over time.

The second dimension considers the level of analysis. Social sciences tend to move from macro- to micro-explanations and to interpret changes on the macro-level as results of actions taken by individual agents and their interactions. These interactions also include reactions and feedback of individual agents in connection with changes in their environment, i.e. changes on the macro-level that form the context of individual decisions and actions. Contexts interacting with health comprise economic conditions and incentives, including those stemming from insurance systems themselves. Again, there is a direct correspondence between this general shift to micro-level research and microsimulation.

The third dimension looks at the shift from analysis to synthesis. When shifting the focus of attention from structure to process, research increasingly tends not to stop at the analysis of these processes and the resulting structures. The identification of the elementary processes that generate the complex dynamics of a system are indispensable for understanding these dynamics, but also have to be 'put together' by way of synthesis. In this way, system dynamics can be projected under different assumptions. As described in more detail below, the life course may be viewed as a combination of a large number of elementary processes. The challenge is to detect the elementary processes and the rules that link them. Microsimulation is the main tool for linking multiple elementary processes in order to generate complex dynamics and to quantify what a given process contributes to the complex pattern of change.

The fourth shift is based on the insight that uncertainty is associated with many events. Agents have only limited control over most events and their exact timing. Hence the individual likelihood that certain events will or will not happen becomes an important issue. This holds true for many phenomena and events studied in demographic research: pregnancy is a good example. While the degree of planning might vary, the exact timing cannot be controlled even though probabilities might be well known. This also applies to health issues, where behavioral choices have a strong impact on health risks. Again, microsimulation can be used to study these random distributional effects. Due to the inclusion of stochastic elements, i.e., Monte-Carlo simulation, resulting in different outcomes for each individual simulation experiment, microsimulation allows for the exploration of the distribution of events rather than its point-estimates, thus leading to a more adequate representation of uncertainty and risk.

Together, these four shifts have a huge impact on the way in which individual lives and interactions of individuals are described and investigated. The corresponding paradigmatic shifts led to the development of the human life course as a central concept or ‘organization principle’.

The human life course

The term ‘life course’ was first used by Cain (Cain, 1964) to encompass anthropological, sociological, and psychological concepts of aging, particularly as they were related to the maturing individual's movement through an expected sequence of social roles. The life course refers to a sequence of socially defined events and roles that the individual enacts over time. It differs from the concept of the life cycle by allowing for many diverse events and roles that do not necessarily proceed in a given sequence, but constitute the sum total of a person's actual experience over time (Elder, 1975). These roles and the transitions from one role to another are central issues in family demography: childhood, partnership formation and dissolution as well as parenthood, just to name some of them. Contrary to the life-cycle concepts, which are widely used, e.g., in economics or psychology, and which are based on a predetermined ‘typical’ sequence of roles, episodes of life or expected behaviors, the life-course concept permits us to study changing role patterns and the interactions between different domains or careers, such as education, jobs, partnerships, births, but also disability. The health status can be seen as an integrated part of this career framework as it interacts with all other careers. The individual life course is determined by four key factors that constitute the key elements of the life-course paradigm:

- location;
- social integration;
- goal orientation; and
- strategic adaptation.

The location in time and place or the cultural background constitutes the first key element that determines the individual life course and closely corresponds the demographic concept of period effects as a dominant concept, especially in historical demography. In historic demography, births, deaths and marriages are reconstructed by the use of archival parish registers, and the economic and political factors that shaped the key demographic events of everyday life are determined. Key topics and insights of this kind of historical research - which concentrates on ‘ordinary people’ rather than leaders and battles - consider the changing roles and functions of families, and in particular of women. In addition, institutional changes caused by demographic changes (e.g., changes in inheritance laws) are investigated. Period factors affect a whole population at the same moment of (calendar) time. Typical health-related period effects stem from natural or nuclear catastrophes and epidemics, but also from economic transitions and crisis, which, of course, might have a different impact on different age cohorts.

The second key element is social integration or the concept of 'linked lives'. It closely corresponds to cohort effects as used in demography. Important insights were gained by comparing and identifying 'typical' life patterns of different cohorts, a method widely used in sociology. Rich empirical studies of variations in life patterns among different birth cohorts helped to elaborate the multidimensional model of the human life course. In health-care models, the concept of 'linked lives' is also of importance with regard to direct family links, as health risks (or the behaviors regarding risk factors) typically vary with living arrangements, as does the availability of informal care. In health-care finance systems, family links in the form of dependent spouses and children are also important. Health issues are also influenced by what can be attributed to cohort effects. Economic crises might affect specific age cohorts and result in specific health problems of these cohorts, but also epidemics can be expected to have a different impact on different cohorts. Health campaigns might have an impact on specific cohorts, such as preventing the current generation of young people from creating dangerous habits.

Individual age is of primary importance in demography (as a third concept beside period and cohort effects) as well as in all life-cycle models, and especially in the psychology of developmental stages. Age is also of central importance with regard to health care risks, and the change of the age distribution of the future population is one of the key issues regarding health care finance. Various scholars have tried to describe the typical life cycle that begins with birth and moves through adolescence, young adulthood, and the middle years to old age and death. By moving to a multidimensional model, the study of the life course has perceivably moved from a tendency to divide the study of development into discrete stages toward the firm recognition that any point in the life span must be viewed dynamically. This holds also true for the individual health status that is not only determined by random effects and age, but heavily depends on past behavior. Generally, the current situation and decisions of a person can be seen as the consequence of past experiences and future expectations, as integration of individual motives and external constraints. In this way, human agency and individual goal orientation are added to the explanatory framework.

The fourth component of the life-course framework was mainly brought in by longitudinal surveys and associated methods: strategic adaptation or the timing of lives. Timing of live events can be understood as both passive and active adaptation for reaching individual or collective goals. Timing is one of the most important strategies in the presence of conflicting goals, i.e., births might be postponed in order to reach other career goals first. Also, behavioral changes directly related to health – like quitting smoking – and rehabilitation/treatment might be scheduled at certain periods (e.g., holidays). The timing of retirement and when to move to a nursing home can serve as another example of a strategic adaptation to a given context under the given limitations of choices. Individuals adapt to the challenges confronting them by timing the events of their lives so as to make the most of opportunity and undergo the least frustration and failure. Whatever a person's social and cultural heritage, friendships and networks, or personal motivations are, they all come together and are experienced through the individual's adaptation to concrete situations and events (Giele and Elder, 1998). In demographic research, the life-course framework links the concepts of time, age and cohort by the fourth component of timing of lives.

The life-course paradigm moved research from single-factor explanations to multidimensional models that are flexible enough and capable of encompassing many different types of cultural, social and individual variations. While human lives can be described in various ways and terminologies, one approach has gained increasing importance and has dominated life descriptions from a live course perspective: the description of lives as event histories. An event is defined as qualitative change that occurs at a specific point in time and that places an individual in a new status. Events are transitions between states, such as marriage and divorce, that change the marital status of a person. Individuals experience events and organize their lives around these events. As Willekens (1999, p. 2) states, "most people spend a considerable part of their lives either preparing for life events or coping with life events" – falling ill being a good example for the latter.

States and events typically belong to different domains or careers, like partnership, job, educational and disability careers, that interact and influence each other. As a result, people may experience problems of synchronization and compatibility of careers. Many of the resulting problems - e.g., the reconciliation of job and family life - are central in explaining demographic phenomena. A typical strategy to cope with incompatibilities is rescheduling activities and events. An example of this strategic adaptation is to postpone births.

The collection of all possible states for each career to be considered in a specific analysis creates a state space that determines all possible trajectories and outcomes of individual life histories along with all possible transitions. Once defined, the description of individual lives consists of 'event history data', i.e., all events are recorded together with the time they occurred or, alternatively, all states are recorded by precisely noting when they began and when they ended. This approach of describing individuals is popular in dynamic microsimulation and allows to overcome the limitations of other approaches, as it allows for the inclusion of duration-dependencies in behavioral models and thereby does not restrict modeling to first order Marcov processes. This can clearly be seen when comparing dynamic microsimulation (following a state-space approach) to cell-based approaches (that put individuals into a grid of cells representing all possible combinations of states). While both approaches use a state-space approach (note that microsimulation is not restricted to this approach), no information on how individuals organized their lives before entering a cell can be recorded in the latter. Health studies are a typical field of research, where individual histories can be expected to have a huge impact on future events. While this might be true for life-course studies in general, health studies are somewhat privileged in this respect, as individual data on health and health care benefits and contributions are (at least theoretically) available in the form of administrative data, i.e., from social security carriers.

5. Microsimulation versus cell-based approaches

With regard to projection models, microsimulation and cell-based macrosimulation are often two alternative methods for making similar statements about future population characteristics (Imhoff and Post, 1998). While population projections in the narrow sense (by age, sex and some few other characteristics such as education) are almost exclusively produced by the cell-based cohort-component method, for more detailed projections, e.g., in the field of health care need and finance, the choice depends on the priorities set on the basis of a detailed evaluation of the strengths and weaknesses of both methods according to the research goal. While health care models are a typical example where both approaches can be found in parallel, there is a broad range of applications where no alternative to microsimulation exists. Good examples are tax-benefit and social security models that include detailed policies and/or require individual accounts over time. Caldwell and Morrison (2000) give the following examples:

- analysis of projected winners and losers of alternative policies on a period-specific or lifetime basis;
- analysis simultaneously focused on families and individuals;
- exploration at the micro-level of the operation of social security programs in the context of a broader tax/transfer system;
- quantification of incentives to work, to save, or to retire at particular life-course or period junctures;
- cross-subsidies across population segments or cohorts;
- feedback effects of government programs on population demographics; and
- longer-term consequences of social trends in marriage, divorce and fertility.

This chapter compares the micro- and cell-based macrosimulation approach and highlights their strengths, weaknesses and relevance in health studies. It is organized according to the following headings:

- The representation of populations;
- The modeling of population dynamics;
- Linking microsimulation and cell-based macrosimulation models;
- The areas of possible applications of the microsimulation approach;
- The strengths and weaknesses of microsimulation compared to cell-based macrosimulation.

The representation of populations

One of the first and most obvious differences between micro-and macro-models lies in the description of the population itself. In microsimulation all individuals are represented by an individual record containing all individual characteristics that might also include links to other individuals/records (e.g., to keep track of kinship networks) or any other variables. In contrast, in cell-based models population is represented by an aggregated cross-classification table, in which the cells represent all possible combinations of the characteristics considered.

A first trade-off can be found with regard to the storage space required by both methods, which is determined by the number of attributes and the population size. While this space is independent of population size in cell-based models, the number of cells – the state space consisting of all possible combinations of attribute values - "explodes" with the number of possible population attributes. In contrast, in microsimulation the number of records is determined by the population (or sample) size, and the storage space will increase only linearly with the number of variables[3] (independent of their possible values). Note that for this reason cell-based models are limited to categorical variables, making microsimulation the only practical choice when the projection model contains continuous covariates.

The importance of population size in microsimulation leads to another distinction between the approaches: microsimulation models are usually based on a population sample rather than on the total population. An exception to this is the Swedish SVERIEGE model, which is based on individual data of the whole Swedish population. The reasons why microsimulation is (usually) based on samples does not only lie in its practicability, but also in the large number of covariates that microsimulation models usually contain. The joint distribution of all state variables and covariates is generally unknown at the population level and necessary data are typically only available from sample surveys (Imhoff and Post, 1998).

Modeling of population dynamics

In cell-based macro-models, the projection model has to evaluate, for a given state-space and cell occupation, how the number that each individual cell contains changes over time. Being limited to categorical variables, dynamics can always be described by a limited set of events describing all possible changes of attribute values – or transitions from one cell to another. Given the importance of event history analysis in microsimulation models, this concept is also frequently applied to microsimulation, although microsimulation is not limited to this approach. Events are random variables that occur with a certain probability. At the population-level one can speak of the 'average' occurrence of a certain event, but this average remains to be ultimately based on the individual occurrences (Imhoff and Post, 1998). Imhoff and Post note that '…

[3] Example: For a population of size N with A attributes and Ci categories for attribute i = 1..A, the state-space would consist of C1*C2*...*CA cells in the macro-model representation while the population would be represented by a matrix of dimension N*A in the micro-model. While the population representation would be more storage-efficient in a common age-sex state-space (of typically 101*2 = 202 cells) for any population (sample) bigger than 100 persons (N*A = 100*2 = 200), this picture would change dramatically when considering more population characteristics. Consider, for example, a model that additionally includes nationality, occupation, education level, income class, parity and health status, then, even if allowing for only 6 categories each, the state space would increase to 6*6*6*6*6*6*101*2 = 9.424.512 cells. In this case, a micro-population of the same storage size could already consist of 9.428.512/8 = 1.178.064 individuals. Doubling the possible categories of only one attribute, i.e., increasing the income categories to 12, would double the whole state-space, while this would be of no effect in the case of the micro-representation.

when making a statement about a certain future number of events, we are in fact making a statement about the expected value of a random variable. In doing so, both the microsimulation and the macrosimulation approach rely upon the Law of Large Numbers. However, they do so in different ways. A macro-model assumes that the size of the population is so large that the projected number of events may be set equal to its expected value. A micro-model assumes that the number of repetitions of the random experiment in the sample is so large that the resulting projected number of events will approximately equal its expected value.'

The processes that can be simulated by cell-based models are restricted to first order Markov-processes, that is, processes without memory. The number a cell contains does not give any information of how long the individuals it represents have been in this cell and from "where" (which cell) they came.

In microsimulation models, the attribute vector is updated for each individual according to a behavioral model formulated at the individual micro-level. If needed, all past information can be stored allowing for the retrieval of the whole event history or biography of individual agents that might enter the behavioral model. This allows to include variables of duration since the previous event, which is seen as a significant source of demographic heterogeneity.

As there are no restrictions as to the variable types microsimulation models can handle, behavioral models can be of various forms. With regard to the implementation of the state-space approach in microsimulation, whether an event occurs or not for an individual is typically determined by Monte-Carlo simulation. This leads to a major difference between the modeling approaches even when modeling the same processes: dynamic microsimulation models do not only produce the expected value. As individual simulation experiments are subject to random variation, repeated simulation experiments can produce information on the distribution of target variables. As will be seen below, this is not always a “convenient" strength of the microsimulation approach.

Linking microsimulation and cell-based macrosimulation models

In various fields of projection modeling, microsimulation and cell-based macrosimulation are often seen as two alternative methods for making similar statements about the future. With both methods having their strengths and limitations, the modeler's choice is not necessarily between these two alternative methods, but it can also be a choice of how to combine these two approaches. A common practice is to align microsimulation projections to projections obtained from macro-models or scenarios (such as variants of "official" demographic projections). This approach allows to produce or reproduce given scenarios with regard to aggregated target variables while including distributional information into the projection.

Various approaches have been made in the field of linking micro-models (e.g., of a household population) to macro-models (e.g., of the economy), the German DMMS Darmstadt Micro Macro Simulator (Heike et al., 1994) being one example. In this approach, models interchange data via a defined interface (micro-macro link). This link can be of various nature, from models where the simulation results of one model feed into the other model without producing feedback reactions, to highly dynamic models, like models of an economy, where behaviors simulated at the micro-level will influence prices determined in a macro-model that will again feed back into micro-behavior.

A different approach might be of interest in cases where data availability limits possible modeling approaches to macro-models (or in cases in which microsimulation would not add anything, as transition rates are only known at the aggregated level), though some additional information can be obtained from separate microsimulation models and incorporated into the macrosimulation model. An example are the attempts to link the PSSRU (Personal Social Services Research Unit, University of Kent) cell-based macro-model and the NCCSU (Nuffield Community Care Studies Unit) microsimulation model for projections on long-term care finance in the UK (Hancock, 2002). In this approach, the means test of long-term care policies is simulated in a microsimulation model and the results are fed into the macro-model of future care demands and costs, thereby including the issue of cost incidence into the analysis.

Areas of possible application of the microsimulation approach

The areas of possible application and/or integration of the microsimulation approach can be illustrated by starting from a very simple cell-based spreadsheet model of health care demand, benefits and finance, for a given time period. This stylized model distinguishes three types of health care needs (A, B, C) that can be measured in care units with unit costs specified separately for the three types. Care demand of type A can be provided informally or formally, the other types (e.g., medication, hospitalization) can only be provided formally. The health care system is financed by a PAYG system with the contribution rate set in order to balance the system in each single period. Additionally, formal care is financed by a deductible of a given percentage of costs. To be able to determine the contribution rate from earnings in order to balance the system, the following information has to be given:

- the number of people;
- average earnings;
- average care needs per type (A-C);
- the share of informal care of type A;
- the unit costs of care per type (A-C);
- the deductible.

In order to allow accounting for demographic changes, the first four parameters are given in the form of age-specific vectors, with age-cohorts being the only distinguished cells of this simple model. The following graphic illustrates this model, with all exogenous variables being highlighted in the colored box, and the four exogenous age-vectors also being graphically represented in figures A-D.

	# Person	av. Earnings	av. Need A	% informal A	av. Need B	av. Need C	tot. Need A	tot. Formal A	tot. Need B	tot. Need C	tot. Earnings	Contribution	Deductible	tot. Benefits (without deductible)
0	50	0	20	90	10	5	1000	100	500	250	0	0	697,5	3952,5
10	60	1000	15	88	8	1	900	108	480	60	60000	-1533,21	604,8	3427,2
20	70	3000	10	86	2	1	700	98	140	70	210000	-5366,24	289,8	1642,2
30	80	4000	10	84	2	2	800	128	160	160	320000	-8177,12	388,8	2203,2
40	70	6000	12	82	3	4	840	151,2	210	280	420000	-10732,5	519,12	2941,68
50	80	5000	15	80	4	6	1200	240	320	480	400000	-10221,4	828	4692
60	70	4000	30	80	5	8	2100	420	350	560	280000	-7154,98	1134	6426
70	50	3000	60	75	10	15	3000	750	500	750	150000	-3833,03	1800	10200
80	30	2500	90	75	20	20	2700	675	600	600	75000	-1916,51	1721,25	9753,75
90	10	2000	100	70	25	25	1000	300	250	250	20000	-511,07	742,5	4207,5
							14240	2970,2	3510	3460	1935000	-49446		49446
		Unit costs A	9											
		Unit costs B	6								contribution rate:	2,555%		
		Unit costs C	3											
		deductible	15%											

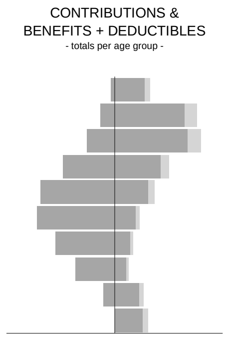

Figure G

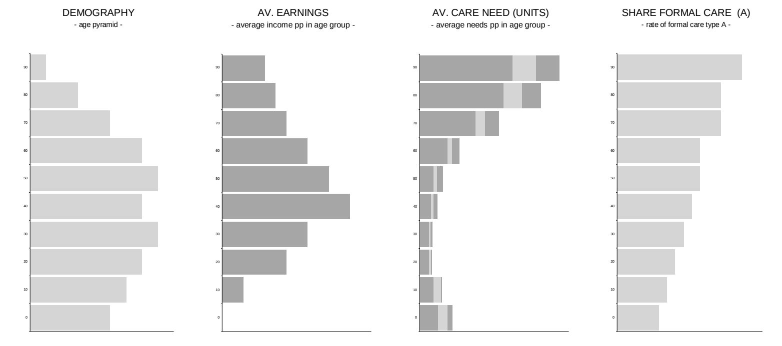

Figure A Figure B Figure C Figure D

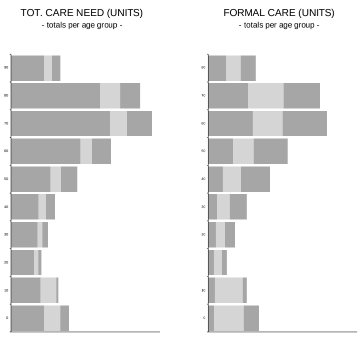

Figure E Figure F

Illustration: A cell based (spreadsheet) calculation model of health care demand and finance for a given period in a balanced PAYG system. Benefits and contributions are calculated for given population numbers, average earnings, care (unit) needs and informal care provision rates per age group as well as unit costs and a deductible rate.

In order to use this calculation model for forecasts, all exogenous information has to be provided for each period considered. For this purpose, forecasts are needed with regard to the age composition of the future population, future earnings, future care needs, the future availability of informal care supply, future unit costs and, if applicable, distributional information if deductibles are subject to a means test. In this context, models can be distinguished by:

- the extent to which they produce these forecasts by themselves or import them from other sources;
- the degree to which the interactions between demographic, health and economic processes are considered;
- the degree to which distributional aspect are included in order to allow for detailed policy simulations including progressive schemes and the inclusion of dependents into health care systems;

- the flexibility with regard to accounting, i.e., the extent contributions and benefits can be attributed to different cohorts or individuals.

Regarding population forecasts, these are usually available in the form of "official central scenario" projections of population by age and sex and most models import these projections or try to internally reproduce the same numbers in order to make the results comparable with other models. This also applies to most microsimulation models that are frequently aligned to such "official" forecasts or are parameterized in order to reproduce them internally.

The idea to model demographic and income processes (including income distributions) in one single model by microsimulation goes back to the first dynamic microsimulation model DYNASIM ('dynamic simulation of income model') developed by Orcutt (1957). While such models are able to account for detailed micro-characteristics such as educational attainment, household composition and health in the determination of human capital and labor supply, and might therefore be suitable to model income differentials, they typically only cover the supply side of the labor market. Various approaches exist in order to link micro-models of the household sector with macro-economic models, the German DMMS (Darmstadt Micro Macro Simulator, Heike et al., 1994) being one example. A similar approach was also followed by Spielauer (2001) in the context of a stylized demonstrational model, linking a dynamic micro-population with (a simplified version of) the IIASA social security accounting model.

Many models focus on future costs of health care systems and set projected costs in relation to forecasted macro-indicators, like the GDP.

A sensitive area is the modeling of future care demand. Cell-based approaches typically associate a specific average "demand mix" with population-cells. This association of a care-mix with cells creates a trade-off between two problems. If only relatively few cells are built according to "easy to project" population characteristics like age and sex, the "demand mix" associated with these cells will typically change over time and a series of assumptions has to be made in order to capture these changes. The other extreme would be to define cells by specific care needs, which moves the problem to modeling the processes that change the association of people to these cells. A typical "way out" is the definition of cells by age groups (and other variables) with care needs attached as observed today, but with future cohorts not automatically associated with cells corresponding to their age. A prominent example is the "Brookings scenario" that associates future cohorts with an assumed lower mortality to cells currently associated with younger age groups.

The approach to use current health information of people pertaining to specific groups or cells in order to project the future health care demand could be directly "translated" into a microsimulation model. Following this approach, each individual in each period would draw his health status or care need by random sampling from the equivalent age and risk group of the last period. While this approach might be appropriate when no fundamental changes in health behaviors are expected, information could be lost if behavioral changes and individual data that can be observed today are not or can not be incorporated. If the sample size allows for more detailed grouping (e.g., current health status and diseases, history information), this approach might lead to a model that distributes care needs adequately to single individuals and would therefore also allow for (meaningful) individual accounting. Such a model could constitute a logical starting point for health care modeling based on existing microdata of social security carriers. A similar approach was already used in the first dynamic microsimulation model DYNASIM (Orcutt, 1957), in which detailed health information is imputed to a simulated future cross-section sample of the US population. One of the main strengths of microsimulation lies, of course, in the fact that it allows to go beyond this approach due to the incorporation of individual-level health models, e.g., models derived by event history analysis of available micro-level health history data.

In order to forecast future health care costs, forecasts are also needed with regard to the future availability of informal care supply. Macro-approaches typically distinguish different household types (e.g., living with a partner and/or with children) in order to account for future changes regarding the population composition by household type, which is assumed to be of key importance with regard to the availability of informal care. This moves the problem to the question of how to produce household projections, if not "given" from other sources. Microsimulation is especially powerful in this kind of modeling, as it can keep track of kinship links and is therefore suitable for detailed kinship projections (also including information on relatives not living in the same household). Future informal care supply will probably also be influenced by a series of other characteristics, such as the employment status of potential providers of informal care. The availability of informal care can be expected to be sensitive especially to changes of female labor market participation. Again, microsimulation might be the appropriate modeling option as it allows to include much more detailed information. Good references regarding kinship microsimulation modeling include Wachter (1995, 1998), Wachter et al. (1998) and Tomassini and Wolf (1999a, 1999b).

Another uncertainty concerns future unit costs of health care services. While microsimulation models might be useful to model future labor supply (many health care services are especially labor intensive), prices are typically estimated exogenously or derived from steady-state assumptions regarding future growth rates.

The last parameter of the stylized model given above is the deductible. As far as such a deductible is concerned, microsimulation can be used for detailed policy analysis regarding its distributional effect and, if subject to a means test, it might be the only modeling approach that allows for the simulation of such detailed policies. An example for such an approach was already given above with the NCCSU microsimulation model (Hancock, 2002). This model is used to simulate the means test of long-term care policies, while the results are fed into a macro-model of future care demands and costs.

Figure G of the Illustration given above shows the resulting contributions and benefits (for hypothetic numbers) for a balanced PAYG heath care system. For a given period, such a system will typically redistribute resources from the "healthier" young to the older age cohorts. It might be desirable to allow for more detailed accounting over the whole life cycles of cohorts (or individuals with different characteristics and behaviors) in order to distinguish inter-and intra-personal redistributions. While this kind of accounting is essential for the study of the distributional impact of health care systems, it might also be essential for assessing the sustainability of such a system. This holds true especially if such a system is newly introduced and initial entrants differ in their contribution and benefit patterns from the population to be covered in the long run.

Strengths and limitations of microsimulation

One of the central strengths of microsimulation lies in the fact that it permits inclusion of more variables than other methods, which is especially important in projection and planning applications being of central importance in health studies, as it allows for more detailed research. For example, when trying to estimate future demand for health care facilities, etc., based on population projections, a large set of household characteristics, such as household size, family composition, age and income can be used. Microsimulation does not impose limits to variable types, allowing also continuous variables or links to other records.

Microsimulation allows for a broad range of behavioral models of any detail or complexity. This flexibility supports the study of the interaction between variables and, consequently, the life-course interactions between various parallel carriers and roles, such as education, work, partnership and parenthood within a changing socio-economic context.

Microsimulation allows the construction of behavioral models at a level on which the relevant decisions are taken, i.e., on the micro-level. There is no need to translate behavioral relations from the micro-level to the macro-level. This also implies that no information is lost through aggregation. The modeling on the micro level also allows for an assessment of the consequences of behavioral changes of specific population groups and thus for the contribution of these changes on the aggregate level.

From the view of policy-makers the main strength of microsimulation lies in its ability to test new policies in a virtual world before they are introduced into practice. In comparison to more traditional policy evaluation modeling exercises, microsimulation is especially powerful in addressing distributional issues, both in a “static” cross-sectional way and over time. The latter makes it a powerful modeling option with regard to sustainability issues of social security systems in the context of demographic change.

Based on micro-data, microsimulation allows flexible aggregation as the information may be cross-tabulated in any form, while in aggregate approaches the aggregation scheme is determined a priori. Simulation results can be displayed and accounted for simultaneously in various ways: in aggregate time series, cross-sectional joint distributions, and individual and family life paths. Flexible aggregation helps to determine “winners and losers” of policy changes by various characteristics. An example is the possibility to study and compare contribution and benefit histories over a whole individual lifespan, allowing for the calculation of individual rates of return.

Microsimulation allows to study the interaction between individuals. While modeling takes place on the individual level, simulation is used to study the resulting dynamics and patterns of change on the macro-level. In the empirical, “data-based” tradition of microsimulation, the possibility to study the interaction between individuals is mainly used to study changes in family and kinship networks. Direct applications can be found in the field of elderly care and other aspects of aging societies, where knowledge of the detailed household and family characteristics is valuable information for designing policies. The knowledge of kinship patterns additionally allows for a detailed study of intergenerational transfers and bequests, but also of hereditary diseases or the heredity transmission of specific health risks. As microsimulation allows for the modeling of interactions between individuals, it is also applied to study the transmission of diseases like AIDS, including the resulting impact on family systems. An example for such a study, projecting the familial impacts of AIDS on the elderly of Thailand, is the work of Wachter et al. (2001).

The advantages described certainly have their price, but fortunately a price that decreases over time, at least with regard to two of the most frequently listed drawbacks of microsimulation: (1) the usually large investments with respect to both manpower and hardware requirements might be considerably reduced over time as hardware prices fall and more powerful and efficient object-oriented computer languages become available; and (2) data problems are reduced over time, as more and better data, and especially longitudinal data, become available. The latter is especially true in the area of health, as social security carriers increasingly keep track of their clients' health histories, including contributions and benefits, in computer databases that might be used by researchers.

Due to the inclusion of stochastic elements - i.e., Monte-Carlo simulation - resulting in different outcomes for each individual simulation experiment, microsimulation allows for the exploration of the distribution of events rather than its point-estimates, thus leading to a more adequate representation of uncertainty and risk. As mentioned above, this is not always a “convenient" strength, as it implies that simulations have to be run various times and results have to be stored for all simulation runs in order to allow for further exploration of the distributional properties of the variables. This is burdensome and, in view of computer capacity still being one of the main bottlenecks of microsimulation, this can not always be done. As dynamic microsimulation models are of a stochastic nature, their outcome is subject to random variation. This stochastic nature of microsimulation models leads to one of its main problem areas, referred to as ‘randomness’. In microsimulation, various sources of randomness can be distinguished. These are:

- Imperfection randomness: this randomness is not specific to microsimulation but also applies to macro-models. Sources are wrong hypothesis on the values of exogenous variables as well as the fact that parameters are usually estimated from empirical data.

- Monte Carlo variability is an inherent randomness in microsimulation that does not produce the expected value, but a random variable with the expected value.

- Randomness originating from the initial population database on which the simulation is starting. Usually based on a population sample, this starting population randomness can only be reduced by increasing the sample size.

While imperfection randomness is unavoidable in all models, its scope is especially large in microsimulation and has thus become a mayor problem of microsimulation.. This is especially true for what is also called specification randomness (Pudney and Sutherland, 1994, quoted from Imhoff and Post, 1998) which is basically caused or increased by the generally large number of variables introduced in most microsimulation models. Measurement errors in the sample accumulate with an increasing number of explanatory variables. And, as a microsimulation model generates its own explanatory variables, each additional explanatory variable requires an extra set of Monte Carlo experiments, with a corresponding increase in Monte Carlo randomness. There is a trade-off between specification randomness ("too many variables") and misspecification errors (too few variables, i.e., too simple models) that leads to the fact that the degree of detail of a projection does not go hand in hand with the overall prediction power of a model. This means that what makes microsimulation especially attractive, namely the large number of variables models can include, comes at the price of specification randomness and the resulting weak prediction power decreasing with the number of variables. This generates, e.g., a trade-off between good demographic predictions and a good prediction regarding distributional issues in the long run, which, in practice, often leads to the use of alignment techniques in order to align the models' aggregate projections to external forecasts. Possible ways out of this "dilemma" are investigated in the second part of this study that, based on a review of existing microsimulation models, presents some lessons for health-related modeling.

Summary and conclusions

This paper constitutes the first part of an investigation of the potential of dynamic microsimulation for modeling and projecting health care demand, health care finance and the economic impact of health behavior. The main purpose of this paper was to provide a brief theoretical background regarding the dynamic microsimulation approach and a comparison of the microsimulation approach with the cell-based macro approach. Starting with a definition of dynamic microsimulation and a classification of the types and approaches, microsimulation modeling was brought into the context of the life-course paradigm. This paradigm, meanwhile being the dominant paradigm in demography, is seen as a useful organizational principle for the study and projection of health-related phenomena using microsimulation. Microsimulation was then compared with cell-based approaches, and the potential strengths as well as drawbacks of the microsimulation approach regarding health care modeling were investigated.

This study shows that microsimulation might turn out to be increasingly appropriate as a modeling approach in a field that is currently dominated by cell-based macro-models. Microsimulation can be used in a wide area of applications ranging from very specialized models that might be linked to macro-models to integrated microsimulation models. The area of possible applications of the microsimulation approach was discussed starting from a simple cell-based macro model. The advantages of the microsimulation approach certainly have their price. There is a trade-off between specification randomness ("too many variables") and misspecification errors (too few variables, i.e., too simple models). or between the prediction power and the detail of the models. This has led to very different ways of how existing microsimulation models have been designed and possibly linked with or aligned to other (macro) models. An investigation of existing microsimulation models is carried out in the second part of this study, that, based on this review, presents some lessons for health-related modeling.

References

Benenson I. and I. Omer (2001), Agent-Based Modeling of Residential Distribution. Department of Geography and Human Environment, University Tel Aviv.

Burch, T.K. (1999), Computer Modeling of Theory: Explanation for the 21st Century. London, Ontario: Population Studies Centre, University of Western Ontario.

Cain, L.D. (1964), Life course and social structure. Handbook of Modern Sociology. Faris, Chicago, Rand McNally: 272-309.

Caldwell and Morrison (2000), Validation of longitudinal dynamic microsimulation models: experience with CORSIM and DYNACAN. In: Microsimulation Modeling for Policy Analysis, Lavinia Mitton, Holly Sutherland and Melvyn Weeks (eds.), Cambridge University Press.

Elder, G.H. (1975), Age differentiation and life course. Annual Review of Sociology, 1: 165-190.

Falkingham, J. and A. Harding (1996), Poverty Alleviation Versus Social Insurance Systems: A Comparison of Lifetime Redistribution. NATSEM discussion paper 12-1996.

Galler, H.P. (1997), Discrete-Time and Continuous-Time Approaches to Dynamic Microsimulation Reconsidered. Technical Paper 13. National Centre for Social and Economic Modeling (NATSEM), University of Canberra.

Giele, J.Z. and G.H. Elder, Jr. (1998), Methods of Life Course Research: Qualitative and Quantitative Approaches. SAGE Publications.

Gilbert, N. and K. Troitzsch (1999), Simulation for the Social Scientist. Open University Press.

Hancock, R., A. Comas-Herrera, R. Wittenberg and R. Pickard (2002), Who will pay for long-term care in the UK? Projections linking macro- and micro-simulation models. Paper prepared for a meeting of the International Network of Research on Elder Care, Hannover, 8-9 February, 2002.

Heike, H.-D., K. Beckmann, A. Kaufmann and T. Sauerbier (1994), Der Darmstädter Mikro-Makro-Simulator - Modellierung, Software Architektur und Optimierung. In: Faulbaum, Frank (ed.): SoftStat'93 - Advances in Statistical Software 4, Fischer, Stuttgart, New York, pp. 161-169.

Imhoff, E.V. and W. Post (1998), Microsimulation Methods for Population Projection. New Methodological in the Social Sciences, pp. 97-138.

Klevmarken, N.A. (1997), Behavioral Modeling in Micro Simulation Models. A Survey. Working Paper 1997:31, Department of Economics, Uppsala University.

Lane, D.C. (1999), Social theory and system dynamic practise. European Journal of Operational Research, 113, 501-527.

Lutz, W., J.W. Vaupel and D.A. Ahlburg, Eds. (1999), Frontiers of Population Forecasting. Population and Development Review, New York: Population Council, 24 (Supplement).

Martini, A. and U. Trivellato (1997), The Role of Survey Data in Microsimulation Models for Social Policy Analysis. Labour, 11(1), pp. 83-112.

O'Donoghue, C. (2001), Redistribution in the Irish Tax-Benefit System. PhD thesis at the London School of Economics.

O'Donoghue, C. and H. Sutherland (1999), Accounting for the family in European income tax systems. Cambridge Journal of Economics, 23: 565-598.

Orcutt, G. (1957), A new type of socio-economic system. Review of Economics and Statistics, 58:773-797

Pechman, J., B. Okner (1974), Who Bears the Tax Burdon? Washington DC: The Brookings Institution.

Pryor, R. J., N. Basu and T. Quint (1996), Development of Aspen: A Microanalytic Simulation Model of the U.S. Economy. SAND96-0434, Sandia National Laboratories.

Spielauer, M. (2000), Microsimulation of Life Course Interactions between Education, Work, Partnership Forms and Children in Five European Countries. IIASA Interim Report IR-00-032, International Institute for Applied Systems Analysis, Laxenburg, Austria.

Spielauer, M. (2001), Microsimulation Modeling of Population, Economic Growth, and Social Security Systems. IIASA Interim Report IR-01-026, International Institute for Applied Systems Analysis, Laxenburg, Austria.

Sutherland H. (1995a), Static Microsimulation Models in Europe: A Survey. Microsimulation Unit Discussion Paper, MU9503, University of Cambridge, Department of Applied Economics.

Sutherland, H. (1995b), The Development of Tax-Benefit Models: A View from the UK. University of Cambridge. Department of Applied Economics. Dae Working Paper No. 9519, pp. 1-19.

Toder, E. et.al. (2000), Long Term Model Development for Social Security Policy Analysis. The Urban Institute.

Tomassini C. and D. Wolf (1999a), Shrinking Kin Networks in Italy Due to Sustained Low Fertility. European Journal of Population, forthcoming (Center for Policy Research Papers in Microsimulation Series Paper No. 5, December 1999).

Tomassini C. and D. Wolf. (1999b), Stability and Change in the Living Arrangements of Older Italian Women: 1990-1995. Genus, forthcoming (Center for Policy Research Aging Studies Program Paper No. 19, April 1999).

Troitzsch, Klaus G. (1996), Social Science Microsimulation. Berlin, New York: Springer.

Vencatasawmy, C. (2002), Modeling fertility in a life course context: some issues. Austrian Institute for Family Studies, Working Paper 17.

Wachter, K.W. (1995), 2030's Seniors: Kin and Step-Kin. Berkeley: University of California.

Wachter, K.W (1998), Kinship Resources for the Elderly: An Update. Berkeley: University of California.

Wachter, K., D. Blackwell and E.A. Hammel (1998), Testing the Validity of Kinship Microsimulation: An Update. Berkeley: University of California, February.

Wachter, K., J. Knodel and M. VanLandingham (2001), Parental Bereavement: Heterogeneous Impacts of AIDS in Thailand. PSC Research Report, Report No. 01-493, December 2001; Population Studies Center at the Institute for Social Research, University of Michigan.

Willekens, F. (1999), Theoretical and Technical Orientation toward Longitudinal Research in the Social Sciences. Population Research Centre, University of Groningen.