

Chapter 15

Continuous-time Microsimulation in Longitudinal Analysis

Frans Willekens

1 Introduction

In longitudinal analysis, individuals are followed in time and are observed either continuously or at points in time. The time to event, the sequence of events and the factors that influence timing and sequence constitute the object of study. Time is a continuous variable (exact time) or discrete variable (time interval). Longitudinal data are used to estimate parameters of event history or life history models. Individual life histories can be represented by sequences of states and sequences of events, that are transitions between states, and described by multistate transition models. The parameters of these models are transition intensities when time is represented by a continuous variable and transition probabilities when time is discrete (for details, see Willekens, 2001). Microsimulation contributes to longitudinal data analysis in a number of ways (Wolf, 1986, 2001). First, it generates individual event histories that are fully consistent with a set of transition intensities (probabilities). Second, it produces estimates of the full distribution of an outcome, in addition to the expected value that is produced analytically by most models. Third, it is helpful in examining the potential seriousness of defective data. Fourth, it may play a role in the imputation of missing

353

The research is part of the MicMac project, an international (European) project funded by the European Commission in the context of the Sixth Framework Programme. For details, see the MicMac website www.micmac-projections.org. I like to thank Dr. Jutta Gampe, Max Planck Institute for Demographic Research, Rostock, for extensive comments on an earlier draft. I also like to thank Douglas Wolf, a referee and the editors for their comments.

data. These contributions are enhanced when microsimulation is viewed as a form of sampling of a virtual population, an approach advocated by e.g. Wolf (2001) and adopted in this chapter.

Microsimulation in continuous time resolves three important problems of discrete-time microsimulation. The first is how to determine the sequence of events (transitions) rather than the state occupancies at successive points in time. The second is the precise measurement of the lengths of episodes between events. In discrete-time microsimulation the duration between events can be determined only approximately, whereas it can be determined precisely in continuous-time microsimulation. The third is how to handle multiple transitions during a same interval. In discrete-time microsimulation, multiple transitions during an interval are either omitted or assumptions about the ordering and the timing of the events are imposed exogenously. **In continuous-time microsimulation, the theory of competing risks determines the timing and sequence.** That allows a more accurate study of temporal sequence of events than in discrete time analyses. In addition the theory allows to model complex event sequences and interactions between events. The fact that microsimulation models in discrete time are not able to handle complex and interdependent event sequences is viewed as an important limitation (Zaidi and Rake, 2001: 19). The fourth problem that continuous-time microsimulation resolves is related to the third. It is the estimation of the number of events during an interval. In addition to resolving these problems, continuous-time microsimulation paves the way to an integrative framework that combines the *analysis* of life history data and the *synthesis* of life histories. That framework is rooted in probability theory and statistical theory and is extensively documented in the literature on survival analysis and event history analysis (see e.g. Blossfeld and Rohwer, 2002; Klein and Moeschberger, 2003; Andersen and Keiding, 2002; Putter et al., 2007; Meira-Machado et al., 2009). By adopting the established framework, microsimulation is embedded in event history analysis, which is an aim worth pursuing (see also Wolf, 1986 and Galler, 1997).

354

In event history models (also known as duration models, survival models and transition models) the variables of interest are the (waiting) time to event and sequences of events and states occupied. The time to event is a random variable. The distribution of the values of the random variable is described by the distribution function, the survival function, the probability density function and the hazard function. The *inverse distribution function* or *quantile function* is of prime importance in continuous-time

microsimulation. The quantile function translates a probability into a real number, whereas the more commonly used distribution function and survival function translate a real number into a probability. The real number is the (waiting) time to event.

The chapter is organized as follows. Section 2 describes the method for generating waiting times to events from duration models. The approach is to use the inverse distribution function of the duration model. Once the function is specified, the generation of waiting times is straightforward. Section 3 presents numerical illustrations. Three illustrations are considered. The first is a simple one: a single event during a period of a given duration. The second illustration is a full multistate transition model with three states: healthy, disabled and dead. The multistate model is a continuous-time Markov model. That model, combined with a random generation of waiting times to transition, produces lifepaths for members of a virtual population. In this illustration the virtual population is homogeneous. The third illustration considers a heterogeneous population. Individuals are characterized by covariates. Some members of the virtual population participate in an intervention programme that includes prevention and treatment. The illustration expresses the nature and level of intervention in terms of the transition rates and assesses the impact of the prevention and treatment programmes on the probability of disability, the time at onset of disability and the numbers of years with disability. Section 4 concludes the chapter. Continuous-time microsimulation is implemented in a number of simulation models. They are listed in Annex A.

2 Inverse distribution or quantile function

The quantile function translates a probability into a real number. In the context of dynamic microsimulation it translates a probability of a transition into a (waiting) time to transition. This section presents the general method and applications to a few well-known transition models. Let U denote a random variable following a uniform distribution on the interval from 0 to 1: $U \sim U[0,1]$. Let T be a random variable measuring the time to an event or transition, where transition refers to a direct transition from an origin state to a destination state. The distribution of T is $F(t)$ with $F(t)$ the probability that T is less than or equal to t , i.e. the probability that the transition occurs in the interval from 0 to t . The survival function is $S(t) = 1 - F(t)$. At

time t , i.e. during the interval from t to $t+dt$, the transition occurs at a rate

$\mu(t) = -\frac{1}{S(t)} \frac{dS(t)}{dt}$. The rate is the hazard rate or transition rate. Hence

$$S(t) = \exp\left[-\int_0^t \mu(\tau) d\tau\right] = \exp[-H(t)]$$

where $H(t)$ is the cumulative hazard function.

The distribution function maps a real number (a particular value t of the random variable T) into a probability. The real number that is mapped into a probability is the *quantile* of the random variable (see Evans et al., 2000: 5). Hence, t is the quantile of T . With t is associated a probability, α say, and the distribution function gives the probability that T does not exceed t . The inverse distribution function or *quantile function* maps a probability (α) into a real number (t). In other words it maps a realization of U (denoted by α) into a realization of T (denoted by t). The inverse distribution function of T , denoted by $F^{-1}(t)$ and $G(\alpha)$, is the value of t (quantile) such that the probability that T takes on a value less than or equal to t , is α : $\Pr\{T \leq G(\alpha)\} = \alpha = F(G(\alpha)) = F(t)$. $G(\alpha)$ gives the value t for which $F(t) = \alpha$. The quantile $G(\alpha)$ is the 100α percentile. The inverse survival function $Z(\alpha)$ is the quantile that is exceeded with probability α : $\Pr\{T > Z(\alpha)\} = \alpha = S(Z(\alpha))$. Inverse distribution functions are widely used in statistics, for instance to determine confidence intervals. Note that $Z(\alpha) = G(1-\alpha)$.

356

The inverse distribution function is used to generate random numbers from the distribution of a random variable. If $G(T)$ is the inverse distribution function of T , then $U = G(T)$ follows a uniform distribution on the interval from 0 to 1. In a Monte Carlo microsimulation a random draw from a distribution function involves two steps. First, a random value for the probability α is drawn from the uniform distribution $U[0,1]$ (note that $G(T)$ follows the distribution $U[0,1]$). Second, using the inverse distribution function $G(\alpha)$, the probability is mapped into a real number t , which indicates the timing of the transition. The first step is independent of the transition model used.

Common waiting time distributions include the exponential, the Weibull and the Gompertz distribution. If the waiting time to transition (T) follows an exponential distribution the transition occurs at a constant rate, i.e. the transition rate is constant. If the waiting time distribution is Gompertz, the transition rate changes exponentially with duration. If it is Weibull, the transition rate varies with duration following a power function

of duration. The exponential distribution is used in most continuous-time dynamic microsimulation models, e.g. DYNAMOD, SOCSIM, LifePaths and PENSIM. In this section I consider the quantile function of three widely used transition rate models: the exponential model, the Gompertz model and the Cox model.

The exponential distribution is thoroughly documented by Balakrishnan and Basu (1996) and Evans et al. (2000). If T is exponentially distributed, the transition rate is constant. Let μ denote the constant transition rate. The survival function is $S(t) = \exp[-\mu t]$ and the distribution function is $F(t) = 1 - \exp[-\mu t]$. The inverse distribution function of T is $F^{-1}(t) = G(\alpha) = -\frac{\ln[1 - \alpha]}{\mu}$.

For a given transition rate (and the exponential model) it translates a transition probability into a waiting time to transition. With a draw α from a the uniform distribution $U[0,1]$ is associated a time to event $t = G(\alpha)$.

In this section on theory, two distributions are considered that are not applied in this chapter: the Gompertz distribution and the Cox model. The Gompertz distribution of waiting times has two parameters, a scale parameter (μ) and a shape parameter (ν). The transition rate changes exponentially; $r(t) = \mu \exp(\nu t)$ (with $\mu \geq 0$). If $\nu = 0$, the Gompertz distribution reduces to the exponential distribution. The survival function is

$$S(t) = \exp\left[\frac{\mu}{\nu} (1 - \exp(\nu t))\right] \text{ and the distribution function is } 1 - S(t).$$

The quantile function is

$$T = \frac{1}{\nu} \left[\ln\left(1 - \frac{\nu}{\mu} \ln(1 - U)\right) \right]$$

where $U \sim U[0,1]$ is a random variable the values of which are uniformly distributed in the range from 0 to 1. A random draw from a Gompertz distribution is obtained in two steps, described above. First, a random number α is drawn from a uniform distribution over $[0,1]$. Second, the value of t is derived from the quantile function. Mueller et al. (1995: 558) generate a sample of waiting times from a Gompertz distribution.

The Cox proportional hazard model is given by $\mu(t|\mathbf{Z}) = \mu_0(t) \exp[\beta' \mathbf{Z}]$ where t is the time (duration), \mathbf{Z} a vector of covariates, β the vector of regression coefficients and $\mu_0(t)$ the baseline hazard function, which is the hazard function for the group of individuals with characteristics equal to the reference categories of the elements of \mathbf{Z} . For details on the Cox model, see any

textbook on survival analysis or event history analysis. The survival function of the Cox proportional hazard model is $S(t|\mathbf{Z}) = \exp[-H_0(t)\exp(\beta'\mathbf{Z})]$ where $H_0(t)$ is the cumulative hazard function. The distribution function is $1 - S(t|\mathbf{Z})$. The quantile function of the Cox proportional hazard model is (Bender et al., 2005) $F^{-1}(t|\mathbf{Z}) = G(\alpha|\mathbf{Z}) = H_0^{-1}[-\ln(\alpha)\exp(-\beta'\mathbf{Z})]$. The quantile function determines the waiting time to a transition that is consistent with a probability of the transition when the probability depends on a constant transition rate and a set of time-independent covariates. The quantile function translates the transition rate (hazard rate) into a waiting time to transition. The translation involves the two steps listed above. The first step translates the transition rate into a realization of a random variable that follows a uniform distribution and the second step translates that realization into a realization of a waiting time distribution following a specific Cox model. For a discussion and application see Bender et al. (2005). A problem is that the Cox model is a semi-parametric model which leaves the baseline hazard unspecified. As a consequence, the cumulative hazard function H_0 is unknown and the inverse cannot be obtained. To obtain waiting times from the transition rates, the baseline hazard function must be specified. If the baseline hazard function is constant, the waiting times generated are exponentially distributed. Bender et al. refer to this model as the Cox-exponential model. They also discuss the Cox-Gompertz and the Cox-Weibull models. If the baseline hazard function is a Gompertz distribution, the transition rate varies exponentially with duration with a level that depends on the covariates.

358

In longitudinal microsimulation, the entire life course of an individual may be simulated before the simulation of the next individual starts. Waiting times are generated for several competing events and the next event and the time to that event are determined by the smallest waiting time. An alternative is to simulate a segment of life for all individuals before the simulation of the next segment starts. A segment can vary in length from a month to several years. An advantage of this approach is that at the beginning of each segment characteristics of other individuals and the context can be considered in determining transition rates. Segments may refer to windows of observation, periods during which for a given transition rate the duration dependence does not change (e.g. piecewise constant transition rates), and to time periods that are dictated by the application. For instance, demographic projections generally consider periods of one year. Consider duration intervals of length h years. If the waiting time t drawn at random from a waiting time distribution is less than h , the transition occurs during

the observation window at time t . If t exceeds h , the transition does not occur. In case of a repeatable event, multiple transitions may occur during a period of h years. Many life events may occur more than once, i.e. job change, childbirth, marriage, and migration. Suppose an event is repeatable and its first occurrence is at t_1 ($t_1 < h$). The event occurs a second time during the interval if a second draw of a random variable from $U[0,1]$ results in a value of t_2 that is less than or equal to $h - t_1$. In that case, the second event occurs at time $t_1 + t_2$.

The assumption of fixed transition rates is for presentation only. Hazard rates are generally assumed to be piecewise constant, i.e. constant during intervals of a given length, usually one or five years. In that case, the exponential distribution is a step function with parameters that differ between intervals. For a particular interval, the random draws from the particular exponential distribution are kept only if the event time is in the interval.

If the event is a repeatable event, it may occur more than once during an interval of length h . Let t_1 denote the time at first occurrence. The probability that the event occurs a second time during the interval is the probability that it occurs during the interval from t_1 to h , which is of length $h - t_1$. That probability is

$$F(h - t_1) = 1 - \exp\left[-\int_{t_1}^h \mu(\tau) d\tau\right].$$

A random number α is drawn from the uniform distribution $U[0,1]$ and the value of t_2 is determined that is consistent with that random number α . If $t_2 < h - t_1$, the event occurs a second time during the interval h , otherwise it does not. The sequence of events during the interval can be simulated in a similar way.

If t is less than the time till the end of the interval, the event occurs at t . Otherwise the event does not occur.

3 Illustration

Recall that a transition is an event that can be characterized by the time at occurrence, the origin state and the destination state. The origin state is denoted by i , the destination state by j and the waiting time to transition by the random variable T_{ij} . Let time be measured in years and fractions of a year. The rate of transition is μ_{ij} . Assume that the transitions occur in continuous time. Observations on times to transitions are manifestations

of T_{ij} and the observed times to transitions are the basis for the statistical estimation of the transition rates. A discussion of the estimation is beyond the scope of this chapter. For methods, see e.g. Blossfeld and Rohwer (2002). The time to event T is generally analysed in terms of its distribution function $F(t)$ and the associated survival function $S(t)$, density function $f(t)$, hazard function $\mu(t)$, quantile function $G(\alpha)$, and expected values of key indicators such as the expected waiting time to a transition $E[T]$. Note that the named distributions are equivalent and can be transformed one into another. The transition rates estimated from the data may be used to generate life histories that are consistent with the transition rates, using the quantile function. This is equivalent to sampling a virtual population with life histories governed by the empirical transition rates. Expected values serve as benchmarks to assess the results of the microsimulation. For large sample sizes the sample values should coincide with the expected values. This section consists of three subsections. Section 3.1 presents a transition with a single origin and a single destination. Multiple origins and multiple destinations are considered in Section 3.2. The aim is to determine the sequence and dates (times) of transition during the interval from 0 to h and to derive subsequently the length of episodes between transitions. The waiting time distributions are truncated at h . Conditional measures such as the expected waiting time to the event, provided the event occurs before h , are based on the truncated distribution. Section 3.3 introduces covariates and applies a transition model and continuous-time microsimulation to assess the effects of an intervention programme on the life histories of members of the virtual population.

3.1 *Single origin and single destination*

A simple model of T is the basic exponential transition rate model (see e.g. Blossfeld and Rohwer, 2002, Chapter 4). The exponential distribution has a single parameter that is independent of time or duration. For presentation, I assume a transition rate of $\mu = 0.2$.

The probability of a transition within a period of time is given by the distribution function. The probability of a transition within a year is $F(1) = 1 - \exp(-0.2) = 0.18$, i.e. 18%. The expected waiting time to a transition is $E[T] = 1/0.2 = 5$ years. It is

$$E[T] = \int_0^{\infty} \tau \exp(-\mu \tau) d\tau = \int_0^{\infty} \tau \exp(-0.2\tau) d\tau = \left| -\frac{1}{0.2} \exp(-0.2\tau) \right|_0^{\infty} = \frac{1}{0.2} = 5$$

The median waiting time to the transition, i.e. the time at which there is a 50% chance that the event occurred, is given by the inverse distribu-

tion function: $G(0.5) = -\frac{\ln[1-0.5]}{0.2} = \frac{\ln(2)}{0.2} = 3.5$ years. The probability

of a transition reaches 25% at the upper quartile $[G(0.25)]$ which is 1.44 years; the probability of a transition reaches 75% at the lower quartile $[G(0.75)]$ which is 6.9 years. If the transition rate is 0.2, there is a 90% probability that the transition occurs before 11.5 years. Note that the probability that the transition occurs before the expected waiting time is 63.2%. The probability that no transition occurs between the start of the interval and τ ($0 \leq \tau \leq h$) is the survival function $S(\tau) = \exp[-0.2 \tau]$.

The expected waiting time to transition during a period of h years is the total time expected to be spent in the origin state during h years. It is

$$E[T] = \int_0^h \exp(-\mu \tau) d\tau = \int_0^h \exp(-0.2\tau) d\tau = -\frac{h}{0.2} \exp(-0.2\tau) \Big|_0^h = -\frac{h}{0.2} [\exp(-0.2h) - 1]$$

361

It is 0.91 if $h = 1$. The value is relatively high because the probability of experiencing the event during a year is relatively low and the waiting time is one year for persons who do not experience the event during the period of one year.

The expected time spent in the origin state during a period of h years is a weighted average of the expected sojourn time in the presence of a transition and the sojourn time in the absence of a transition: $E_e[T] = h * S(1) + E_e[T] [1 - S(1)]$ where $E_e[T]$ is the expected waiting time to the transition provided the transition occurs during the interval. The sojourn time in the absence of the transition is h years. The expected sojourn time provided the transition occurs is

$$E_e[T] = \frac{E[T] - h S(1)}{1 - S(1)} \text{ years. If } h = 1, E_e[T] = 0.48. \text{ Under the exponential}$$

model, the transitions are concentrated in the first half of the year. If a transition occurs during a year, the probability that it occurs in the first half of the year is

$$F(0.5)/F(1) = [1 - \exp(-\frac{1}{2}0.1)] / [1 - \exp(-0.1)] = 0.52 \text{ , which is 52\%}.$$

If the transition is repeatable, it may occur more than once during a period of h years. Suppose the rate is fixed at 0.2. The probability of at least two occurrences within a year is 1.75% and the probability of at least three occurrences is 0.11%. The probability of no occurrence during a period of

one year is the survival function $\exp(-\mu) = \exp(-0.2) = 0.82$ or 82%. The probability of precisely one occurrence is 16.37% and the probability of exactly two occurrences is 1.64%. The probability of precisely $n(t)$ occurrences during the period from 0 to t is given by the Poisson distribution:

$$\Pr\{N(t) = n(t)\} = \frac{(\mu t)^{n(t)} \exp[-\mu t]}{n(t)!}$$

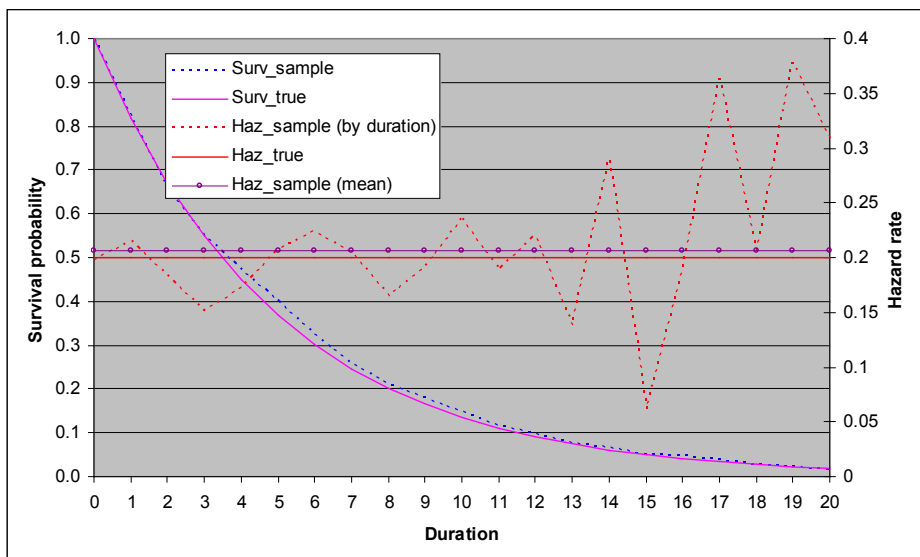
where $n(t)!$ is factorial $n(t)$ which is the product $1 * 2 * 3 \dots * n(t)$.

The time intervals between occurrences are independent and exponentially distributed. Let D_n denote the duration between the $n-1^{\text{st}}$ and the n -th occurrence. The expected length of the interval between any two occurrences is $1/\mu$ (provided the length of "observation" is not constrained to a period of a given length). The time to the occurrence of rank n , T_n , is the sum of independent exponentially distributed variables D_n and is a gamma-distributed random variable. Hence the time to the n -th occurrence is the gamma distribution with parameters μ and n .

362

Consider a random sample of 1,000 subjects from the virtual population that experiences a single, non-repeatable transition at a constant rate of 0.2 per year. The (waiting) time to transition is drawn from an exponential distribution with parameters $\mu = 0.2$, using the method described above. Figure 1 shows the true survival and hazard distributions and the 'empirical' distributions based on the sample of the virtual population.

Figure 1: The exponential model ($\mu = 0.2$)



The “empirical” survival distribution (Surv_emp) based on simulated waiting time data is close to the true distribution as expected. For each duration the transition rate is calculated by dividing the number of transitions in the virtual population of 1,000 by the number of survivors in the mid-period. These duration-specific transition rates vary erratically around the true value of 0.2, especially at higher durations, which are drawn less frequently than lower durations. The mean transition rate is 0.204 (Haz_estim). It is estimated from the simulated waiting time data using the basic exponential transition rate model described by Blossfeld and Rohwer (2002, Chapter 4).

Suppose the transition is repeatable and we want to determine for each subject the number of occurrences during a period of one year ($h=1$). Each subject in the virtual sample is followed (observed) for a period of one year and transitions are recorded. For each transition that occurs, the time to transition and the rank of the transition are recorded. Statistical measures are calculated from the sample and the sample values are compared to expected values that are based on the theoretical distribution. Table 1 shows the results for three samples of size 1,000. The expected values are derived from the exponential distribution and the Poisson distribution, whatever distribution applies. In the first sample, 829 subjects do not experience a transition during the year and 171 experience at least one transition. Most of the subjects who experience at least one transition experience a single transition (152), 18 experience 2 transitions and 1 experiences 3 transitions. The total number of transitions experienced by the 1,000 subjects during that one year is 191 ($152 + 18 \cdot 2 + 1 \cdot 3$). In the second random sample, 203 subjects experience at least one transition. Of them, 189 experience a single transition, 12 experience 2 transitions and 2 have 3 occurrences during the year. The total number of transitions in that year is 219. The *expected* distribution of the subjects by number of transitions is given in the last column. The expected distribution of number of transitions is derived from the Poisson distribution. The number of transitions that the 1,000 subjects may expect to experience during the interval of one year, given the transition rate of 0.2, is 200 ($=1,000 \cdot 0.2$).

The *expected* time to the first occurrence during the year, provided the transition occurs, is 0.48 years (see above). The expected times to subsequent occurrences are more difficult to obtain analytically. The sample mean of the waiting times to the transition follows directly from the microsimulation (sampling). In Table 1 the values are shown for three samples. Note that in Sample 3, the time to the third occurrence is less than the time to the second occurrence. The reason is that the two subjects that experience three occur-

rences during the year experience the first and the second occurrence earlier than the other subjects. Note also that the difference between the times at two consecutive occurrences does not yield the interval between transitions. To determine the interval, the times to transitions must be conditioned on a next transition. For instance, the time between the first and second transition should be calculated including only subjects that experience a second transition during the interval. These subjects may experience the first transition at different times (usually earlier) than subjects that do not experience a second transition. To illustrate the relation between the times to transitions and the number of transitions, consider Sample 1. The mean waiting time to the first transition is 0.50. The 18 subjects that experience two transitions during the year experience the first transition at 0.45 years, which is earlier than the overall average of 0.50 years. The 15 subjects with a single transition during the year experience the transition at 0.51 years, on average. The subject with three occurrences during the year experiences the occurrences at 0.51 years, 0.60 years and 0.96 years. The observation that subjects with more transitions during a given period experience the first transition earlier than other subjects is a general one.

364

Table 1: Number of occurrences and times to transition. Random samples of 1,000 transitions and expected values

Number of subjects by number of occurrences within a year	Random sample 1	Random sample 2	Random sample 3	Expected values
0	829	797	828	819
1	152	189	153	164
2	18	12	17	16
3	1	2	2	1
4	0	0	0	0
5	0	0	0	0
Total	1000	1000	1000	1000
Total number of occurrences within a year	191	219	193	200
<i>Time to transition</i>				
1	0.504	0.478	0.483	0.483
2	0.672	0.705	0.700	
3	0.960	0.740	0.596	
4	-	-	-	
5	-	-	-	

3.2 Multiple origins and multiple destinations

If an origin state may have several exits or if a transition may result in one of multiple destinations, each exit or destination may be viewed as competing to be the exit or destination. In other words, the exits or destinations represent competing risks. In the presence of multiple destinations, the destination must be determined in addition to the time to transition. The time to event (transition) is an exponential random variable. It follows an exponential distribution with parameter the total exit rate, which is the sum of destination-specific transition rates (see further). The number of persons selecting a particular destination in a set of possible destinations is a multinomial random variable (or binomial in case of 2 competing risks). The competing risk model is generally formulated in terms of latent times to transition (see e.g. Klein and Moeschberger, 2003: 50ff.).¹ Let T_j , $j = 1, 2, \dots, J$ be a random variable denoting the unobservable time to occurrence of the transition to destination j , where J is the total number of destinations. T_j is a latent variable. In the theory of competing risks, observations on transitions consist of (1) the shortest time to transition, i.e. $t = \min(t_1, t_2, \dots, t_j)$ and (2) the destination. The destination is represented by the random variable S . If a transition to state j has taken place at t_j , then t_j is a realization of T and the destination is j ($S = j$). The basic competing risk parameter is the hazard rate for risk j

365

$$\mu_j(t) = \lim_{\Delta t} \frac{P[t \leq T < t + \Delta t, S = j | T \geq t]}{\Delta t}$$

which states that, in the presence of multiple destinations, the hazard rate is the product of the rate of transition in the small interval Δt and the probability that the destination is j , provided that the transition has not occurred before t . The total hazard rate is

$$\mu(t) = \sum_{j=1}^J \mu_j(t)$$

The destination-specific hazard rate at time t may be written as the product of the total hazard rate $\mu(t)$ and a probability $p_j(t)$ that the destination state after the transition is j , provided the transition occurs at time t :

$$\mu_j(t) = \mu(t) p_j(t)$$

In continuous-time microsimulation, two approaches may be distinguished to determine the time to transition and the destination. They are equivalent, however. The first, used in e.g. LifePaths, uses $\mu_{ij}(t)$, i.e. the

1 The approach is also used in the presentations of the LifePaths microsimulation model.

destination-specific hazard rates, and generates waiting times to transition for every possible destination j . The shortest waiting time is selected to determine the actual time to event and the destination. In case an interval of length h is considered, the transition occurs if t_j is in the interval. The second approach uses two random variables: $\mu(t)$ to determine the time to transition and $p_j(t)$ to determine the destination. The first random variable (time to event) is drawn from an exponential distribution to determine the timing of the transition. The destination is determined by a random draw from a uniform distribution $U \sim U[0,1]$. Let the draw be denoted by u . If u is less than $p_1(t)$, the transition is to the first destination, if $p_1(t) \leq u < p_1(t) + p_2(t)$, the transition is to the second destination, if $p_1(t) + p_2(t) \leq u < p_1(t) + p_2(t) + p_3(t)$, the transition is to the third destination, etc.

The competing risk model may easily be extended to a multistate model with hazard rates depending on state of origin and state of destination. The transition rate $\mu_{ij}(t)$ is the rate at which individuals, who occupy state i at time t , make a transition to state j . If the sample population is stratified by state occupied, the transition rates are conditioned on the state of origin, and the multistate model resembles the competing risk model. The transition rate by origin and destination $\mu_{ij}(t)$ may be written as the product of the rate of leaving state i [$\mu_{i+}(t)$] and destination probability conditional on the state of origin [$p_{ij}(t)$] and conditional on leaving. For instance, the rate of a direct transition from state 1 to state 2 may be written as: $\mu_{12} = \mu_{1+} \cdot p_{12}$, where μ_{1+} is the exit rate from 1 and p_{12} is the probability that a subject leaving 1 transits to destination 2.

By way of example suppose the state space consists of three states: healthy (1), disabled (2) and dead (3). At the start of the process being simulated, all subjects are in state 1. The process is simulated for a period of 10 years. Suppose the transition rates are constant and equal to: $\mu_{12} = 0.12$, $\mu_{13} = 0.03$, $\mu_{21} = 0.06$ and $\mu_{23} = 0.06$. The rate of leaving state 1 (μ_{1+}) is 0.15 and the rate of leaving state 2 (μ_{2+}) is 0.12. The transition rates imply that 80% of the subjects leaving state 1 (healthy) move to state 2 (disability) and 20% move to state 3 (dead). Consider a sample of 1,000 subjects in state 1 at the start of the process. The *expected* state occupancies at different years are shown in Table 2. They are calculated by the following equation:

$$\mathbf{K}(t+1) = \mathbf{P} \mathbf{K}(t) = \exp[-\mathbf{M}] \mathbf{K}(t)$$

where $\mathbf{K}(t)$ is a vector of state occupancies indicating the number of subjects in each of the three states at time t . \mathbf{P} is the matrix of transition probabilities. In this illustration, they are estimated from the transition rates using

the linear approximation of the exponential model (for the derivation, see Wilkens, 2006): $\exp[-\mathbf{M}] \approx [\mathbf{I} + \frac{1}{2}\mathbf{M}]^{-1} [\mathbf{I} - \frac{1}{2}\mathbf{M}]$

The transition rates are assembled in the transition matrix \mathbf{M} :

$$\mathbf{M} = \begin{bmatrix} \mu_{11} & -\mu_{21} & -\mu_{31} \\ -\mu_{12} & \mu_{22} & -\mu_{32} \\ -\mu_{13} & -\mu_{23} & \mu_{33} \end{bmatrix} = \begin{bmatrix} 0.15 & -0.06 & 0 \\ -0.12 & 0.12 & 0 \\ -0.03 & -0.06 & 0 \end{bmatrix}$$

where $\mu_{ii} = \sum_{j=1}^J \mu_{ij}$. The matrix of transition probabilities for one-year intervals is

$$\mathbf{P} = \begin{bmatrix} p_{11} & p_{21} & p_{31} \\ p_{12} & p_{22} & p_{32} \\ p_{13} & p_{13} & p_{33} \end{bmatrix} = \begin{bmatrix} 0.863 & 0.053 & 0.000 \\ 0.105 & 0.890 & 0.000 \\ 0.031 & 0.057 & 1.000 \end{bmatrix}$$

The expected state occupancies at the beginning of each year from 0 to 10 are given in Table 2.

Table 2: State occupancies, expected values

Year	Healthy	Disabled	Dead
0	1000	0	0
1	863	105	31
2	751	185	64
3	658	244	98
4	581	286	133
5	517	316	167
6	463	336	201
7	417	348	235
8	379	353	268
9	346	354	300
10	317	352	331

To determine the *sample* values of the state occupancies at different durations, consider again a sample of 1,000 subjects in state 1 at the beginning of the process. Assume that the subjects do not differ with respect to the transitions between healthy, disabled and dead. All healthy people experience the same incidence rate of disability and the same death rate. Disabled persons experience the same recovery rate and death rate. We construct the lifepath of the 1,000 subjects during a 10-year period. Continuous-time microsimulation is used to determine, for each subject, the time to transition to disability, recovery, or death. Two random variables are generated. The

first is the time to transition drawn from an exponential distribution that is characteristic for the state of origin. Healthy subjects leave the state of being healthy at a time that is determined by $\mu_{1+} = 0.15$. For each subject, the exit time is drawn from an exponential distribution with parameter $\mu_{1+} = 0.15$. The destination state is determined by drawing a random number from a uniform distribution. If the number is between 0 and 0.8, the healthy subject who discontinues to be healthy becomes disabled. If the number is between 0.8 and 1.0, the subject dies. The recovery and death of subjects with disability are determined in a similar way. The time to transition or exit time is drawn from an exponential distribution with parameter $\mu_{2+} = 0.12$ and the direction is determined by drawing a random variable from a uniform distribution. Recovery occurs when the random variable is between 0 and 0.6. If the value of the variable exceeds 0.6, the subject dies. Table 3 shows the state occupancies every year of the period from 0 to 10. The sample observations are close to the expected values. Note that all virtual subjects are followed for a period of 10 years. Censoring is at year 10.

368

Table 3: State occupancies, sample values

Year	Healthy	Disabled	Dead
0	0	0	0
1	863	112	25
2	758	185	57
3	649	261	90
4	551	313	136
5	481	347	172
6	439	355	206
7	393	368	239
8	358	373	269
9	333	369	298
10	306	370	324

Of the 1,000 healthy subjects at the start of the process, 676 become disabled for at least some period and 172 die while healthy (including those who never became disabled and the 156 who recovered from disability before death). 153 die in disability. At the end of the observation at year 10, 306 are healthy, 370 disabled and 324 dead.

The sampling from the virtual population through microsimulation gives information on the population that cannot or cannot easily be obtained otherwise. The 1,000 subjects experience 1,157 transitions (Table 4). Most are

from healthy to disability. 676 transitions initiate an episode of disability. Of the 676 episodes of disability, 156 end in recovery, 153 in dead and 367 are truncated when year 10 is reached. The table also shows the health status at death: 172 are healthy and 153 disabled.

Table 4: Transitions, sample

ORIGIN	DESTINATION			Total
	Healthy	Disabled	Dead	
Healthy	0	676	172	848
Disabled	156	0	153	309
Dead	0	0	0	0
Total	156	676	325	1157

The healthy subjects that become disabled during the observation window of 10 years, become disabled after 3.9 years, on average. Disabled subjects that recover, recover at 5.9 years, on average. Although disabled subjects have a mortality that is twice that of healthy subjects, they die later than healthy subjects: 5.8 years versus 4.2 years. The difference is due to the relatively late onset of disability, the low recovery rate and the high death rate for healthy subjects that is independent of time since onset of the process. As a consequence, many healthy subjects die before the mean age at onset of disability.

A major advantage of continuous-time microsimulation is the possibility of multiple transitions within a year. The number of multiple transitions is relatively rare: 9.9% of the subjects experience two transitions and 0.4% three transitions.

The total number of years spent alive during the observation period is 8.3 years. An average subject spends 5.4 years healthy and 2.9 years disabled.

During the 10-year period, the 1,000 subjects follow 10 different lifepaths. The lifepath is the sequence of states. If H denotes healthy, D disabled and + dead, then a lifepath may be represented by a character variable. The most common path is HD (healthy – disabled). A total of 325 subjects follow that path. 217 subjects remain healthy throughout the 10-year period; they do not experience any transition. The lifepath is H. Of the 172 subjects who die while healthy, 161 die before experiencing a disability, 11 after recovery from a first period of disability and 3 after recovery from a second period of disability (Table 5). The table also shows the mean ages at transition,

followed by the character denoting the transition. Note that subjects with multiple transitions generally experience the transition earlier than subjects with a single transition.

Table 5: Lifepaths during 10-year period, sample of 1,000 subjects

Pathway	Number	Name	Mean age at transition				
1	325	HD	4.24D				
2	217	H					
3	161	H+	4.03+				
4	150	HD+	2.68D	5.67+			
5	84	HDH	3.32D	6.79H			
6	40	HDHD	2.36D	4.88H	7.25D		
7	11	HDH+	1.96D	4.15H	5.77+		
8	7	HDHDH	1.49D	2.85H	5.74D	7.67H	
9	3	HDHD+	1.64D	3.86H	4.97D	6.78+	
10	2	HDHDHD	3.38D	3.92H	6.50D	8.04H	8.17D

3.3 Covariates and interventions

Hazard rates generally depend on personal attributes or covariates. Time-invariant covariates include sex and place of birth. Time-varying covariates include level of education, marital status, employment status, place of residence and health status. The time-path of time-varying covariates is a continuous-time process. It may be approximated by a discrete-time process when piecewise constant hazard rates are used and the covariates are allowed to change at the beginning of time intervals only. In that case, the values of the covariates are updated at the beginning of each interval and the hazard rate is obtained depending on the new values of the covariates. When covariates are allowed to change at any time during the interval, i.e. in continuous time, the interval is split in two or more subintervals and the hazard rates are derived for each subinterval. In other words, the hazard rates are updated whenever covariate values change. That procedure of interval splitting is similar to episode-splitting in event-history modeling (Blossfeld and Rohwer, 2002: 140ff.). The technique involves the splitting of episodes at every point in time where one of the time-varying covariates changes its value. Each of the original episodes is replaced by a contiguous set of subepisodes (splits) with appropriate values of the covariates. Interval splitting is implemented in LifePaths.

By way of example, consider the disability model and consider an intervention programme that reduces the incidence of disability and increases the rate of recovery once disability has struck. Subjects enrol in the programme at different ages and they remain enrolled till the end of the study period, which is 10 years. If a healthy subject enrolls in the programme, the incidence rate of disability drops by 50%. Hence the rate after enrolment is 0.5 times the rate before enrolment. Subjects that are disabled and enrolled in the programme have a higher rate of recovery. The rate is assumed to be three times the rate for subjects not enrolled. Since subjects may enrol at any age, being enrolled is a time-varying covariate. The transition rates after enrolment may be written as

$${}_k m_{12}^e(x) = m_{12}(x) \exp[(\ln 0.5) * {}_k X(x)]$$

$${}_k m_{21}^e(x) = m_{21}(x) \exp[(\ln 3.0) * {}_k X(x)]$$

Where x denotes age, ${}_k m_{ij}^e(x)$ the rate at which subject k changes from state i to state j at age x , and ${}_k X(x)$ is the time-varying covariate that is equal to one if subject k is enrolled at age x and is 0 otherwise.

371

To determine who enrolled in the programme and at what age, i.e. to determine the values of ${}_k X(x)$, a random sample is drawn from the virtual population. It is assumed that, at the population level, 10% of the subjects not yet enrolled at the beginning of a year enrol in the programme. The enrolment rate is independent of the health status, but of course depends on the enrolment status. If the treatment programme is conditional on participation in the prevention programme, only healthy subject may enrol and the enrolment rate is dependent on the health status. Since the enrolment rate is independent of the health status, the expected proportion of subjects enrolled after a period of 10 years is 61.4% ($=100 * [1 - 1 / (1 + 0.10)^{10}]$). In the sample, 712 subjects enrolled, i.e. 71%. Table 6 shows the number of new yearly enrolments in the sample of the virtual population.

In the presence of the intervention programme, less subjects in the sample of the virtual population enter disability (613 versus 676) and more recover from disability (231 versus 156). More subjects die healthy (180 versus 172) because more subjects are healthy and are healthy longer. In the sample population, more subjects may expect to be healthy after 10 years (437 versus 306) and less are disabled (239 versus 370). Because of the intervention, subjects spend more years healthy (6.3 years versus 5.4 years) and less in disability (2.2 years versus 2.9 years). The expected number of

years spent alive during the observation period decreases a little (8.4 years versus 8.3).

If the entire population enrolls in the intervention programme at the start of the observation period, the effect is more significant. In that case the number of healthy years is 7.2 and the number of years in disability is 1.3. The total number of years lived during the observation period is 8.5.

Table 6: Number of new enrolments by year

Year	Enrolment
0	124
1	84
2	80
3	83
4	62
5	58
6	51
7	64
8	34
9	42
10	30
TOTAL	712

372

4 Conclusion and discussion

Continuous-time microsimulation has some advantages over microsimulation in discrete time. The main advantage is that the dates of events and the sequences of events can be determined accurately using the theory of competing risks and continuous-time multistate transition models. Whereas in discrete-time microsimulation sampling from a uniform distribution determines the event occurrences, in continuous-time microsimulation the sampling is from a waiting time distribution. The main tool for continuous-time microsimulation is the inverse distribution function or quantile function. For a given transition model, the function translates the probability of a transition during an interval into a waiting time. Different transition models have different quantile functions. The method based on the quantile function is a general method that applies to all waiting time models and other

models as well. The chapter illustrates the method using the exponential model. Other transition models for which quantile functions can be defined, may be applied in continuous-time microsimulation.

The ultimate aim of microsimulation is to produce a virtual population that closely resembles a real population and to use the virtual population to study characteristics of the real population and to perform experiments *in silico* that are not possible in real populations (*in vivo*). The major weakness of microsimulation is the dependence on the model. If the model is a weak representation of a population, then the results of microsimulation lack validity. In microsimulation, a good model is a necessary condition but it is not sufficient. The sample size also matters. The virtual population must be sufficiently large to closely resemble the real population. With today's computer technology, that weakness can easily be overcome.

With the advent of the R programming environment, continuous-time microsimulation is becoming more easy to implement. The standard R library generates quantile functions for a wide variety of probability distributions, including the exponential and the Weibull distributions but not the Gompertz and the Cox model. Some packages in the R library are particularly useful. For instance, the *msm* (multistate Markov model) package contributed by Jackson (2009) includes a function to estimate a continuous-time Markov model from empirical data and another function (*sim.msm.r*) that uses the model to simulate individual event sequences. Packages recently developed by Putter et al. (2007) and Meira-Machado et al. (2009) also estimate multistate transition models that may be used to simulate life histories. These developments substantially reduce the programming costs of continuous-time microsimulation. Computing time is not substantially larger than in discrete-time microsimulation. The reason is that the application of the quantile function does not involve any iteration. These developments are expected to enhance the use of continuous-time microsimulation in the study of life histories.

References

- Andersen, P.K./Keiding, N. (2002) 'Multi-state models for event history analysis', *Statistical Methods in Medical Research* 11: 91-115.
- Antcliff, S. (1993) 'An Introduction to DYNAMOD: A Dynamic Microsimulation Model', DYNAMOD Technical Paper no. 1, National Centre for Social and Economic Modelling, University of Canberra.
- Balakrishnan, K./Basu, A.P. (1996) *The exponential distribution. Theory, methods and applications*. New York: Gordon and Breach Publishers.
- Bender, R./Augustin, T./Blettner, M. (2005) 'Generating survival times to simulate Cox proportional hazard models', *Statistics in Medicine* 24: 1713-1723.
- Blossfeld, H.-P./Rohwer, G. (2002) *Techniques of event history modeling. New approaches to causal analysis*. Second edition. Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Evans, M./Hastings, N./Peacock, B. (2000) *Statistical distributions*. Third edition. New York: Wiley.
- Galler, H.P. (1997) 'Discrete-time and continuous-time approaches to dynamic microsimulation reconsidered'. Technical Paper no. 13, National Centre for Social and Economic Modelling (NATSEM), Faculty of Management, University of Canberra.
- Gribble, S. (1997) 'LifePaths: A longitudinal microsimulation model using a synthetic approach'. Paper presented in the conference "Microsimulation in Government Policy and Forecasting: International Conference on Combinatorics, Information Theory and Statistics" Maine Portland, Maine, USA, July 18-20, 1997
- Hammel, E.A. (1990) *SOCSIM II*, Working Paper no. 29, Graduate Group in Demography, University of California, Berkeley.
- Hammel, E.A./Hutchinson, D./Wachter, K./Lundy, R./Deuel, R. (1976) 'The SOCSIM Demographic-Sociological Microsimulation Program Operating Manual'. Institute of International Studies Research Monograph No. 27, University of California, Berkeley, California.
- Holmer, M./Janney, A./Cohen, B. (2006) 'PENSIM overview'. Policy Simulation Group, U.S. Department of Labor, Washington, D.C.
- Jackson, C. (2009) The msms package. Version 0.8.2. Reference manual. Available at the Comprehensive R Archive Network, <http://cran.r-project.org/>
- Kelly, S. (2003) 'Forecasting wealth in an ageing Australia. An approach using dynamic microsimulation'. Paper presented at the 7th Nordic Seminar on Microsimulation Models, Helsinki, June 2003.
- Klein, J.P./Moeschberger, M.L. (2003) *Survival analysis. Techniques for censored and truncated data*. Second edition. New York: Springer Verlag.
- Klevmarken N.A./Olovsson, P. (1996) 'Direct and behavioral effects of income tax changes - simulations with the Swedish model MICROHUS', in: Harding, A. (ed.) *Microsimulation and Public Policy*. Amsterdam: North-Holland-Elsevier.
- Meira-Machado, L./Uña-Álvarez, J. de/Cadarso-Suárez, C./Andersen, P.K. (2009) 'Multi-state models for the analysis of time-to-event data', *Statistical Methods in Medical Research* 18 (2): 195-222.
- Mueller, L.D./Nusbaum, T.J./Rose, M.R. (1995) 'The Gompertz equation as a predictive tool in demography', *Experimental Gerontology* 30 (6): 553-569.
- Putter, H./Fiocco, M./Geskus, R.B. (2007) 'Tutorial in biostatistics: competing risks and multi-state models', *Statistics in Medicine* 26: 2389-2430.

- Rowe, G./Nguyen, H. (2004) 'Longitudinal analysis of labour force survey data', *Survey Methodology (Statistics Canada)* 30 (1): 105-114.
- Spielauer, M. (2006) 'The "life-course" model, a competing risk cohort microsimulation model: source code and basic concepts of the generic microsimulation programming language ModGen'. Working Paper WP 2006-046, Max Planck Institute for Demographic Research, Rostock. Available at <http://www.demogr.mpg.de/papers/working/wp-2006-046.pdf>
- Statistics Canada (2001) 'The LifePaths Microsimulation Model: An Overview'. http://statcan.ca/english/spsd/LifePathsOverview_E.pdf
- Wachter, K.W./Blackwell, D./Hammel, E.A. (1998) 'Testing the validity of kinship microsimulation: an update'. Manuscript, www.demog.berkeley.edu/~wachter/
- Willekens, F.J. (2001) 'Theoretical and technical orientations toward longitudinal research in the social sciences', *Canadian Journal of Population* 28 (2): 189-217
- Willekens, F.J. (2006) 'The multistate model for biographic projections'. Manuscript.
- Wolf, D.A. (1986) 'Simulation methods for analyzing continuous-time event history models', *Sociological Methodology* 16: 283-308.
- Wolf, D.A. (2001) 'The role of microsimulation in longitudinal data analysis', *Canadian Studies in Population* 28: 165-179.
- Zaidi, A./Rake, K. (2001) 'Dynamic microsimulation models: a review and some lessons for SAGE'. SAGE Discussion Paper no. 2 [SAGEDP/02]. ESRC SAGE Research Group, The London School of Economics, London. www.lse.ac.uk/depts/sage

Annex A: Continuous-time microsimulation models

Few continuous-time microsimulation models exist. They include the SOCSIM model developed by Hammel et al. at Berkeley (Hammel et al., 1976; Hammel, 1990),^{1,2} the demographic PopSim part of the DYNAMOD model developed at NATSEM (Antcliff, 1993), MICROHUS of Uppsala University (Klevmarken and Olovsson, 1996), LifePaths of Statistics Canada (Gribble, 1997; Statistics Canada, 2001) and PENSIM of the US Department of Labor (Holmer et al., 2006). PENSIM uses the same algorithm as LifePaths (Holmer et al., 2006: 3). The algorithm consists of drawing a sample of waiting times to event and comparing waiting times, generated by hazard models, to determine the timing and sequence of events. For a description of several of these models, including LifePaths, see Zaidi and Rake (2001). Zaidi and Rake assert that "The LifePaths's choice of the continuous time is definitely desirable from a theoretical point of view, although the use of continuous time puts heavy demand on the underlying data and computer resources." (Zaidi and Rake, 2001: 16). The methodology of microsimulation in continuous time was discussed as early as 1986 by Wolf (1986). Researchers at Statistics Canada developed a general-purpose environment for programming microsimulation models, called Model Generator (ModGen). The ModGen language is a superset of the C++ programming language. This environment provides a common code-base for modellers which they can use to generate microsimulation models that are variants of LifePaths. Statistics Canada uses this environment to generate several special-purpose models such as the Population Health Model (POHEM) that uses the demographic module of LifePaths but replaces the mortality equations with a highly detailed model of morbidity and mortality. Spielauer (2006) provides a step-by-step documentation of a continuous-time microsimulation model programmed in ModGen and applied to study fertility change using survey data. Dynamic microsimulation models are transition rate (hazard) models; they use information on members of different birth cohorts to generate life histories of individuals. For a nice illustration, see Rowe and Nguyen (2004), as well as Spielauer (2006).

376

1 For an extensive bibliography, see www.demog.berkeley.edu/~wachter/socrefs.html

2 In SOCSIM, time is measured in integral months (Wachter et al., 1998: 10). The same approach is used in DYNAMOD (Kelly, 2003: 4).