New Zealand Population Review, 31(1):77-124.

Biographic Forecasting: Bridging the Micro-Macro Gap in Population Forecasting

FRANS WILLEKENS*

Abstract

The paper outlines a new model for demographic projections by detailed population categories that are required in the development of sustainable (elderly) health care systems and pension systems.

The methodology consists of a macro-model (MAC) that models demographic changes at the population level and a micro-model (MIC) that models demographic events at the individual level. Both models are multistate models that rely on rates of transition between states of existence or stages of life. MAC focuses on transitions among functional states by age and sex. The transitions determine the distribution of cohort members among functional states. The output of MAC consists of ***cohort biographies***. MIC addresses demographic events and other life transitions at the individual level. It is a micro-simulation model that produces ***individual biographies***. This paper describes approaches to functional population projection and provides a detailed description of the multistate model. It also contains an overview of the MicMac project.

Demographic projections are usually confined to populations disaggregated by age, sex and sometimes race/ethnicity. The general methodology, the cohort-component method, is well-established. The basic approach is to distinguish birth cohorts, to determine the number of survivors in the base year and to determine *for each cohort* and for each future year the number of persons by age and sex that (1) enter a population through birth or migration and (2) leave a population due to death or emigration. The number of entries and exits are based on *rates* of birth, death and emigration by age and sex, and number of immigrants by age and sex. The estimation of empirical rates from data (often incomplete

---

* Netherlands Interdisciplinary Demographic Institute (NIDI), The Hague and Population Research Centre, University of Groningen. Email: willekens@nidi.nl

or defective data) and the prediction of future rates involve important methodological issues.

The projected population by age and sex serves as an input in functional population projections, which are related to particular functions or activities in society. They include projections of the population by functional status such as labour force status, educational status, health status and status in the household. These projections are made to determine the projected need for some "function" – a product, a service, an allowance, an activity or a facility (Kono 1993). Examples include:

- The future size of the labour force to determine the supply of labour and the demand for jobs.
- The future size of the population retired from the labour force to determine the demand for pensions.
- The future size of school enrolments to determine the demand for teachers and buildings and to determine the population composition by level of education and hence the human capital.
- The future size of the population by health status and/or disability status to determine the demand for health care including the number of physicians and hospital beds.
- The future number of households by size and type to determine the demand for housing and durable consumer goods.
- The future number of people eligible for assistance of different type. Eligibility criteria frequently include age, sex and functional status (eg. level of income, health status).
- The future size of vulnerable groups in society.

Despite the wide variety of functions, from a methodological viewpoint traditional functional projections differ from each other only in minor detail. Traditionally, people of a given age and sex are allocated to functional states using a set of prevalence rates, a distribution function or another allocation mechanism. The method is referred to as distribution method. In the forecasting literature the method is also known as *static* as opposed to methods that are based on transition rates (or incidence rates) and that are *dynamic* (such as the multistate model) (eg. Zeng Yi *et al.* 1997). Examples of the static method based on distributions are the headship rate method for household projections, labour force projections based on labour force participation rates, educational projections based on enrolment rates, the ratio method of sub national population projection, and health status projections based on prevalence rates ('Sullivan method'). The distribution function may change over time to capture real or assumed shifts in behaviour or conditions. For illustrations of the static method in functional population projections, see Bogue *et al.* (1993:Chapter 18).

In the dynamic method, the distribution of people among functional states is not imposed by a distribution function but is the outcome of transitions people make in life. People move between functional states and as a consequence, the structure of the population changes. The rates of transition determine the population dynamics and the rates may change in time and may vary between subpopulations. In the dynamic method, several states of existence are distinguished and the transitions between the states are considered. The method is known as the multistate method. Because of the pivotal role of transitions, multistate models picture more closely the *mechanism of demographic change* taking place in the real world. As a result, they are better suited for integrated population projections in which functional states and interactions between functional states play a crucial role. In addition, the transitions provide a way to assess the impact on population dynamics of behavioural changes brought about by technological, economic or cultural change, or policies. The multistate method has become the standard methodology among demographers (Rogers 1975, 1995; Willekens and Drewe 1984; Keyfitz 1985; Ahlburg *et al.* 1999; Van Imhoff and Keilman 1991; Zeng Yi 1991). It has been applied for projecting regional populations, and projections by educational status, household status, labour force participation and health/disability status. The multistate model is currently receiving much interest in epidemiology and public health (for a review, see Commenges 1999; Hougaard 1999, 2000). A major difference between multistate models in demography and epidemiology is that in demography age is a key variable, whereas in epidemiology it is not (yet).

The choice of static versus dynamic method has been the subject of long debates in demographic analysis and forecasting. In labour force projections, the debate was most intensive in the early 1980s after the publication in 1982 by the Bureau of Labor Statistics of multistate tables of working life (Smith 1982). In health status projections, the debate is of a more recent date (eg. Crimmins *et al.* 1994; Mathers and Robine 1997). Some authors attempted to reconcile the two approaches (eg. Newman 1988). In his review of functional population projections, Kono asserts that "Because of complex and precise data demands, however, almost no multistate models which could be used reliably in official national projections, beyond regional projections, have been developed" (Kono 1993:18.2). In this paper, we aim to show that multistate models represent an adequate basis for the specification of functional population projection models despite data demands.

Forecasting involves dealing with uncertainty since the future is inherently uncertain. Two basic issues arise. The first is to *quantify* the uncertainty, ie. to indicate the degree of precision of the projected figures. The precision is high when a predicted figure is likely to be true. The second is to *reduce* the uncertainty, i.e. to increase the predictive performance of forecasts. The quantification of uncertainty has received much attention in the scholarly literature. The traditional approach is to specify a few sets of vital rates and migration rates that represent possible futures (scenarios). A more recent approach is to generate probabilistic projections that are based on the assumption that point forecasts are available for the relevant vital rates and migration rates, and the expected uncertainty of the forecast can be characterized in terms of variances and certain simple covariance structures for the error terms. The outcome is a predictive interval that specifies the probability that the future population will be between x and y million (eg. Alho and Spencer 1997; Lee 1998; Alho 2003). In order to determine the nature of the distribution characterizing vital rates and the width of the distribution, three alternative approaches have been proposed in the literature. One approach is based on statistical time series analysis, the second uses an extrapolation of errors observed in past projections, while the third derives uncertainty bands from expert judgement. A synthesizing approach that includes the key elements of all three approaches has been outlined by Lutz *et al.* (1997).

The reduction of uncertainty has been studied less systematically in the scholarly literature although it is a core issue in the production of more accurate and reliable forecasts. Strategies for reducing uncertainty include (1) a better understanding of the mechanisms that govern demographic change and (2) a better measurement of vital rates (demographic parameters) for subpopulations. The second strategy, which involves better data, received extensive coverage in the literature. The first strategy was much less the subject of systematic investigation. It involves a better use of scientific knowledge in demographic projections. Twenty years ago, Keyfitz asked the question "Can knowledge improve forecasts?". He stated that "For policy purposes, causal knowledge is essential; for forecasting it is desirable, of course, but the forecast is not necessarily a failure if the causal mechanisms remains undiscovered. Observed regularities serve perfectly well for forecasting as long as they continue to hold." (Keyfitz 1982:747). For many years, the search for regularities dominated the demographic forecasting literature. There is nothing wrong with that. As Keyfitz observed "Pending the discovery of a truly behavioral way of estimating the future, we cannot afford to be ashamed of extrapolating the observed regularities of the past" (Keyfitz 1982:747). About ten years ago and in the context of forecasting the health of the elderly population, Manton, Singer and Suzman are less at ease when they summarized the state of the art as follows:

> Current forecasting procedures are often based on empirical extrapolations and do not directly reflect physiological processes at the individual level or the mixture of individuals in a cohort. The failure to deal with individual aging trajectories, and their cohort mix, makes it difficult to use epidemiological and biomedical evidence on the impact of health changes on the organism in forecasts."(Manton *et al.* 1993:25).

The effective use of substantive knowledge on causal mechanisms remains a challenge. Most demographic projection models have limited scope for incorporating substantive knowledge on causal mechanisms. Multistate models have that scope.

Over the years, researchers tried to improve the *predictive performance* of the models they developed by incorporating substantive knowledge on biological and behavioural mechanisms underlying demographic change. When Keyfitz published his views, no generally accepted framework existed that encompassed the many factors affecting demographic processes in a dynamic way and allowed a causal analysis. Today that framework exists. During the past decades, the life course paradigm has become a "metatheoretical perspective" that integrates biological processes, past experiences (antecendents) and historical context (Giele and Elder 1998:21). It provides a way to combine biological processes, cognitive processes and social processes that shape the lives and behaviours of people. It also provides a logical framework to integrate risk factors (particular attributes) and exposure analysis (duration at risk of particular events and risk levels). In the behavioural and social sciences and in epidemiology, biological and behavioural mechanisms are increasingly being studied from a life-course perspective (eg. Giele and Elder 1998; Elder 1999; Kuh and Ben-Shlomo 1997; Barker 1998; Ben-Shlomo and Kuh 2002; Kuh and Hardy 2002; WHO, 2002; Halfon and Hochstein 2002). The concept of a life course refers to the way in which the countless aspects of our lives are interwoven and shaped by biological, technological, cultural and institutional influences and how their interaction results in an organic system that evolves in time. The factors affecting our lives include personal characteristics, individual histories, contextual factors and collective histories. Since the life course is embedded in a historical context, the effects of these factors are revealed more clearly if different cohorts (or generations) are considered. The life course paradigm continues to be a successful organizational principle for research. It has been proposed as a paradigm for demographic forecasting (Willekens 2002). It is gradually becoming a paradigm in policy making in the private and public sector. It is the dominant framework that underlies life planning, a subject that is receiving a growing interest as the population ages and the role of government in social support is being debated. Governments and financial institutions are providing tools for assessing lifetime financial and other consequences of life events such as marriage, divorce, childbirth, and retirement. New government policies are introduced that adopt a life course perspective (Rowe 2003). For example, health policies are increasingly targeting risk factors that affect health in later life (eg. smoking and obesity), and the provision of pensions is increasingly being discussed within the context of life planning extending over the entire life course.

This paper defends three complementary views:

1. *Functional population projections are essentially projections of cohort biographies.* Functional population projections pertain to different domains of life including work, family, health and residence. The aim of functional population projections is to project how many members of a real or synthetic cohort occupy the functional states at a given age and a given point in time. In other words, the aim of functional population projections is to project ***state occupancies***. The sequence of state occupancies by cohort members as they age describes a ***cohort biography.***[1] Consequently, functional population projections are essentially projections of cohort biographies (biographic projections).

2. *Whenever possible, functional population projections should utilize multistate models since they picture more closely the mechanisms of demographic change.* Static projection models rely on observed state occupancies (prevalence rates, ratio method). However, the state occupancies are the outcome of ***state transitions*** at earlier ages. The dynamic method derives state occupancies from transitions between functional states. The transitions are governed by transition rates (or incidence rates) and transition probabilities.
3. *The life course offers a logical framework for incorporating substantive knowledge in forecasts.* Knowledge on biological and behavioural mechanisms can indeed improve forecasts and the life course is the way to incorporate substantive knowledge. Techniques of event history or life history modeling permit causal analysis (Blossfeld and Rohwer 2002). They can be extended to forecasting.
4. *Functional population projections should evolve to projections of individual biographies.* Cohort members differ in personal attributes and living conditions (context). The best approach to account for these differences is to distinguish individuals and to characterize each individual by a bundle of attributes. These *virtual* or *synthetic* individuals bridge the micro-macro gap in population forecasting. The life courses of these individuals may be projected using techniques of micro-simulation. The aggregation of the individual biographies that result yields a bottom-up estimate of the cohort biography.

This paper suggests a shift from conventional population projection, with its emphasis on numbers of people, to a projection of the lives of people. It links macro-level models of population dynamics with micro-level models of the individual life course. The life course is viewed as a sequence of *states* (or stages) and *events* that involve transitions from one state to another. The advantages of such a shift are the following.

i. Population heterogeneity can be dealt with in a better way than when other approaches are adopted. Traditional projections assume that members of the same cohort have identical demographic behaviour. Within-cohort differences are introduced by stratifying the population into subpopulations on the basis of significant attributes such as sex, marital status, health status, and region of residence. Membership of a subpopulation is usually not fixed forever. During the life course, people move between subpopulations. They marry and divorce, change health status and migrate. Members of the same stratum or subpopulation are not homogeneous either; they may differ in many ways and the differences are likely to affect their chances for survival, the number of children they have, and other aspects of demographic behaviour. An investigation and representation of these differences at the population level becomes infeasible quickly. An approach that focuses on individual actors, their lifestyles and life courses, facilitates the implementation of heterogeneity.

ii. Population dynamics is the outcome of changes in the relative size of subpopulations (composition effect) and changes in the behaviour of members of a subpopulation (rate effect). Population forecasting involves the prediction of or hypotheses about changes at the individual level. Since demographic events are embedded in the life course, these predictions are difficult to make without a life course perspective. For instance, it has been stressed that mortality projections should use the growing insights in the physiological mechanisms underlying aging and their relation to mortality (Manton 1993:79). Considerable progress has been made. Yashin (2001) reviews mortality models that incorporate theories of aging. What applies to mortality, applies to fertility, migration, marital status change, and other demographic events.

iii. Life course projections provide information not available in regular functional population projections. Functional projections provide information on state occupancies and types and numbers of state transitions at some future time. Life course projections also provide sequences of states occupied (pathways) and estimates of expected sojourn times in the different states or episodes of life. The episodes may relate to unemployment, disease, dependency, or poverty. For instance, population projections generally include estimates of dependency ratios, but the expected duration of dependency remains unknown. Life course projections provide information on expected sojourn times in dependency. For instance, unemployment projections would be more relevant to policy making if durations of unemployment spells would be predicted in addition to the proportion unemployed.

iv. The projection of the lives of people has significance in its own right, independent of its contribution to improved population forecasts. The prediction of the probability of an event in a given period or a lifetime

> has significance beyond its meaning for changes at the population level. Examples include the prediction of the impact of risk factors on the incidence of a chronic disease, the likelihood of deviant or criminal behaviour, the probability that a marriage ends in a divorce, the probability of entry into relative poverty, etc.. The number of papers on these subjects is immense. There exists however a common modelling approach that is often used. It is the prediction of probabilities based on hazard functions estimated conditionally on risk factors and other covariates that affect the rate of occurrence, sometimes augmented by unobserved random effects. The incidence is often linked to events and experiences early in life (including foetal life and infancy). The term "programming" has been used to describe a process whereby a stimulus at a sensitive or critical period of development has lasting or lifelong significance (eg. Barker 1998:13). One of the best examples of the programming phenomenon is the permanent change that is induced by under-nutrition in early life.

The paper is organised as follows. Section 2 presents the approach that is adopted in biographic forecasting. Life is viewed as a sequence of states and events. They may pertain to one particular domain of life or to a combination of domains such as work and family. A sequence in one domain is referred to as a career. Life consists therefore of a set of interdependent careers (Elder 1985).

Section 3 presents the multistate model. The projection model is an extension of the cohort-component model and the model for functional population projections. The basic parameters are **transition rates** and **transition probabilities**. These rates and probabilities must be estimated from the data and consequently the model rates/probabilities are equal to empirical rates/probabilities. The core of the method is the multistate model and regression models that predict rates (or probabilities) of transition. In this paper, no distinction is made between the multistate life table model and the projection model. We make use of two perspectives on the life table. The first is a population perspective: the life table describes the characteristics of a stationary population. The second is a life history perspective: the life table describes the life history of members of a synthetic cohort, i.e. the cohort biography. In traditional projections, the first perspective dominates. The life table serves as the source of the parameters of the projection model. In this paper, the life history perspective is followed. Although that perspective has been around for decades and the multistate life table has been used to describe life histories (eg. Willekens and Rogers 1978), it has not caught on.

Section 4 discusses a generic approach to accommodate substantive knowledge and causal mechanisms in population forecasting. The approach relies on transition rate models and transition probability models and follows the perspective on causal analysis adopted by Blossfeld and Rohwer (2002).

The empirical base for demographic forecasting consists of data of various types. Two broad categories are distinguished: data on events and data on discrete-time transitions. The multistate life table contains methods for estimating transition probabilities from transition rates. Section 5 describes a method for estimating transition rates from probabilities. It is the inverse method, developed by Singer and Spilerman (1979).

The central position of transition rates in demographic analysis is also illustrated in Section 6. It is asserted that the transition rates are the logical parameters to integrate scenario-setting and various types of uncertainty in demographic forecasting. The view is held that probabilistic forecasting should concentrate on quantifying the uncertainty in transition rates and transition probabilities. The transition rates also constitute a bridge between the population level and the individual level. At the individual level of analysis, e.g. to predict individual biographies, use is made of transition rates that depend on individual attributes. At the population level, e.g. to forecast cohort biographies, expected values of transition rates across individuals are used. The micro-macro link is further described in Section 7.

Section 8 presents a brief description of the MicMac project funded by the European Commission and implemented by a consortium of research centres in Europe. An introduction to the software is also given. Section 9 concludes the paper.

The Approach: Forecasting Biographies/Life Histories

Generic models of the life course view an individual as a carrier of attributes and the life course as sequences of events and states/episodes. Sequences are defined in different domains of life and they co-exist, co-evolve and interact. A population consists of individuals and the population structure is the aggregate effect of individual life courses. For the purpose of analysis and projection, a population is stratified in birth cohorts. A cohort may be further stratified by functional state or state of existence. A cohort evolves because members make transitions between functional states and eventually die. Existing cohorts are replaced by new birth cohorts. Functional population projections are projections of state occupancies.

Multistate models describe the life course. Age is the main time scale used to position life events in time. In its simple form, the multistate model describes the collective life history or biography of a cohort and disregards intra-cohort variations. The description of the cohort biography is facilitated by the multistate life table and extensions of the life table. Multistate models have been used for population projection. Examples include LIPRO (Van Imhoff and Keilman 1991), MUDEA (Willekens and Drewe 1984), and PROFAMY (Zeng Yi *et al.* 1997). The models are designed to describe and project changes in the population composition at the macro level. They concentrate on the position individuals occupy in the collective biography at consecutive points in time and the population structure that results. They do not address the prognosis of events and episodes in the lives of people, which is the subject of life course projections. The link between the traditional macro-level models and the new micro-level models is the multistate life table, and more particularly the **occurrence-exposure rate**. Occurrence-exposure rates are also known as event rates, hazard rates and transition rates, depending on the field of study. The occurrence-exposure rate bridges the micro-macro gap in population forecasting. The occurrence-exposure rate may be used to describe the dynamics at the macro-level and the transitions at the micro-level. Observed differences between cohort members are considered in terms of risk factors and covariates and the risk ratios or relative risks associated with different levels of risk factors and covariates. Unobserved differences are described by mixture models and random effect models. Mixture models classify people in a finite number of categories. Random effect models assume a continuous distribution of individual differences.

The life table underlies both the dynamics at the population level and the individual life history. That is consistent with the population and the life history perspectives on the life table (see above). The multistate life table and the multistate projection model are adequately documented in the literature (eg. Roger, 1975; Schoen 1988; Manton and Stallard 1988). In this paper, no fundamental distinction is made between the life table and the projection model. The life table is viewed as a projection model.

The combination of a hazard model and a multistate life table constitutes the main ingredient of the proposed method for functional population projections and the prediction of the life course. The combination of regression models and life tables was introduced more than thirty years ago by Cox (1972) and, for multistate life tables, more than 10 years ago by Gill (1992). Both authors give a central position to occurrence-exposure rates or transition rates. Many models in demography and epidemiology rely on probabilities or on types of rates that differ from occurrence-exposure rates. The method that allocates a key position to occurrence-exposure rates is sometimes referred to as the person-years approach.

By way of illustration of biographic forecasting, we consider chronic diseases. The prognosis of a chronic disease involves the prognosis of the occurrence of the disease, the age at onset of the disease, and the number of years with the disease. It also involves the identification of the factors that increase or reduce likelihood of the disease. Risk factors are among them, but also the factors that influence the length of life. The prognosis of a disease cannot be separated from a prognosis of the length of life. Figures on lifetime risks that are often cited in the media or scholarly literature involve statements or assumptions on length of life. To clarify the interdependence, the occurrence-exposure is the key. To demonstrate the significance of this statement, which may be common knowledge in demography, a reference is made to recent discussions in the medical literature. Many studies that estimate the probability of an event during a period of a given length or the lifetime risk of an event overestimate the probability by an inadequate treatment of the competing risks. The matter is at the core of life-course forecasting. We consider a few examples.

The probability that a person in the Netherlands develops cancer is 45.1 per cent for males and 30.4 per cent for females, if the person survives to the age of 85. The probabilities are conditional on survival to the age of 85. If mortality before the age of 85 is taken into account the probabilities of developing cancer before the age of 85 is 33.2 per cent for males and 27.8 per cent for females (Schouten *et al.* 1994). The first type of probability is known as the cumulative incidence (CI), the second is the life table probability. As people live longer, the probability of developing cancer will increase, even when the age-specific cancer rates do not change. It is the composition effect, referred to in the introduction of this paper. More people survive to ages when the risk is high. Schouten *et al.* also estimate that the probability that women in the Netherlands develop breast cancer before the age of 85 is 9.4 per cent in the absence of mortality before 85 and 7.9 per cent in the presence of mortality. An often-cited figure in the United States is a probability of breast cancer of 12.8 per cent. The estimate is from the National Cancer Institute (NCI) and is the *lifetime* probability of breast cancer (e.g. National Cancer Institute 2001; Morris *et al.* 2001). It is based on the NCI's Surveillance Program (SEER) and cancer rates from 1995 through 1997, and it takes into account that not all women live to older ages, when breast cancer risk becomes the greatest (National Cancer Institute 2001). Cancers that develop at a higher age are more prevalent among population groups that live longer. For instance, the lifetime risk of prostate cancer is higher among Whites than Blacks because fewer Blacks reach the ages where prostate cancer develops rapidly (Wun *et al.* 1998:183). The lifetime risk of parkinsonism and Parkinson's disease is only slightly higher in men than in women because of the opposite effects of higher incidence and higher mortality in men (Elbaz *et al.* 2002). It is well-known that healthy life may increase the lifetime probability of chronic diseases that start at higher ages, such as cardiovascular disease and cancer. The greater longevity of women is the primary cause of their greater lifetime probabilities of congestive heart failure and stroke (Peeters *et al.* 2002). And when the disease occurs, women lose a greater number of years of life than men. A final illustration highlights the complex interrelation between smoking, cardiovascular disease and mortality. Although smoking is known to increase the risk of cardiovascular disease at each age, over a lifetime never-smokers have approximately the same risk of cardiovascular disease as always smokers, simply because they live longer (Mamun *et al.* 2002). Furthermore, if fewer smokers die from lung cancer, the lifetime risk of heart disease among former smokers may rise. It is the association between the risk factors and different chronic diseases that underlie the tempo distortions of mortality, identified by Bongaarts and Feeney (2002) and discussed by Vaupel (2002). It links the tempo distortions to heterogeneity and selection.

The cumulative incidence or cumulative risk is often used to determine the likelihood of a disease. It is the number of new cases during a given period (5 years, 10 years, or lifetime) divided by the initial population free of a disease (*event-free population*). The CI is generally not adjusted for the presence of competing risks, such as death. It is therefore free of the influence of mortality.[2] Formulated differently, it assumes that there is no competing risk of death. In the absence of death or another competing event, the amount of time at risk is the same for every member of the group and members are assumed to live for the entire lifespan. It is a conditional probability; conditional on survival. As a consequence, the CI or cumulative risk overestimates the risk of developing a disease. The effect of the competing risks is taken into account by constructing a multiple decrement life table. The life table risk or lifetime risk, as it is often called in epidemiology, is lower than the CI. For a discussion, see Schouten *et al.* (1994), Lloyd-Jones *et al.* (1999), Beiser *et al.* (2000:1499) and Elbaz *et al.* (2002). Beiser *et al.* (2000) distinguish an unadjusted cumulative incidence (UCI) and an adjusted cumulative incidence (ACI). The UCI overestimates the incidence of an event because not all people live till the maximum age. The ACI proposed by the authors account for the competing risk of death by calculating the CI using the multiple-decrement life table. In the absence of a mortality pattern of the study population, the authors suggest using the mortality experience of a "standard" population, leading to a standardised lifetime risk. This brief discussion of cumulative risks and lifetime risks or lifetable probabilities illustrate the importance of competing events in the estimation of probabilities. The use of occurrence-exposure rates as defined above, is a guarantee for the correct estimation of probabilities. In the field of epidemiology, the person-years analysis of incidence rates has been described by Breslow and Day (1987). For a detailed recent description in the context of the prediction of lifetime incidence, see Beiser *et al.* (2000). Occurrence-exposure rates are used throughout this paper.

The illustration demonstrates the types of issues that arise in biographic forecasting. Prediction or prognosis of events and experiences, e.g. episodes of poverty, unemployment, or impairment, is not common yet in demography, but is common in epidemiology, public health, and medical practice. In those fields, it is still considered "a complicated business" (De Backer and de Bacquer 1999). The probabilities that are estimated from data and projected in the future are used by physicians to determine the need for intervention or treatment. The probabilities are also used in public health to determine public health concerns and to assess the public health and financial consequences of the presence of risk factors.

Transition rates depend on risk factors and other determinants. A risk factor is defined as a factor that is causally related to an outcome. The concept originated in epidemiology, where the identification of the causal link is an important element of the etiology of a disease. In many cases, a causal link cannot be determined and the association between predictor and outcome is a statistical one. In the prediction of the life course, risk factors and other factors are evaluated in terms of their predictive performance and not their explanatory power. Two comments are warranted here. First, the link between a risk factor and the outcome is probabilistic. It means that the presence of a risk factor changes (usually increases) the *probability* of an event or the *expected* duration of an episode. The significance of an event lies in the consequences to the life history of an individual (Peeters *et al.* 2002). Second, several risk factors may change during the course of life. Modifiable risk factors are particularly relevant in the design of health policies and public health programmes. They should also be considered in forecasting since the health outcomes (and mortality) depends on the modifiable risk factors. For instance, when more people stop smoking and start eating healthy, the long-term consequences will be increased survival, possibly associated with longer periods of chronic disease.

The Multistate Model

In this section, we derive the multistate model for a sample of m individuals born at the same time (same year, say). We adopt a probabilistic perspective which has been introduced in multistate demographic modelling by Hoem and Jensen (1982), Namboodiri and Suchindran (1987) and others (eg. Chiang 1984).

State Occupancies

Let $Y_k(x)$ be a time-varying indicator variable representing the state occupied by individual k ($k = \{1, 2, \ldots, m\}$) at age x. Individual k is not necessarily a specific person but a combination of attributes. Instead of age, we may use another time scale. In that case, x indicates the time elapsed since the reference event. The possible states are given by the state space **S** $=\{1, 2, 3, \ldots, I\}$, with I the size of the state space. The state space includes all possible states. If death is considered, it includes the state of dead. Dead is an absorbing state and cohort members who die remain in that state. The polytomous random variable $Y_k(x)$ is a discrete variable that can take on as many non-zero values as there are states in the state space. $Y_k(x)$ is zero if individual k died before age x.

The number of individuals in state i (i = 1, 2, …, I) at age x is denoted by $N_i(x)$. It is equal to

$$N_i(x) = \sum_{k=1}^{m} I_{Y_k(x)} = \sum_{k|Y_k(x)=i}^{m} \mathrm{I}$$

where m is the number of individuals in the birth cohort and $I_{Y_k(x)}$ is an indicator function which is 1 if $Y_k(x)$ is i and 0 otherwise. $N_i(x)$ is a random variable.

A second approach exists to denote the state occupied at a given age. It defines a binary random variable $Y_{ki}(x)$. It is equal to 1 if individual k is in state i at age x and 0 otherwise. The number of individuals in state i at x is

$$N_i(x) = \sum_{k=1}^{m} Y_{ki}(x)$$

The expected value of $Y_{ki}(x)$ is the probability that individual k is in state i at age x. It is the *state probability*. Two types of state probabilities are distinguished: unconditional and conditional. The unconditional state probability is the probability that cohort member k occupies state i at age x; it is denoted by ${}_k\ell_i(x)$ and

${}_k\ell_i(x) = E[Y_{ki}(x)] = \Pr\{Y_{ki}(x)=1\} = \Pr\{Y_k(x)=i\}$.

It is a composite probability that depends on survival. The conditional state probability is the probability that cohort member k occupies state i at age x, provided k is alive at x. It is denoted by ${}_k\pi_i(x)$. The relation between the two probabilities is: ${}_k\ell_i(x) = {}_k\ell_+(x) * {}_k\pi_i(x)$, where ${}_k\ell_+(x)$ is the probability that cohort member k is alive at age x. If all cohort members are identical, i.e. if the cohort is homogeneous, the state probabilities are the same for all individuals: ${}_k\ell_i(x) = {}_+\ell_i(x) = \ell_i(x)$ for all k and ${}_k\pi_i(x) = \pi_i(x)$ for all k. The traditional multistate cohort-component model relies on the homogeneity assumption. In the absence of intra-cohort variation, the expected number of cohort members in state i at age x is $K_i(x) = E[N_i(x)] = Q * \ell_i(x) = Q * \ell_+(x) * \pi_i(x)$, with Q the cohort size or radix. This expression is part of the traditional multistate life table and is implicit in the multistate cohort-component model (macro model). If individuals differ in a few characteristics only or if a few characteristics suffice to predict the state occupied at age x, then ${}_k\pi_i(x) = \pi_i(x,Z)$, where Z represents a specific combination of characteristics or covariates. The probability that individual k occupies state i at exact age x depends on the covariates only and individuals with the same covariates have the same state probability. Covariates will be introduced later.

In functional population projections, the state probabilities are estimated directly from the data (exogenous) if the static method is adopted and are generated by a multistate model (endogenous) if the dynamic method is used. Note that the prevalence rates and headship rates in static functional population projections are in fact state probabilities. We consider the estimation of state probabilities from sample data. Consider a sample of m individuals. We do not consider covariates, implying that all individuals are identical. Covariates are introduced below. In addition, age is omitted for convenience. The number of individuals observed in state i is

$$N_i = \sum_{k=1}^{m} Y_{ki}$$

The probability of observing n_1 individuals in state 1, n_2 in state 2, n_3 in state 3, etc., is given by the multinomial distribution

$$\Pr\{N_1 = n_1, N_2 = n_2, \ldots\} = \frac{m!}{\prod_{i=1}^{I} n_i!} \prod_{i=1}^{I} \pi_i^{n_i}$$

where n_i is the observed number of individuals in i. π_i is the probability that an individual is found in state i; it is the expected value of Y_i: $\pi_i = E[Y_i]$. The restrictions $\sum \pi_i = 1$ and $\sum N_i = \sum n_i = m$ apply. The most likely values of the parameters π_i, given the data, are obtained by maximizing the likelihood that the model predicts the data, which is the maximum likelihood method. The value of π_i (i = 1, 2, …, I) that maximizes the above multinomial distribution is $\hat{\pi}_i = \frac{n_i}{m}$. $\hat{\pi}_i$ is the estimate of the ??????

The expected (predicted) number of individuals occupying state i is $E[N_i] = \hat{\pi}_i\, m$. The variance of N_i is $Var[N_i] = \pi_i(1-\pi_i)m$. It is estimated as $\hat{\pi}_i(1-\hat{\pi}_i)m$. The variance of Y_i is $Var[\pi_i] = Var[N_i/m] = Var[N_i]/m^2 = [\pi_i(1-\pi_i)]/m$. It is estimated as $\hat{\pi}_i(1-\hat{\pi}_i)m$. The variance declines with increasing sample size.

The parameters used in demographic projections are frequently based on vital statistics or census data and not on sample surveys. For large m, the estimate of the state probability π_i has low variance, and estimation errors may be omitted. That is common practice in population projections. Other measurement errors should be considered, however.

State Transitions

In this section, we derive expressions for transition probabilities and transition rates. First we provide a logical link between state occupancies and state transitions. The link is useful since the static method of functional population projections focuses on state occupancies whereas the dynamic method focuses on state transitions and obtains state occupancies from the initial condition and a sequence of state transitions.

The state occupied at a given age generally depends on the states occupied at previous ages, in addition to personal attributes at the given age and prior experiences and conditions captured in the life history. Hence the probability of being in state j at age y depends on the states occupied at previous ages x_1, x_2, x_3 etc: $Pr\{Y(y) = j \mid Y(x_3), Y(x_2), Y(x_1); Z\}$ $y > x_i$ $i = 1, 2, 3$ where Z denotes contemporary and prior characteristics and experiences. It is often assumed that only the most recent state occupancy is relevant (denoted by x):

$$Pr\{Y(y) = j \mid Y(x_3), Y(x_2), Y(x_1); Z\} = Pr\{Y(y) = j \mid Y(x); Z\}$$

If the state occupied at x is i, then

$$\Pr\{Y(y) = j \mid Y(x) = i\} = p_{ij}(x, y)$$

$p_{ij}(x,y)$ is the probability that an individual who occupies state i at x occupies state j at y. It is the discrete-time transition probability. The interval can be of any length but is generally one or five years.

Transitions may be measured in continuous time and in discrete time. The distinction is consistent with the traditional distinction between two approaches to microsimulation modeling: continuous-time modeling and discrete-time modeling (eg. Galler 1997; O'Donoghue n.d.:13). We first consider **continuous time**. Let ${}_kY_{ij}(x)$ be a time-varying indicator variable which takes on the value 1 if individual k makes a move from state i to state j at exact age x, i.e. in the infinitesimally small interval following x. It is zero otherwise. The interval is sufficiently small to exclude multiple transitions. During the interval, at most one transition may occur. The number of transitions by members of the birth cohort is

$$N_{ij}(x) = \sum_{k=1}^{m} {}_kY_{ij}(x)$$

The expected value of ${}_kY_{ij}(x)$ is the probability that individual k makes a transition from i to j at age x. It depends on being alive at x and being in i at that age. The conditional transition probability is the probability of a move from i to j provided individual k is alive and in state i at age x. It is the transition probability:

$$_k\mu_{ij}(x) = \lim_{(y-x)\to 0} \frac{\Pr\{Y_k(y) = j \mid Y_k(x) = i\}}{y-x} = \lim_{(y-x)\to 0} \frac{_k p_{ij}(x,y)}{y-x}$$

It is the transition probability per unit time for very small intervals. The probability that individual k who occupies i at exact age x moves to j at that age, $_k\mu_{ij}(x)$, is known as the instantaneous rate of transition or ***transition intensity.***[3]

The unconditional transition probability is $_k\ell_{ij}(x)$ which may be written as $_k\ell_{ij}(x) = {}_k\ell_{i+}(x) * {}_k\mu_{ij}(x) = {}_k\ell_{+}(x) * {}_k\pi_i(x) * {}_k\mu_{ij}(x)$ where the first term is the probability of surviving to age x, the second the conditional state probability and the third the transition intensity. It is the event rate during the infinitesimally small interval following exact age x.

If all cohort members are identical, the transition probabilities are the same for all individuals: $_k\ell_{ij}(x) = \ell_{ij}(x)$ for all k and $_k\mu_{ij}(x) = \mu_{ij}(x)$ for all k. The traditional multistate cohort-component model (macro model) relies on the homogeneity assumption. In the absence of intra-cohort variation, the expected number of cohort members making a transition at age x from state i to state j is $K_{ij}(x) = E[N_{ij}(x)] = Q * \ell_{ij}(x) = Q * \ell_+(x) * \pi_i(x) * \mu_{ij}(x)$ where Q is the cohort size or radix.

In some applications, such as migration, the transition intensity $\mu_{ij}(x)$ is decomposed into two components: a generation component and a distribution component. The generation component is intensity of leaving the state of origin (exit rate). The distribution component is the probability of a given destination, conditional on leaving the state of origin. The transition intensity may be written as $\mu_{ij}(x) = \mu_{i+}(x) * \xi_{ij}(x)$ with $\mu_{i+}(x)$ hte instantaneous rate of leaving state i and $\xi_{ij}(x)$ the probability that an individual who leaves state i selects j as the destination. It is the conditional probability of a *direct transition* from i to j. Direct transition differ from discrete-time transitions discussed later in this paper. Direct transitions are events while discrete-time transitions refer to states occupancies at two points in time. Within an interval, several direct transitions may occur. In migration analysis and multiregional demography, direct transitions are generally referred to as *moves* (Rogers *et al.* 2002). Probabilities of a direct transition are estimated in LIFEHIST, a packaged developed by Rajulton at the University of Western Ontario. Note that $\xi_{ij}(x) = \frac{\mu_{ij}(x)}{\mu_{i+}(x)}$. Note also that the above expression is that of a competing risk model or a transition rate model with multiple destinations (Blossfeld and Rohwer 2002). In the terminology of competing risks, the first term is the rate of event and the second term (destination) indicates the type of event. The separation of the transition intensity into two parts is particularly useful when the factors that affect the occurrence of an event (movement out of a state of existence) differ from the factors that affect the type of event (direction of change or destination after the event). In that case the event occurrence and the direction of change are two distinct causal processes and $\mu_{i+}(x)$ and $\xi_{ij}(x)$ can be estimated independently (Hachen 1988:29; Sen and Smith 1995:372). The transition rate is studied using a rate model whereas the destination probability is studied using a logit or logistic regression model.

Transitions may also be measured in **discrete time** by comparing the states occupied at two consecutive ages. Consider the interval between ages x and y. Let ${}_kY_{ij}(x,y)$ be a time-varying indicator variable which takes on the value 1 if individual k occupies state i at exact age x and state j at exact age y. It is zero otherwise. State j may be the state of "dead" which is absorbing. Note that k refers to any cohort member and is not restricted to cohort members occupying state i at x. Below we consider the sub cohort of occupants of i at x.

The number of discrete-time transitions between origin state i and destination state j during the interval from x to y is equal to the (initial) number of cohort members (cohort size m) in state i at exact age x and state j at exact age y. It is denoted by $N_{ij}(x,y)$. It is

$$N_{ij}(x,y) = \sum_{k=1}^{m} {}_kY_{ij}(x,y)$$

${}_kY_{ij}(x,y)$ and $N_{ij}(x,y)$ are random variables. Let ${}_ky_{ij}(x,y)$ be an observation on ${}_kY_{ij}(x,y)$ and let $n_{ij}(x,y)$ denote the observed number of individuals in state i at age x and state j at age y:

$$n_{ij}(x,y) = \sum_{k=1}^{m} {}_ky_{ij}(x,y) = \sum_{k=1}^{m} {}_ky_i(x)\ {}_ky_j(y)$$

where ${}_ky_i(x)$ is 1 if individual k is observed to occupy state i at exact age x and 0 otherwise. The probability of observing $n_{ij}(x,y)$ cohort members occupying state i at exact age x and state j at exact age y for various i and j is given by the multinomial distribution provided the transitions are independent:

$$\Pr\{N_{12}(x,y)=n_{12}(x,y), N_{13}(x,y)=n_{13}(x,y),\ldots\}=\frac{m!}{\prod_{i=1}^{I}\prod_{j=1}^{J} n_{ij}(x,y)!}\prod_{i=1}^{I}\prod_{j=1}^{J}\ell_{ij}(x,y)^{n_{ij}(x,y)}$$

where □$_{ij}$(x,y) is the probability that cohort member occupies i at x and j at y with $\sum_i\sum_j \ell_{ij}(x,y)=1$. $n_{ij}(x,y)$ is the observed number of individuals (among the original cohort of m individuals) occupying i at age x and j at age y with $\sum_i\sum_j N_{ij}(x,y)=\sum_i\sum_j n_{ij}(x,y)=m$. The most likely values of the parameters □$_{ij}$(x,y), given the data, are obtained by maximizing the likelihood that the model predicts the data. It is the maximum likelihood method. The likelihood function is

$$\prod_{i=1}^{I}\prod_{j=1}^{J}\ell_{ij}(x,y)^{n_{ij}(x,y)}$$

The values of ℓ_{ij} (i,j = 1, 2, …, I) that maximize the above multinomial distribution is $\hat{\ell}_{ij}(x,y)=\frac{n_{ij}(x,y)}{m}$. The quantity $\hat{\ell}_{ij}(x,y)$ is the estimate of the probability that a cohort member occupies state i at age x and state j at age y, which is the expected value of $Y_{ij}(x,y)$: $\ell_{ij}(x,y) = E[Y_{ij}(x,y)]$. Note that $\ell_{ij}(x,y)$ is an unconditional probability since it relates to an initial cohort member.

The population is usually stratified by age and the base population is the number of cohort members surviving to exact age x. In that case, m is replaced by m(x) and the transition probability is the (conditional) probability that a cohort member *surviving at age x* occupies state i at age x and state j at age y. It is the probability that any person of age x occupies state j at age y, irrespective of the state occupied at x. In multistate demography, this is known as a population-based life-table measure (Willekens 1987:136ff). The transitions may be conditioned, not only on survival, but also on the state occupied at age x. It is a status-based life-table measure and is denoted by $\pi_{j|i}(x,y)$ and it is the probability that a cohort member of age x (i.e. surviving at age x) and occupying state i, will be in j at age y. The probability will be denoted by $p_{ij}(x,y)$. The probability of a given observed set of transitions is

$$\Pr\{N_{i1}(x,y)=n_{i1}(x,y), N_{i2}(x,y)=n_{i2}(x,y),\ldots|Y_i(x)=1\}=\frac{m_i(x)!}{\prod_{j=1}^{J} n_{ij}(x,y)!}\prod_{j=1}^{J}p_{ij}(x,y)^{n_{ij}(x,y)}$$

The above equation represents the competing risks model or multiple destination model (eg. Blossfeld and Rohwer 2002).

Transition Intensities, Rates and Probabilities

The probability that individual k in i transfers to j during an infinitesimally small interval following x is the instantaneous rate of transition:

$$_{k}\mu_{ij}(x) = \lim_{(y-x)\to 0} \frac{\Pr\{Y_k(y) = j \mid Y_k(x) = i\}}{y-x} = \lim_{(y-x)\to 0} \frac{_{k}p_{ij}(x,y)}{y-x} \qquad \text{for i not equal to j.}$$

In this section, we assume that all cohort members are identical. Within-cohort variation is absent. The instantaneous rate of transition is also known as the transition intensity and the force of transition. The term $\mu_{ii}(x)$ is defined such that $\sum_{j} \mu_{ij}(x) = 0$

Hence

$$-\mu_{ii}(x) = -\sum_{j \neq i} \mu_{ij}(x) = \lim_{(y-x)\to 0} \frac{p_{ij}(x)-1}{y-x}$$

The quantity $\mu_{ii}(x)$ is non-negative. The quantity $\mu_{ii}(x)$ is sometimes referred to as the intensity of passage because it relates to the transition from i to any other state different from i (eg. Namboodiri and Suchindran 1987:38). Schoen (1988:65) refers to it as the “force of retention.

The intensities are the basic parameters of a continuous-time multistate process. Under the restrictive Markov assumption, the probability that an individual leaves a state depends only on the state. It is independent of other characteristics. In this paper, the transition probability also depends on age.

The matrix of instantaneous rates with off-diagonal elements $-\mu_{ij}(x)$ and with $\mu_{ii}(x)$ on the diagonal is known as the generator of the stochastic process $\{Y_k(x); x \geq 0\}$ (Çinlar 1975:256). The matrix is denoted by $\boldsymbol{\mu}(x)$. It has the following configuration:

$$\boldsymbol{\mu}(\mathrm{x}) = \begin{bmatrix} \mu_{11}(\mathrm{x}) & -\mu_{21}(\mathrm{x}) & . & . & -\mu_{I1}(\mathrm{x}) \\ -\mu_{12}(\mathrm{x}) & \mu_{22}(\mathrm{x}) & . & . & -\mu_{I2}(\mathrm{x}) \\ . & . & . & . & . \\ . & . & . & . & . \\ -\mu_{1I}(\mathrm{x}) & -\mu_{2I}(\mathrm{x}) & . & . & \mu_{II}(\mathrm{x}) \end{bmatrix}$$

In multistate demographic models, the diagonal also includes death rates and emigration rates:

$$\mu_{ii}(x) = \mu_{id}(x) + \mu_{io}(x) + \sum_{j \neq i}^{J} \mu_{ij}(x)$$

where $\mu_{id}(x)$ is the mortality rate at age x in state i, and $\mu_{io}(x)$ the instantaneous rate of leaving state i to outside of the system.

Note that

$$\lim_{(y-x)\to 0} \frac{\mathbf{P}(x,y) - \mathbf{I}}{y-x} = -\boldsymbol{\mu}(x)$$

The matrix of discrete-time transition probabilities is:

$$\mathbf{P}(\mathrm{x, y}) = \begin{bmatrix} p_{11}(x,y) & p_{21}(x,y) & \cdot & \cdot & p_{N1}(x,y) \\ p_{12}(x,y) & p_{22}(x,y) & \cdot & \cdot & p_{N2}(x,y) \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ p_{1N}(x,y) & p_{2N}(x,y) & \cdot & \cdot & p_{NN}(x,y) \end{bmatrix}$$

and $\mathbf{P}(x,y) = \mathbf{I}$.

An element of $\mathbf{P}(x,y)$, $p_{ij}(x,y)$, denotes the (conditional) probability that an individual who is in state i at exact age x is in state j at exact age y. The Markovian assumption implies the following relationship between $\mathbf{P}(x,x+\nu)$ and $\mathbf{P}(x+\nu,y)$:

$$\mathbf{P}(x,y) = \mathbf{P}(x,x+\nu) * \mathbf{P}(x+\nu,y) \qquad x < x + \nu < y$$

Subtraction of $\mathbf{P}(x+\nu,y)$ from both sides of the equation yields

$$\frac{\mathbf{P}(x,y) - \mathbf{P}(x+\nu,y)}{\nu} = \frac{[\mathbf{P}(x,x+\nu) - \mathbf{I}]\mathbf{P}(x+\nu,y)}{\nu}$$

and

$$\lim_{\nu\to 0} \frac{\mathbf{P}(x,y) - \mathbf{P}(x+\nu,y)}{\nu} = \lim_{\nu\to 0} \frac{[\mathbf{P}(x,x+\nu) - I]\,\mathbf{P}(x+\nu,y)}{\nu}$$

or

$$\frac{d\mathbf{P}(x,y)}{dx} = -\boldsymbol{\mu}(x)\mathbf{P}(x,y)$$

Recall

$$\lim_{(y-x)\to 0} \frac{\mathbf{P}(x,y) - \mathbf{I}}{y-x} = -\boldsymbol{\mu}(x)$$

Multiplying both sides with the vector of state probabilities at age x, $\mathbf{P}(x)$, leads to:

$$\frac{d\mathbf{P}(x)}{dx} = -\boldsymbol{\mu}(x)\mathbf{P}(x)$$

where $\mathbf{P}(x)$ is a vector of state probabilities.

The model is a system of differential equations. In multistate demography, two avenues are followed to solve the system. Both introduce age intervals (Rogers and Willekens 1986:370ff). The first avenue postulates a piecewise constant intensity function, $\boldsymbol{\mu}(t) = \boldsymbol{\mu}(x)$ in the interval from x to y ($x \leq t < y$). This implies an exponential distribution of demographic events within each age interval. The model that results is referred to as the exponential model. The second avenue postulates a piecewise linear survival function. A piecewise linear survival function is obtained when demographic events are uniformly distributed within the age intervals. The model that results is referred to as the linear model. The first avenue is followed by

$$\mathbf{P}(x,y) = \begin{bmatrix} p_{11}(x,y) & p_{21}(x,y) & \cdot & \cdot & p_{N1}(x,y) \\ p_{12}(x,y) & p_{22}(x,y) & \cdot & \cdot & p_{N2}(x,y) \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ p_{1N}(x,y) & p_{2N}(x,y) & \cdot & \cdot & p_{NN}(x,y) \end{bmatrix}$$

Van Imhoff (1990) and Van Imhoff and Keilman (1991) among others; the second by Willekens and Drewe (1984) among others. The state occupancies and the sojourn times must be estimated simultaneously from the population at the beginning of the interval and the events during the interval.

To solve the system of differential equations, it may be replaced by a system of integral equations:

$$\mathbf{P}(x,y) = \mathbf{I} - \int_0^{y-x} \boldsymbol{\mu}(x+t)\,\mathbf{P}(x,x+t)\,dt$$

To derive an expression involving transition rates, we write

$$\mathbf{P}(x,y) = \mathbf{I} - \left[\int_0^{y-x} \boldsymbol{\mu}(x+t)\,\mathbf{P}(x,x+t)\,dt\right]\left[\int_0^{y-x} \mathbf{P}(x,x+t)\,dt\right]^{-1}\left[\int_0^{y-x} \mathbf{P}(x,x+t)\,dt\right]$$

$$\mathbf{P}(x,y) = \mathbf{I} - \mathbf{M}(x,y)\,\mathbf{L}(x,y) \qquad (1)$$

where $\mathbf{M}(x,y)$ is the matrix, with elements $m_{ij}(x,y)$, of average transition rates during the interval from x to y and $\mathbf{L}(x,y) = \int_0^{y-x} \mathbf{P}(x,x+t)\,dt$ is the sojourn time spent in each state between ages x and y per person in each state at age x.

i. *Exponential model*

The transition intensities $\boldsymbol{\mu}(x)$ are assumed to remain constant during the age interval from x to y and to be equal to the model transition rates $\mathbf{M}(x,y)$.

It is furthermore assumed that they are can be estimated by empirical occurrence-exposure rates for that age interval. This assumption is consistent with the general assumption in demography that life-table rates are equal to empirical rates. In this paper no separate notation is used for model rates and empirical rates. The matrix of transition probabilities between x and y is

$$\mathbf{P}(x,y)=\exp\left[-(y-x)\mathbf{M}(x,y)\right]$$

where **M**(x,y) is the matrix of empirical occurrence-exposure rates or transition rates for the age interval from x to y and $\mu_{ij}(t) = m_{ij}(x,y)$ for $x \leq t < y$ and □(t) = **M**(x,y) for $x \leq t < y$.

A number of methods exists to determine the value of exp[-**M**] (eg. Director and Rohrer 1972:431ff; Aoki 1976:387; Strang 1980:206). We use the Taylor series expansion. Note that for matrix **A**, exp(**A**) may be written as a Taylor series expansion

$$\exp(\mathbf{A})=\mathbf{I}+\mathbf{A}+\frac{1}{2!}\mathbf{A}^2+\frac{1}{3!}\mathbf{A}^3+\ldots$$

Hence

$$\exp[-(y-x)\mathbf{M}(x,y)]=\mathbf{I}-(y-x)\mathbf{M}(x,y)+\frac{(y-x)^2}{2!}\left[\mathbf{M}(x,y)\right]^2-\frac{(y-x)^3}{3!}\left[\mathbf{M}(x,y)\right]^3+\ldots$$

(see also Schoen 1988:72).

The transition rates **M**(x,y) are estimated from the data. The transition rate $m_{ij}(x,y)$ is equal to the ratio of the number of moves or direct transitions from i to j during the interval from x to y, and the duration of exposure spent in the state of origin i:

$$m_{ij}(x,y)=\frac{n_{ij}(x,y)}{L_i(x,y)}$$

where $n_{ij}(x,y)$ is the observed number of moves from i to j during the interval and $L_i(x,y)$ is the duration in i exposed to the risk of moving to j. It is the sojourn time in i during the (x,y) interval. Exposure is measured in person-months or person-years. In case of two states, the rate equation may be written as follows:

$$\begin{bmatrix} m_{11}(x,y) & -m_{21}(x,y) \\ -m_{12}(x,y) & m_{22}(x,y) \end{bmatrix}=\begin{bmatrix} n_{11}(x,y) & -n_{21}(x,y) \\ -n_{12}(x,y) & n_{22}(x,y) \end{bmatrix}\begin{bmatrix} L_1(x,y) & 0 \\ 0 & L_2(x,y) \end{bmatrix}^{-1}$$

where $m_{11}(x,y) = m_{12}(x,y)$ and $m_{22}(x,y)=m_{21}(x,y)$. In multistate demography, the state of dead is gererally not treated as a separate state and the death rate is included in the diagional elements, e.g. $n_{11}(x,y) = m_{12}(x,y)+m_{1d}(x,y)$

where $m_{1d}(x,y)$ is the number of deaths among cohort members occupying state 1 and aged x to y.
In matrix notation: $\mathbf{M}(x,y)=\mathbf{n}(x,y)\left[\mathbf{L}(x,y)\right]^{-1}$

Let $\overline{\mathbf{L}}(x,y)$ be the vector of sojourn times containing the diagonal elements of **L**(x,y) and let **K**(x) be a vector with the state occupancies at age x by surviving cohort members as its elements:

$$\mathbf{K}(x)=\begin{bmatrix}K_1(x)\\K_2(x)\end{bmatrix}$$

with $K_i(x)$ the number of cohort members in state i at exact age x. The vector of sojourn times by all cohort members in the various positions is obtained by the following equation:

$$\overline{\mathbf{L}}(x,y)=\left[\int_0^{y-x}\mathbf{P}(x,x+t)\,dt\right]\mathbf{K}(x)$$

Since the transition intensities are constant in the interval from x to y, the equation may be written as follows:

$$\overline{\mathbf{L}}(x,y)=\left[\int_0^{y-x}\exp(-t\,\mathbf{M}(x,y)\,dt\right]\mathbf{K}(x)$$

Integration yields

$-\left[\mathbf{M}(x,y)\right]^{-1}\left|\exp\left[-t\,\mathbf{M}(x,y)\right]\right|_0^{y-x}$ which is equal to

$$\left[\mathbf{M}(x,y)\right]^{-1}\left[\mathbf{I}-\exp\left[-(y-x)\,\mathbf{M}(x,y)\right]\right]$$

Hence the sojourn times in the various states during the (x,y)-interval are given by:

$$\overline{\mathbf{L}}(x,y)=\left[\mathbf{M}(x,y)\right]^{-1}\left[\mathbf{I}-\exp\left[-(y-x)\,\mathbf{M}(x,y)\right]\right]\mathbf{K}(x)$$

ii. *Linear model*

To solve equation (1), one may introduce an approximation of **L**(x,y). A simple approximation is that **P**(x,x+t) is linear on the interval x ≤ x+t < y. Hence $\mathbf{L}(x,y)=\int_0^{y-x}\mathbf{P}(x,x+t)\,dt$ may be approximated by a linear integration:

$$\mathbf{L}(x,y)=\frac{y-x}{2}\left[\mathbf{I}+\mathbf{P}(x,y)\right]$$

Introducing this expression in equation (1) gives

$$\mathbf{P}(x,y)=\mathbf{I}-\frac{y-x}{2}\mathbf{M}(x,y)\left[\mathbf{I}+\mathbf{P}(x,y)\right]$$

$$\mathbf{P}(x,y) = \mathbf{I} - \frac{y-x}{2}\mathbf{M}(x,y)\,\mathbf{I} + \frac{y-x}{2}\mathbf{M}(x,y)\,\mathbf{P}(x,y)$$

$$\mathbf{P}(x,y) + \frac{y-x}{2}\mathbf{M}(x,y)\,\mathbf{P}(x,y) = \mathbf{I} - \frac{y-x}{2}\mathbf{M}(x,y)\,\mathbf{P}(x,y)$$

$$\mathbf{P}(x,y) = \left[\mathbf{I} + \frac{y-x}{2}\mathbf{M}(x,y)\right]^{-1}\left[\mathbf{I} - \frac{y-x}{2}\mathbf{M}(x,y)\right]$$

The linear approximation implying the assumption that the events are uniformly distributed over the interval is adequate when the transition rates are small or the interval is short. It can be shown that the linear model is an approximation to the exponential model that retains the first three terms of the Taylor series expansion (Annex I).

In the previous section we described the separation of the origin-destination specific transition rate into two components, a generation component and a distribution component. The discrete-time transition probabilities are related to the probabilities of direct transition in an interesting way. The off-diagonal elements of **M**(x,y) may be replaced by $-m_{i+}(x,y)\ \xi_{ij}(x,y)$ where $m_{i+}(x,y)$ is the rate of leaving i (exit rate), which is assumed to be constant in the interval from x to y. The diagonal elements are $m_{i+}(x,y)$. The **μ**-matrix may be written as

$$\begin{bmatrix} \mu_{11}(x,y) & -\mu_{21}(x,y) & . & . & -\mu_{I1}(x,y) \\ -\mu_{12}(x,y) & \mu_{22}(x,y) & . & . & -\mu_{I2}(x,y) \\ . & . & . & . & . \\ . & . & . & . & . \\ -\mu_{1I}(x,y) & -\mu_{2I}(x,y) & . & . & \mu_{II}(x,y) \end{bmatrix} = \begin{bmatrix} \xi_{11}(x,y) & -\xi_{21}(x,y) & . & . & -\xi_{I1}(x,y) \\ -\xi_{12}(x,y) & \xi_{22}(x,y) & . & . & -\xi_{I2}(x,y) \\ . & . & . & . & . \\ . & . & . & . & . \\ -\xi_{1I}(x,y) & -\xi_{2I}(x,y) & . & . & \xi_{II}(x,y) \end{bmatrix} \begin{bmatrix} \mu_{1+}(x,y) & 0 & . & . & 0 \\ 0 & \mu_{2+}(x,y) & . & . & 0 \\ . & . & . & . & . \\ . & . & . & . & . \\ 0 & 0 & . & . & \mu_{I+}(x,y) \end{bmatrix}$$

with ξ(x,y) the probability of at least one direct transition from i to j during the interval from x to y.

Population Projection

$$\mathbf{P}(x,y) = \begin{bmatrix} p_{11}(x,y) & p_{21}(x,y) & . & . & p_{N1}(x,y) \\ p_{12}(x,y) & p_{22}(x,y) & . & . & p_{N2}(x,y) \\ . & . & . & . & . \\ . & . & . & . & . \\ p_{1N}(x,y) & p_{2N}(x,y) & . & . & p_{NN}(x,y) \end{bmatrix}$$

$$= \exp\left[-(y-x) \begin{bmatrix} \xi_{11}(x) & -\xi_{21}(x) & . & . & -\xi_{I1}(x) \\ -\xi_{12}(x) & \xi_{22}(x) & . & . & -\xi_{I2}(x) \\ . & . & . & . & . \\ . & . & . & . & . \\ -\xi_{1I}(x) & -\xi_{2I}(x) & . & . & \xi_{II}(x) \end{bmatrix} \begin{bmatrix} \mu_{1+}(x) & 0 & . & . & 0 \\ 0 & \mu_{2+}(x) & . & . & 0 \\ . & . & . & . & . \\ . & . & . & . & . \\ 0 & 0 & . & . & \mu_{I+}(x) \end{bmatrix} \right]$$

The distribution of a population of a given age at a given point in time is represented by the vector of *state occupancies* K(x,t)

$$\mathbf{K}(x,t) = \begin{bmatrix} K_1(x,t) \\ K_2(x,t \\ \dots \\ K_N(x,t) \end{bmatrix}$$

where $K_i(x,t)$ is the number of persons of age x in state i at time t.

In this section we consider a projection model that allows for international migration. The number of people at a given age above 0 depends on an initial condition, deaths, interstate transitions, emigrations during an interval and immigrants during an interval.

i. *Exponential model*

The exponential model of multistate population growth is derived from a system of differential equations

$$\frac{d\,\mathbf{K}(t)}{d\,t} = -\mathbf{M}(t)\,\mathbf{K}(t) + \mathbf{F}(t)\,\mathbf{I}_m(t)$$

where $\mathbf{K}(t)$ is a vector of state occupancies, i.e. the number of individuals in the various states, $\mathbf{I}_m(t)$ is the vector of immigrants at time t by state of existence, and $\mathbf{M}(t)$ and $\mathbf{F}(t)$ are coefficient matrices. We consider the age interval from x to y and introduce piecewise constant rates $\mathbf{M}(x,y)$. The projection model expresses the state occupancies of the cohort members at age y in terms of the state occupancies at age x and the transition rates during the interval from x to y. The solution to the system of differential equations is

$$\mathbf{K}(y) = \exp\left[-(y-x)\mathbf{M}(x,y)\right]\mathbf{K}(x) + \int_0^{y-x} \exp\left[-(y-t)\mathbf{M}(x,y)\right]\mathbf{I}_m(t)\,dt$$

Since the instantaneous rates are assumed constant in the interval (x,y) and if immigration $\mathbf{I}_m$ is uniformly distributed then

$$\mathbf{K}(y) = \exp\left[-(y-x)\mathbf{M}(x,y)\right]\mathbf{K}(x) + \int_0^{y-x} \exp\left[-(y-t)\mathbf{M}(x,y)\,dt\right]\mathbf{I}_m \text{ and}$$

$$\mathbf{K}(y) = \exp\left[-(y-x)\mathbf{M}(x,y)\right]\mathbf{K}(x) + \left[\mathbf{M}(x,y)\right]^{-1}\left[\mathbf{I} - \exp\left[-(y-x)\mathbf{M}(x,y)\right]\right]\mathbf{I}_m$$

$$= \mathbf{G}(x,y)\mathbf{K}(x) + \mathbf{F}(x,y)\mathbf{I}_m$$

G(x,y) represents the contribution of cohort members present at age x to the population at age y and **F**(x,y) represents the contribution of immigrants aged x to y.

ii. *Linear model*

The multistate projection model predicts the state occupancies from information on state occupancies at a previous point in time and immigration:

$$\mathbf{K}(y) = \mathbf{P}(x,y)\,\mathbf{K}(x) + \mathbf{F}(x,y)\,\mathbf{I}_m$$

where **P**(x,t) is the matrix of transition probabilities for persons aged x, and $\mathbf{I}_m$ is a vector representing the number of immigrants during a unit interval. **F**(x,y) is the coefficient matrix that denotes the contribution of immigrants during a given period to the population at the end of the period. The coefficient matrices are related to the transition rates (Willekens 1998):

$$\mathbf{P}(x,y) = \left[\mathbf{I} + \frac{y-x}{2}\mathbf{M}(x,y)\right]^{-1}\left[\mathbf{I} - \frac{y-x}{2}\mathbf{M}(x,y)\right]$$

and

$$\mathbf{F}(x,y) = \left[\mathbf{I} + \frac{y-x}{2}\mathbf{M}(x,y)\right]^{-1}$$

Transition Probability Models and Transition Rate Models in Population Forecasting

Now we introduce covariates, including contextual variables. Covariates may be introduced in two ways: by stratifying the population by the relevant covariate(s), provided the covariates are discrete variables, or by a regression equation. In population projections, stratification is generally used. For instance, a population is usually stratified by sex and birth cohort (age) and demographic parameters are sex- and birth cohort-specific. Regression models are more economic at a cost of precision. Stratification involves as many parameters as there are cells in the cross-classification of covariates. The same is true only in a saturated regression model, which is a model that has as many independent parameters as there are unknowns (cells in the cross-classification). A saturated regression model involves several interaction effects, which may be redundant in most practical applications.

In the regression model, the covariates are denoted by **Z** ($\mathbf{Z} = \{Z_1, Z_2, Z_3, \ldots\}$). A covariate Z_p may represent a single attribute or a combination of attributes (to denote interaction effects). In addition to covariates one may include time (when the rates are time-varying) and/or cohort (when rates are estimated for different cohorts) among the explanatory variables or predictors. Furthermore, the transition rates may depend on the entire previous life course: $\mathbf{M}(x,\Theta[0,x])$ where $\Theta\,[0,x]$ represents the life course from birth to age x. We consider state probabilities and transition rates. State probabilities are related to covariates using a logit model or logistic regression. Transition rates are related to covariates using transition rate models that are related to the family of Poisson regression models. Logit models are also applied to link transition probabilities to covariates.

iii. *State probabilities*

The **state probability** at age x, $\pi_i(x,\mathbf{Z})$, is the probability that an individual of age x and with covariates **Z** occupies state i. The logit equation relates the state probabilities to covariates.

$$logit[\pi_i(x,\mathbf{Z})] = \ln\frac{\pi_i(x,\mathbf{Z})}{\pi_r(x,\mathbf{Z})} = \eta_i(x,\mathbf{Z}) = \beta_{i0}(x) + \beta_{i1}(x)Z_1(x) + \beta_{i2}(x)Z_2(x) + \beta_{i3}(x)Z_3(x) + \ldots$$

where $\dfrac{\pi_i(x,\mathbf{Z})}{\pi_r(x,\mathbf{Z})}$ is the odds that a cohort member of age x occupies state i rather than the reference state r (reference category). The logit transformation assures that the state probabilities lie between 0 and 1, and that their sum is equal to one. The value of η_i may range from $-\infty$ to $+\infty$, but the value of π_i stays within 0 and 1. To obtain the probabilities, the logit scale is converted into the probability scale:

$$\pi_i(x,\mathbf{Z}) = \frac{\exp[\eta_i(x,\mathbf{Z})]}{\exp[\eta_1(x,\mathbf{Z})] + \exp[\eta_2(x,\mathbf{Z})] + \ldots} = \frac{\exp[\eta_i(x,\mathbf{Z})]}{\sum_{j=1}^{I}\exp[\eta_j(x,\mathbf{Z})]}$$

where the 1 is associated with the reference category. The model is the multinomial logistic regression model.

iv. Transition rates

The **transition rates** may depend on time [$\mu_{ij}(x,t)$] and may also depend on personal characteristics. For instance $\mathbf{M}(x,\mathbf{Z})$ is the matrix of transition rates for individuals of age x and the vector of background variables **Z** ($\mathbf{Z} = \{Z_1, Z_2, Z_3, \ldots\}$). The dependence of transition rates on personal attributes and other explanatory variables is described by transition rate models, better known as hazard models (Tuma and Hannan 1984; Blossfeld and Rohwer 2002). These models are Poisson regression models. Transition rate models include the basic (exponential) transition rate model, the piecewise constant rate model, and the Cox regression model. Transition rate models also include parametric model of time- (or age-)dependence such as the Gompertz model, the Weibull model, the Coale-McNeil model (for fertility and nuptiality), the Heligman-Pollard model (for mortality) and the Rogers-Castro model (for migration). Transition rate models are estimated from empirical data. The data may be vital statistics, census data or surveys. The estimation of transition rates from survey data require recently developed theories of statistical inference.

The basic parameters of the multistate model are the transition rates $m_{ij}(x,y)$. As described in a previous section, the transition rate may be written as the product of an exit rate $m_i(x,y)$ and a conditional transition probability. The exit rate is modeled using a transition rate model for a single event (leaving the state of origin). The elementary transition rate model is the basic exponential model, with the rate being independent of age (Blossfeld and Rohwer 2002). The regression model linking an exit rate to covariates is

$$m_i = \exp[\beta_{i0} + \beta_{i1} Z_1 + \beta_{i2} Z_2 + \ldots]$$

and a regression model linking transition rates to covariates is

$$m_{ij} = \exp[\beta_{ij0} + \beta_{ij1} Z_1 + \beta_{ij2} Z_2 + \ldots]$$

The models may be written as log-linear models

$$\ln m_i = \beta_{i0} + \beta_{i1} Z_1 + \beta_{i2} Z_2 + \ldots$$

$$\ln m_{ij} = \beta_{ij0} + \beta_{ij1} Z_1 + \beta_{ij2} Z_2 + \ldots$$

The transition rate is the ratio of number of events over total exposure time. These two components may be studied separately, as is done in the log-rate model (eg. Yamaguchi 1991:Chapter 4). It that model, it is assumed that changes in the number and timing of direct transitions (events) do not significantly affect the total exposure time. The assumption is realistic when exposure time is large compared to the number of transitions. If a variation in number or timing of transitions does not affect total exposure, the latter component may be considered fixed and may be treated as an *offset* in probability models including regression models. The problem of modeling transitions reduces to the prediction of the number of events (counts) which is the numerator of the transition rate. The number of direct transitions that occur during a unit interval is often represented by a Poisson random variable. The number of events that may occur during the interval is not restricted in any way. Subjects in a (sample) population may experience more than one event during the unit interval. The Poisson model is

$$\Pr\{N_{\mathrm{ij}} = \mathrm{n}_{\mathrm{ij}}\} = \frac{\lambda_{ij}^{n_{ij}}}{n_{ij}!} \exp[-\lambda_{\mathrm{ij}}]$$

where N_{ij} denotes the number of transitions from i to j, n_{ij} is the observed number of transitions, and λ_{ij} is the expected number of transitions. The latter is the parameter of the Poisson model. It is assumed that the transitions are independent. The parameter may be made dependent on covariates:

$$E\lfloor N_{ij} \rfloor = \lambda_{ij} = \exp\lfloor \beta_{ij0} + \beta_{ij1} Z_1 + \beta_{ij2} Z_2 + \ldots \rfloor$$

The model may be written as a log-linear model:

$$\ln \lambda_{ij} = \beta_{ij0} + \beta_{ij1} Z_1 + \beta_{ij2} Z_2 + \ldots$$

In principle, Z_p can be any covariate. In conventional log-linear analysis, all covariates are discrete or categorical. The observations on transitions may therefore by arranged in a contingency table. The covariates refer to rows, columns, layers and combinations of these (to represent interaction effects).

The log-rate model is a log-linear model with an offset:

$$E\left[\frac{N_{ij}}{PY_i}\right] = \frac{\lambda_{ij}}{PY_i} = \exp\left[\beta_{ij0} + \beta_{ij1} Z_1 + \beta_{ij2} Z_2 + \ldots\right]$$

where PY_i denotes exposure time in state i (origin state). Since PY_i is fixed, the equation may be rewritten as follows:

$$E\lfloor N_{ij} \rfloor = \lambda_{ij} = PY_i \exp\lfloor \beta_{ij0} + \beta_{ij1} Z_1 + \beta_{ij2} Z_2 + \ldots \rfloor$$

The age dependence may be introduced in two ways: non-parametric and parametric. In the first approach, the population is stratified by age and a transition rate is estimated for each age separately. In the parametric approach, age dependence is represented by a model. A common model is the Gompertz model, which imposes onto the transition rate an exponential change with duration. The Gompertz model has two parameters and each may be made dependent on covariates (For a detailed treatment, see Blossfeld and Rohwer 2002). Other parametric models of duration dependence may be used. In studies of marriage and fertility, the Coale-McNeil model is often used to describe the age dependence of the marriage or first birth rate (eg. Liang 2000). In migration studies, the model migration schedule is a common representation of the age dependence of the migration rate (eg. Rogers and Castro 1986). Each parameter of the model may be related to covariates. In practice, only one or a selection of parameters is assumed to depend on covariates.[4]

In some cases, the researcher is not interested in the age dependence of transition rates, but in the effect of covariates on the level of transition. Rather than omitting age altogether, as in the basic exponential model, the transition rate is allowed to vary with age but the effect of the covariates on the transition rate does not vary with age. The transition rate model that results is a Cox proportional hazard model. It is written as

$$m_{ij}(x) = m_{ijo}(x)\exp\left[\beta_{ij0} + \beta_{ij1}Z_1 + \beta_{ij2}Z_2 + \ldots\right]$$

where $m_{ijo}(x)$ is the baseline hazard. It is the set of age-specific transition rates for the reference category. Note that if the age dependence (age structure) of transition is independent of the dependence on covariates (motivational structure), the baseline hazard may be represented by a parametric model and the two components may be estimated separately.

From Transition Probabilities to Transition Rates

In this section, we assume that transition is measured in discrete-time. Examples include the census (based on the residence at time of census and 5 years prior to the census). From that information, the approximate transition rates can be derived. The problem is equivalent to one in which we are given $\mathbf{P}(x,y)$ and $\mathbf{M}(x,y)$ is required. The derivation starts with the exponential expression $\mathbf{P}(x,y) = \exp\left[-(y-x)\mathbf{M}(x,y)\right]$. The exponential expression may be approximated by the linear model:

$$\mathbf{P}(\mathrm{x},\mathrm{y}) = \left[\mathbf{I} + \tfrac{1}{2}\mathbf{M}(\mathrm{x},\mathrm{y})\right]^{-1}\left[\mathbf{I} - \tfrac{1}{2}\mathbf{M}(\mathrm{x},\mathrm{y})\right]$$

The approximation is adequate when the transition rates are small or the interval is short.

The derivation of the rate of transition during an interval from information in regions of residence at two consecutive points in time is known as the inverse problem: transition rates are derived from transition probabilities (Singer and Spilerman 1979).

$$\mathbf{M}(x,y) = \tfrac{y-x}{2}\left[\mathbf{I} - \mathbf{P}(x,y)\right]\left[\mathbf{I} + \mathbf{P}(x,y)\right]^{-1},$$

provided $\left[\mathbf{I}+\mathbf{P}(x,y)\right]^{-1}$ exists. The inverse relation may be used to infer transition probabilities for intervals that are different from the measurement intervals. For instance, if the states occupied are recorded at age x and at age y, the inverse relation may be used to infer the average transition rates **M**(x,y) and to derive the transition probabilities over a one-year period. The expression is $\mathbf{P}(x,x+1)=\exp\left[-\mathbf{M}(x,y)\right]$ where **M**(x,y) is estimated from **P**(x,y) using the inverse method. The method assumes that transition rates are constant during the (x,y)-interval and that the linearity assumption is an adequate approximation of the exponential model.

Transition Rates: The Bridge to Demographic Scenarios and Stochastic Projections

The transition rates (or in some cases the transition probabilities) also represent the ultimate scenario variables. Scenarios are often formulated in terms of demographic indicators such as life expectancy and TFR, but these measures must be translated into age (and sex) specific transition rates. Scenarios may be expressed directly in terms of the transition rates or in terms of the covariates that predict the transition rates. For instance, if one is interested in future pension payments, the age at retirement may be considered a scenario variable or it may be predicted by a transition rate model with the rate of retirement depending on a set of personal attributes and policy variables.

The transition rates (or transition probabilities) also constitute the ultimate variables for stochastic projection. Predicted values of demographic indicators such as life expectancy or TFR (predicted by expert opinions, a mathematical model extrapolating past values and errors, or a combination) must be translated into age (and sex) specific transition rates. Stochastic projections are based on the assumption that the transition rates are not point estimates but interval estimates following particular probability distributions around mean or expected values (e.g. normal distribution or beta distribution). A two-step procedure involving the random selection of values from a distribution and the projection of the population using these values constitutes a simulation experiment. Repeated simulation experiments produce information on the distribution of target variables such as the population aged 65+, the dependency ratio, the number of years spent in pension and the number of years with severe disability.

The coefficients of the regression models predicting the transition rates or transition probabilities may also vary in time or over individuals with the same attributes. In that case the model with fixed effects (fixed effects model) changes into a varying effect model (which includes as a special case the random effects model that describes the distribution of effects of covariates among a group of people with the same attributes). The model may become very complex and may be manageable only in a micro-simulation mode.

Individual Biographies and Microsimulation

The multistate model is also the basis for the prediction of individual biographies. In the previous sections, an individual was represented by k. To project k's biography, information on the attributes of k, on other explanatory variables and maybe some assumptions about patterns of change are used to determine for each age the transition rate experienced by k. Transition rates are predicted from explanatory variables using a regression model (transition rate model). The technique is used extensively in medical sciences (eg. Anderson *et al.* 1990; Mamun 2004). The method is also related to micro-econometric models of labour market dynamics (eg. Flinn and Heckman 1982a; 1982b).

The multistate model specified above is a macro-simulation model that uses point estimates of the transition rates and predicts expected values, e.g. the *expected* number of individuals in a given state at a given future time. The transition rates may depend on covariates, eg. sex. The expected value is the mean value in a population of individuals. Individual values, e.g. states occupied by individuals in a population at a given future point in time, are obtained by randomly allocating events (transitions) to members of the population in a way that is consistent with (a) the point estimates of the transition rates (mean or expected value) (prior estimate) and (b) the distribution of the individual values around the mean. It implies that the transition rates follow a given probability distribution. Individual values are generally not produced for all members of a population but for a (random) sample. Using a random number generator, a unique value is selected from a distribution for each individual in the sample population. If for a given individual, the value is less than the mean value in the population (point estimate), the event is allocated to the individual and the individual makes the transition from the state of origin to the state of destination. If the value drawn from the distribution exceeds the mean value, the individual remains in the state of origin. The transition rates that result after the events have been allocated to the sample population (posterior estimates) generally differ slightly from the prior estimates. The reason is sample variation. If the sample is sufficiently large, the posterior estimates coincide with the prior estimates. That is why microsimulations require large sample populations to produce reliable estimatesfor the population characteristics.

Wolf (2000) views microsimulation as the generation of data. He sees microsimulation fundamentally as an exercise in sampling: "Microsimulation consists of drawing a sample of realizations of a prespecified stochastic process" (Wolf 2000:2). The same view is adopted in the MicMac project. The data consists of realizations of an underlying probability mechanism. Probability models constitute the core of microsimulation and predictions generated by the probability models represent the output of microsimulation. Wolf also observed that the emphasis in most microsimulation is on the outputs generated by the simulation, rather than on the process of model development, estimation, and assessment. He argues that microsimulation has much to offer in these modeling steps. The MicMac project is an exercise in model development, estimation and assessment. The estimation of model parameters (transition rates in continuous-time models and transition probabilities in discrete-time models) from data uses the theory of statistical inference when required, i.e. when sampling is involved.

The MicMac Project

The Partners and the Work Packages

The biographic projection model is developed and implemented by a consortium of research institutes in Europe. The following institutes participate (coordinator in parentheses): NIDI in The Hague (Nicole van der Gaag), IIASA and Vienna Institute of Demography (VID) (Wolfgang Lutz), Bocconi University in Milan (Francesco Billari), Department of Public Health, Erasmus Medical Center in Rotterdam (Wilma Nusselder), Max Planck Institute for Demographic Research in Rostock (Gabrielle Doblhammer-Reiter), INED in Paris (Laurant Toulemon). General coordination is by the author. Model development is concentrated at NIDI. IIASA will develop methods and procedures to efficiently derive scientifically sound argument-based expert views on future trends of demographic variables utilising insights from the fields of cognitive science, group dynamics and quantitative decision analysis. Other partners will use the instruments developed at IIASA to produce argument-based scenarios and develop uncertainty distributions around most likely values of demographic parameters of multistate models.

The Erasmus Medical Center, in cooperation with the Max Planck Institute for Demographic Research, will determine the age (risk) profiles of key events in morbidity and mortality in the life course and determine the relative risks of these events in relation to proximate risk factors (e.g. smoking, blood pressure, body mass index) and to more distant determinants such as socio-economic status (SES.) This will serve as input for the illustrative projections using the MicMac micro-simulation approach. In addition they will develop mortality and morbidity scenarios using the instruments for argument-based scenarios developed at IIASA.

Bocconi University, in cooperation with VID and INED will provide the input needed to make forecasts on fertility and family and household structure. In addition, argument-based scenarios will be developed.

Both the macro-model (Mac) and the micro-model (Mic) project biographies in terms of state occupancies and transitions. Mac projects cohort biographies. It projects the (expected) number of members of a birth cohort that occupy various states at a future date. It uses mean or expected values of transition rates and transition probabilities. Mic projects individual biographies. It predicts the state occupied by a given individual at a given future point in time. To project individual biographies, Mic uses information on the attributes of the individual, on other explanatory variables and maybe some assumptions about patterns of change to determine for each age the transition rate experienced by the individual. The transition rates are predicted from explanatory variables using a regression model (transition rate model).

VID will also prepare multi-state population projections by level of education. This research will deliver to all other components of MicMac a substantive contribution about changes in the educational composition of the total population (with special emphasis on the working age population). It will define alternative scenarios about future transition rates considering ongoing plans for school reforms and it will cover the analysis of interactions between education and the timing of fertility over the life course.

Software Development

A number of software packages for demographic projection already exist (Willekens and Hakkert 1992; http://www.eat.org.mx/software/softpyd.htm for an update). Packages that implement the cohort- component method include PDPM/PC (Population and Development Projection Methods for Microcomputers), developed under auspices of the United Nations, and PEOPLE, developed by Richard Leete, for national and subnational population projections.

Packages for *multistate* demographic modelling have been developed largely as part of methodological research. The first packages were developed at IIASA in the 1970s by Ledent and Willekens (SPA by Willekens and Roger, 1978; LIFEINDEC by Willekens 1979). They included the multistate life table and multistate projections. The work was carried further at NIDI. That resulted in LIPRO by Van Imhoff and Keilman (1992), FAMY by Zeng Yi (1991), PROFAMY by Zeng Yi (Zeng Yi *et al.* 1997) and MUDEA by Willekens (Willekens and Drewe 1984; Willekens 1995). LIPRO was originally developed for multistate household projection but has been designed as to be generally applicable. It is user-friendly and has been applied in several countries of Europe. At IIASA the first interactive user-friendly software for multistate population projections - DIALOG (Scherbov *et al.* 1986) was developed by Scherbov.

Software for stochastic population projections is not generally available. PEP (Program for Error Propagation) developed by Alho (n.d.) is not generally accessible but has been extensively documented. It is used in the DEMWEL project in the Fifth Framework Programme of the European Commission. The manual is available on the internet: http://joyx.joensuu.fi /~ek/pep/userpep.htm

Dynamic longitudinal microsimulation models simulate life histories of individuals and families. Macro-simulation models, such as demographic projection models, deal with individuals grouped by concerned attributes, for instance a group of persons of the same age and parity and/or marital status. Micro-simulation models simulate life course events and keep detailed records of demographic status transitions for each individual of the sample population. The models trace the influence of a large number of decisions or events on the life course of people. Microsimulation models that are designed to simulate life histories adopt a life course perspective and involve behavioural rules that determine the occurrence of events and ages at events. Dynamic longitudinal microsimulation models have been the subject of extensive reviews (Van Imhoff and Post 1998). Recent reviews were presented at the *Dynamic Microsimulation Modelling Technical Workshop*, held in January 2002 at the London School of Economics (http://www.lse.ac.uk/ Depts/sage/conference/workshop.htm) and at the *International Micro-simulation Conference on Population Ageing and Health: Modelling Our Future*, held in Canberra, Australia in December 2003. Among the microsimulation models, SOCSIM, developed at the University of California at Berkeley, is one of the oldest but continuously updated models, and LifePaths of Statistics Canada is one of the best known. The latter has inspired the development of microsimulation models in several countries. Other models are SABRE in the UK, DESTINIE in France, NATSIM in Australia, SWITCH in Ireland and MOSART in Norway. MicMac differs from these models in the special interest in micro-macro linkages.

As part of the MicMac project, a user-friendly software package will be developed for biographic forecasting (cohort biographies and individual biographies). Biographic projections provide a generic approach to functional population projections. The software implementing Mac includes a facility to link the rates/probabilities to explanatory variables using rate models or probability models (regression models). The link is operationalized in a separate module that can be activated by the user.

The software will be object-oriented. Three broad objects are distinguished: a **pre-processor** to produce the data base for projection (input), the **processor** that represents the projection engine and stores the results in a data base (output), and the **post-processor** to process the projection results. At a more specific level, objects refer to algorithms or procedures to obtain particular estimates or predictions, to determine prediction error, to impose empirical regularities (e.g. using model schedules), to calculate summary indicators, etc. The individual objects will be documented and assembled in an **object library**. That approach enables and simplifies the further development of MicMac after the project is completed (by persons not involved in the initial development). It is an important strategy towards sustainability.

What Does the Project Contribute?

The proposed project signifies a step change in projection methodology. It specifically adds the following:

1. The multistate cohort-component model will be extended to project cohort biographies covering a number of domains of life (e.g health, family and education) that constitutes the basis for integrated demographic projections involving a number of domains of life.
2. The multistate cohort-component model will be extended to project individual biographies. This represents a scientifically sound strategy towards disaggregated demographic projections that cover several domains of life and that consider detailed characteristics of the population.
3. Uncertainty will be dealt with in an innovative way, using recent findings from cognitive science and the study of group dynamics. In addition, uncertainty will be measured at the level of demographic events. That level is appropriate since demographic forecasting errors are caused by errors in predicting births, deaths and other demographic events. An important advantage of this approach is that it will be easier to communicate the results of stochastic projections to policy makers and the general public.
4. For the first time ever, demographic projection models will be available that take advantage of the important advances during past decades in the statistical analysis of lifetime data and event histories. That development concentrated on the explanation of transition rates (hazard rates) and transition probabilities.
5. For the first time ever, a generic software package will become available to explore demographic futures of Europe that cover detailed population categories. Although the software will be presented as a easy-to-use package, it has a modular structure and consists of a large number of objects. The availability of an object library will enhance the further development of MicMac and other demographic software that may use these objects.
6. MicMac will be readily applicable to investigate the impact of lifestyle factors and other major health determinants on specific chronic diseases and the health status in general, and to assess the effects of different types of interventions on the health status of different groups in the population. What differentiates the proposed methodology from

other projection and impact assessment methods is (1) the focus on incidence (transitions) rather than prevalence and (2) the correct treatment of the effects of competing risks by adopting a multistate setting.

Conclusion

Biographic forecasting is a new approach to demographic forecasting that integrates traditional forecasting of the population by age and sex and functional population forecasting. Demographic change is increasingly difficult to forecast because the idiosyncratic nature of demographic behaviour. Family formation, migration and attempts to grow old healthy are part of a lifestyle that also includes work and other domains of life. The interest in life strategies originates from the awareness that critical decisions in life, i.e. decisions related to life events, are not taken in isolation but are based on lessons learned from past experiences (antecedents) and general conceptions about future developments in different life domains. Increasingly, demographic events are embedded in a life plan. The individual programming of life events, viewed by Légaré and Marcil-Gratton (1990) as a challenge for demographers in the twenty-first century, may be situated within the broader context of the individual design and implementation of life strategies. The second demographic transition, with its emphasis on choice biographies and changing interpersonal relationships, may also be viewed as a consequence of individualization and the emergence of individual life strategies.

The new demographic regime characterized by choice biographies and life planning raises new challenges for demographic forecasters. Traditional projections by age and sex do not adequately capture the complexity of life. As a result the uncertainty increases. Probabilistic projections quantify the uncertainty but are not able to reduce the uncertainty. Reduction of uncertainty and increase of forecasting performance require more realistic projection models, i.e. models that are better able to integrate substantive knowledge and to capture the causal links that underlie childbirth, death, migration and the other events that shape the lives of people. This new generation of models consists of transition models or multistate models that capture transitions people make in life and the developmental processes and pathways that characterize individual lives. The multistate life table and the multistate projection model, combined with regression models of transition rates, are adequate candidates for the development of a new generation of demographic projection models. Cox's (1972) paper on regression models for life tables caused a revolution in survival analysis. The emerging discipline of survival analysis was provided with tools it needed for studying the effects of prognostic factors on individual survival in clinical trials designed to evaluate new cancer therapies. The model became a central research tool. The paper by Gill (1992) on regression models for multistate life tables went largely unnoticed although it showed that the Cox model and the associated techniques of statistical inference can be immediately applied to studying transitions in multistate demographic models. It is time to move away from narrow perspectives in demographic forecasting and to broaden the perspective by bridging micro- and macro-level analysis and by effectively integrating substantive knowledge and statistical perspectives and techniques into demographic modeling. The life course provides the logical framework, from a substantive and an analytical perspective. Techniques of statistical inference may be used to obtain parameters of projection models from observational data to complement vital statistics and census data and to provide for a richer empirical basis for demographic forecasts.

Acknowledgement

I would like to thank Nico Keilman for his comments on an earlier draft of this paper.

Notes

1. The concept of cohort biography was introduced by Ryder (1965).

2. In demography, this condition is referred to as the "pure state".

3. The intensity is a conditional probability of a move during a small interval: $\lambda_{ij}(x) = \lim \frac{P(x \leq X > x + \Delta x, J = j \mid X \geq x, I = i)}{\Delta x}$ where X, I and J are random variables denoting age, state of origin and state of destination, respectively.

4. TDA (Transition Data Analysis) has a facility for user-defined rate models (Rohwer and Pötter, 1999, Section 6.17.5). The programme may be downloaded from prof. Rohwer's homepage: http://www.stat.ruhr-uni-bochum.de/ The manual (extensive) can be downloaded from the same site. Willekens (2002) has written a brief introduction to TDA with examples.

References

Ahlburg, D.A., Lutz, W. and Vaupel, J.W. (1999) "Ways to Improve Population Forecasting: What Should be Done Differently in the Future?" In Lutz, W., Vaupel, J.W. and Ahlburg, D.A. (eds) *Frontiers of Population Forecasting* A Supplement to Vol. 24, 1998, *Population and Development Review*. New York: The Population Council:191-198

Alho, J.M. and Spencer, B.D. (1997) "The Practical Specification of the Expected Error of Population Forecasts". *Journal of Official Statistics.* PLEASE PROVIDE DETAILS OF VOLUME AND PAGE NUMBERS

Alho, J. (n.d.) "PEP – A program for error propagation". Available from internet: http://joyx.joensuu.fi/~ek/pep/userpep.htm

Anderson, K.N., Odell, P.M., Wilson, P.W.F. and Kannel, W.B. (1990) "Cardiovascular Disease Risk Profiles". *American Heart Journal*, 121:293-298.

Aoki, M. (1976) *Optimal Control and System Theory in Dynamic Economic Analysis*, New York: North Holland.

Barker, D.J.P. (1998) *Mothers, Babies and Health in Later Life*, Edinburgh: Churchill Livingstone.

Beiser, A., D'Agostino, R.B., Seshadri, S., Sullivan, L.M. and Wolf, P. (2000) "Computing Estimates of Incidence, Including Lifetime Risk: Alzheimer's Diseases in the Framingham Study. The Practical Incidence Estimators (PIE) Macro". *Statistics in Medicine* 19:1495-1522

Ben-Shlomo, Y. and Kuh, D.L. (2002) "A Life Course Approach to Chronic Disease Epidemiology: Conceptual Models, Empirical Challenges and Interdisciplinary Perspectives". *International Journal of Epidemiology*, 31:285-293.

Blossfeld, H.P. and Rohwer, G. (2002) *Techniques of Event History Modeling. New Approaches to Causal Analysis*, New Jersey: Lawrence Erlbaum, Mawah. Second Edition.

Bogue D.J., Arriaga, E.E. and Anderton, D.L. (eds) (1993) *Readings in Population Research Methodology*. Social Development Center, Chicago, and UNFPA, New York.

Breslow, N.E. and Day, N.E. (1987) "Statistical Methods in Cancer Research. International Agency for Research on Cancer". *Scientific Publication 82*, Lyon.

Chiang, C.L. (1984) *The Life Table and its Applications*, Malabar, Fl: R.E. Krieger Publishing.

Çinlar, E. (1975) *Introduction to Stochastic Processes*, New Jersey: Prentice-Hall, Englewood Cliffs.

Commenges D. (1999) "Multi-state Models in Epidemiology." *Lifetime Data Analysis* 5:315-327.

Cox, D.R. (1972) "Regression Models and Life-tables (with discussion)". *Journal of the Royal Statistical Society (B)* 34:187-220.

Crimmins, E.M., Hayward, M.D. and Saito, Y. (1994) Changing Mortality and Morbidity Rates and the Health Status and Life Expectancy of the Older Population". *Demography* 31(1):159-75.

De Backer, G. and de Bacquer, D. (1999) "Lifetime-Risk Prediction: A Complicated Business". *The Lancet* 353:82.

Elbaz, A., Bower, J.H., Maraganore, D.M., McDonnell, S.K., Peterson, B.J., Ahlskog, J.E., Schaid, D.J. and Rocca, W.A. (2002) "Risk Tables for Parkinsonism and Parkinson's Disease". *Journal of Clinical Epidemiology* 55:25-31.

Elder, G.H. Jr. (1985) *Life Course Dynamics: Trajectories and Transitions, 1968-1980*, Ithaca, New York: Cornell University Press. Extracts from Chapter 1 reprinted as "Perspectives on the Life Course" in Bogue, D.J. Arriaga, E.E. and Anderton, D.L. (eds), *Readings in Population Research Methodology*, Chicago: Social Development Center, and New York: UNFPA.

______ (1999) "The Life Course and Aging; Some Reflections". *Distinguished Scholar Lecture*, American Sociological Association, August 1999.

Flinn, C. and Heckman, J. (1982a) "New Methods for Analyzing Structural Models of Labor Force Dynamics". *Journal of Econometrics* 18:115-168.

______ (1982b) "Models for the Analysis of Labor Market Dynamics". In Bassmann, R. and Rhodes, G. (eds) *Advances in Econometrics* 3, Greenwich, Conn: JAI Press.

Galler, H.P. (1997) "Discrete-Time and Continuous-Time Approaches to Dynamic Microsimulaiton Reconsidered". *Technical Paper 13*, National Centre for Social and Economic Modeling (NATSEM), University of Canberra. http://www.natsem.canberra.edu.au/pubs/tps/tp13/tp13.pdf

Giele, J.Z. and Elder, G.H. Jr. (1998) "Life Course Research. Development of a Field". In: Giele. J.Z. and Elder, G.H. Jr. (eds) *Methods of Life Course Research. Qualitative and Quantitative Approaches*, Thousand Oaks, Ca: Sage Publications.

Gill, R.D. (1992) "Multistate Life Tables and Regression Models". *Mathematical Population Studies* 3(4):259-276.

Halfon, N. and Hochstein, M. (2002) "Life Course Health Development: An Integrated Framework for Developing Health, Policy, and Research". *The Milbank Quarterly* 80(3). Available at http://www/milbank.org/quarterly/8003feat.html

Hachen, D.S. (1988) "The Competing Risk Model". *Sociological Methods and Research* 17(1):21-54. Reprinted in Bogue, D.J., Arriaga, E.E. and Anderton, D.L. (eds) *Readings in Population Research Methodology*. Social Development Center, Chicago, and UNFPA, New York. 21.85-21.101.

Hoem, J.M. and Jensen, U.F. (1982) "Multistate Life Table Methodology: A Probabilist Critique". In Land, K.C. and Rogers, A. (eds). *Multi-dimensional Mathematical Demography*. New York: Academic Press.

Hougaard, P. (1999) "Multi-State Models: A Review". *Lifetime Data Analysis* 5(3):239-264.

______ (2000) *Analysis of Multivariate Survival Data*. New York: Springer Verlag,.

Keyfitz, N. (1982) "Can Knowledge Improve Forecasts?". *Population and Development Review* 8(4):729-751.

______ (1985) "The Multistate Model". *Applied Mathematical Demography*. Second Edition. New York: Springer Verlag. Chapter 12. (pp.350-367).

Kono, S (1993) "Functional Population Projections. Editor's Introduction". In Bogue, D.J., Arriaga, E.E. and Anderton, D.L. (eds) (1993) *Readings in Population Research Methodology*. Chicago: Social Development Center, and New York: UNFPA, pp. 18.1–18.2.

Kuh, D.L. and Ben-Shlomo, Y. (1997) *A Life Course Approach to Chronic Disease Epidemiology. Tracing Origins of Ill-Health from Early to Adult Life*. Oxford: Oxford University Press.

Kuh. D.L. and Hardy, R. (2002) *A Life Course Approach to Women's Health*. Oxford: Oxford University Press.

Lee, R.D. (1998) "Probabilistic Approaches to Population Forecasting". *Population and Development Review. Supplement to Volume 24* 165-190.

Légaré, J. and Marcil-Gratton, N. (1990) "Individual Programming of Life Events: A Challenge for Demographers in the Twenty-First Century". In Maltoni, C. and Selikoff, I.J. (eds). Scientific Issues of the Next Century: Convocation of World Academies. *Annals of the New York Academy of Sciences*, 610: 99-105.

Liang, Z. (2000) *The Coale-McNeil Model. Theory, Generalization and Application*, Amsterdam: Thela Thesis.

Liaw, K.-L. and Ledent, J. (1980) "Discrete Approximation of a Continuous Model of Multistate Demography". *Professional Paper 80-14*, Laxenburg, Austria: International Institute for Applied Systems Analysis.

Lloyd-Jones, D.M., Larson, M.G., Beiser, A. and Levy, D. (1999) "Lifetime Risk of Developing Coronary Heart Disease". *The Lancet* 353:89-92 (see also discussion in *The Lancet* 353:924-925).

Lutz, W., Sanderson, W.C. and Scherbov, S. (1997) "Doubling of World Population Unlikely". *Nature* 387:803-805.

Mamun, A.A. (2004) *Life History of Cardiovascular Disease and its Risk Factors*, Amsterdam: Rozenberg Publishers.

Manton, K.G. and Stallard, E. (1988) *Chronic Disease Modeling: Measurement and Evaluation of the Risks of Chronic Disease Processes*, New York: Oxford University Press.

Manton, K.G., Singer, B.H. and Suzman, R.M. (eds) (1993) *Forecasting the Health of the Elderly Population.*, New York: Springer Verlag.

Mathers, C.D. and Robine, J.M. (1997) "How Good is Sullivan's Method for Monitoring Changes in Population Health Expectancies?". *Journal of Epidemiology and Community Health* 51(1):80-86.

Namboodiri, K. and Suchindran, C.M. (1987) *Life Table Techniques and their Applications*, Orlando: Academic Press.

Newman, S.C. (1988) "A Markov Process Interpretation of Sullivan's Index of Morbidity and Mortality". *Statistics in Medicine* 7:787-794.

O'Donoghue, C. (n.d.) "Dynamic Microsimulation. A Methodological Survey". http://www.beje.decon.ufpe.br/v4n2/cathal.pdf

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

______ (1995) *Multiregional Demography. Principles, Methods and Extensions*, Chichester: Wiley.

Rogers, A. and Castro, L. (1986) "Migration". In Rogers, A. and Willekens, F. (eds) *Migration and Settlement. A Multiregional Comparative Study*, Dordrecht, The Netherlands: D. Reidel Publishing Company pp. 157-208.

Rogers, A. and Willekens, F.J. (1986) "A Short Course in Multiregional Mathematical Demography". In Rogers, A. and Willekens, F.J. (eds) *Migration and Settlement. A Multiregional Comparative Study*, Dordrecht: Reidel Publishing Company, 355-384.

Rohwer, G. and Pötter, U. (1999) *TDA User's Manual.* Ruhr-Universität Bochum. Fakultät für Sozialwissenschaften, Bochum, Germany.

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

Schoen, R. (1988a) *Modeling Multigroup Populations*, New York: Plenum Press.

Schouten, L.J., Straatman, H., Kiemeney, L.A.L.M. and Verbeek, A.L.M. (1994) "Cancer Incidence: Life Table Risk versus Cumulative Risk". *Journal of Epidemiology and Community Health*, 48:596-600

Shieh, L.S., Yates, R. and Navarro, J. (1978) "Representation of Continuous Time State Equations by Discrete-Time Equations". *IEEE* 8(6):485-492.

Singer, B. and Spilerman, S. (1979) "Mathematical Representations of Development Theories". In Nesselroade, J.R. and Baltes, P.B. (eds) *Longitudinal Research in the Study of Behavior and Development*, New York: Academic Press, 155-177.

Smith, S. (1982) *Tables of Working Life: The Increment-Decrement Model.* Washington: U.S. Government Printing Service for U.S. Department of Labour, Bureau of Labor Statistics.

Strang, G. (1980) *Linear Algebra and its Applications*, Second edition, New York: Academic Press.

Tuma, N. and Hannan, M.T. (1984) *Social Dynamics: Models and Methods.* New York: Academic Press.

Van Imhoff, E. (1990) "The Exponential Multidimensional Demographic Projection Model. *Mathematical Population Studies* 2(3):171-182.

Van Imhoff, E. and Post, W. (1998) "Microsimulation Methods for Population Projection". *Population: An English Selection*, special issue *New Methodological Approaches in the Social Sciences*, 97-138.

Van Imhoff, E. and Keilman, N. (1991) *LIPRO 2.0: An Application of a Dynamic Demographic Projection Model to Household Structure in the Netherlands.* Amsterdam: Sweets&Zeitlinger.

Willekens, F.J. (1995) MUDEA. *Version 2.0. Manual and Tutoria*l. Faculty of Spatial Sciences, University of Groningen.

______ (1998) "Demographic Projection Models: A Technical Introduction". Manuscript.

______ (2002) "Forecasting the Life Course". Paper presented at the *Annual Meeting of the Population Association of America*, Atlanta.

______ (2004) *Biographies. Real and Synthetic.* Manuscript.

Willekens, F.J. and Hakkert, R. (1992) *Directory of Demographic Software.* Outcome of the IUSSP Working Group on Demographic Software and Microcomputing), published in 1998 on the internet: http://www.unfpacst.cl (site United Nations Population Fund, CELADE, Chili).

Willekens, F.J. and Rogers, A. (1978) "Spatial Population Analysis. Methods and Computer Programs". IIASA, *Research Report RR*-78-18.

Willekens, F.J. and Drewe, P. (1984) "A Multiregional Model for Regional Demographic Projection". In ter Heide, H. and Willekens, F. (eds) *Demographic Research and Spatial Policy.* London: Academic Press, 309-334.

WHO (World Health Organization) (2002) *Life Course Perspectives on Coronary Heart Disease, Stroke and Diabetes. The Evidence and Implications for Policy and Research.* WHO/NMH/NPH/02.1, Department of Noncommunicable Disease Prevention and Health Promotion, World Health Organization, Geneva.

Wolf, D. (2000) *The Role of Microsimulation in Longitudinal Data Analysis.* http://www.ssc.uwo.ca/sociology/longitudinal/wolf.pdf

Wun , L-M., Merrill, R.M. and Feuer, E.J. (1998) "Estimating Lifetime and Age-Conditional Probabilities of Developing Cancer". *Lifetime Data Analysis*, 4(2):169-186.

Yamaguchi, K. (1991) *Event History Analysis*, Newbury Park: Sage Publications,.

Zeng Yi (1991) *Family Dynamics in China. A Multistate Life Table Analysis*, Wisconsin University Press, Madison, WI.

Zeng Yi, Vaupel, J.W. and Zhenglian, W. (1997) "A Multidimensional Model for Projecting Family Households. With an Illustrative Numerical Application". *Mathematical Population Studies* 6(3):187-216.

Annex I
The linear model as an approximation of the exponential model

Using Taylor series expansion, it can be shown that the linear model is an approximation of the exponential model. Two methods are considered

i. Method 1

The exponent $\exp[-h\mathbf{M}]$ can be written as

$$\exp[-h\mathbf{M}] = 1 - h\mathbf{M} + \tfrac{1}{2}(h\mathbf{M})^2 - \tfrac{1}{6}(h\mathbf{M})^3 + \ldots = \sum\nolimits_{k=0}^{\infty} \frac{-(h\mathbf{M})^k}{k!}$$

The geometric progression of $[\mathbf{I} + \tfrac{1}{2}h\mathbf{M}]^{-1}$ is

$$[\mathbf{I} + \tfrac{1}{2}h\mathbf{M}]^{-1} = \mathbf{I} - \tfrac{1}{2}h\mathbf{M} + \tfrac{1}{4}(h\mathbf{M})^2 - \tfrac{1}{8}(h\mathbf{M})^3 + \ldots. \text{ provided that } \left|\tfrac{1}{2}h\mathbf{M}\right| < 1$$

$$[\mathbf{I} + \tfrac{1}{2}h\mathbf{M}]^{-1}[\mathbf{I} - \tfrac{1}{2}h\mathbf{M}] = \mathbf{I} - h\mathbf{M} + \tfrac{1}{2}(h\mathbf{M})^2 - \tfrac{1}{4}(h\mathbf{M})^3 + \tfrac{1}{8}(h\mathbf{M})^4 - \ldots. = \sum\nolimits_{k=0}^{\infty} \frac{-(h\mathbf{M})^k}{(k-1)!}$$

with $(-1)! = 1$

ii. Method 2

Liaw and Ledent (1980) show that a method developed in engineering for the discrete approximation of continuous-time state equations may be applied to show the relation between the exponential model and the linear model. It is the Matrix Continued Fraction (MCF) method developed by Shieh *et.al.* (1978). To make the MCF method transparent, Liaw and Ledent consider the expansion of the number 1.2345 into a continued fraction:

$$1.2345 = 1 + \frac{2345}{10000} = 1 + \frac{1}{100000/2345} = 1 + \frac{1}{4 + \frac{620}{2345}} = \ldots.$$

After a few divisions, one gets

$$1.2345 = 1 + \cfrac{1}{4 + \cfrac{1}{3 + \cfrac{1}{1 + \cfrac{1}{3 + \ldots}}}} = H_1 + \left[H_2 + \left[H_3 + \left[H_4 + \ldots\right]^{-1}\right]^{-1}\right]^{-1}$$

The retention of the first few H_j results in a fairly good approximation of the original number. For instance, the retention of the first three H_j gives the number 1.2308.

Application of the MCF method to approximate $\exp[-(y-x)\mathbf{M}(x,y)]$ gives

$$\exp[-(y-x)\,\mathbf{M}(x,y)] = \mathbf{H}_1 + \left[\mathbf{H}_2 + \left[\mathbf{H}_3 + [\mathbf{H}_4 + \dots\ ]^{-1}\right]^{-1}\right]$$

Shieh *et al.* (1978) show that

$$\mathbf{H}_1 = \mathbf{I}\,,\ \mathbf{H}_2 = [-(y-x)\mathbf{M}(x,y)]^{-1}\,,\ \mathbf{H}_3 = -2\mathbf{I}\,,\ \mathbf{H}_4 = [-3(y-x)\mathbf{M}(x,y)]^{-1}\,,\ \mathbf{H}_5 = 2\mathbf{I}$$

Let $\mathbf{G}_j$ be the estimate of $\exp[-(y-x)\mathbf{M}(x,y)]$ by retaining only the first j $\mathbf{H}$ matrices. Then we get

$$\mathbf{G}_2 = \mathbf{H}_1 + [\mathbf{H}_2]^{-1} = \mathbf{I} - (y-x)\mathbf{M}(x,y)$$

$$\mathbf{G}_3 = \mathbf{H}_1 + \left[\mathbf{H}_2 + [\mathbf{H}_3]^{-1}\right]^{-1} = [\mathbf{H}_2\mathbf{H}_3 + \mathbf{I}]^{-1}\,[\mathbf{H}_1\mathbf{H}_2\mathbf{H}_3 + \mathbf{H}_1 + \mathbf{H}_3]$$

$$\mathbf{G}_3 = \left[\mathbf{I} + \frac{y-x}{2}\mathbf{M}(x,y)\right]^{-1}\left[\mathbf{I} - \frac{y-x}{2}\mathbf{M}(x,y)\right]$$

which is the linear approximation of the exponential model. The linear model is therefore obtained by retention of the first three $\mathbf{H}_j$ in the MCF method.

There is a difference between ignoring the higher $\mathbf{H}_j$ in the MCF method and disregarding the tail of the Taylor series expansion. Shieh *et al.* (1978) observe that

$$\mathbf{G}_3 = \mathbf{I} + [-(y-x)\mathbf{M}(x,y)] + \frac{1}{2!}[-(y-x)\mathbf{M}(x,y)]^2 + \sum_{j=3}^{\infty}\frac{1}{2^{j-1}}[-(y-x)\mathbf{M}(x,y)]^j$$

which differs from the Taylor series expansion.