Flexible transition timing in discrete-time multistate life tables using Markov chains with rewards

Schneider, Daniel C.; Myrskylae, Mikko; van Raalte, Alyson

2023-03-07

United Nations Publications

http://hdl.handle.net/10138/572451

Schneider, D C, Myrskylae, M & van Raalte, A 2023, 'Flexible transition timing in discrete-time multistate life tables using Markov chains with rewards', Population Studies, vol. 78, no. 3, pp. 413-427. https://doi.org/10.1080/00324728.2023.2176535

Routledge
Taylor & Francis Group

Population Studies
A Journal of Demography

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/rpst20

Flexible transition timing in discrete-time multistate life tables using Markov chains with rewards

Daniel C. Schneider, Mikko Myrskylä & Alyson van Raalte

Population Studies, 2023
https://doi.org/10.1080/00324728.2023.2176535

Flexible transition timing in discrete-time multistate life tables using Markov chains with rewards

Daniel C. Schneider [1], **Mikko Myrskylä** [1,2] **and Alyson van Raalte** [1]
[1]Max Planck Institute for Demographic Research, [2]University of Helsinki

Discrete-time multistate life tables are attractive because they are easier to understand and apply in comparison with their continuous-time counterparts. While such models are based on a discrete time grid, it is often useful to calculate derived magnitudes (e.g. state occupation times), under assumptions which posit that transitions take place at other times, such as mid-period. Unfortunately, currently available models allow very few choices about transition timing. We propose the use of Markov chains with rewards as a general way of incorporating information on the timing of transitions into the model. We illustrate the usefulness of rewards-based multistate life tables by estimating working life expectancies using different retirement transition timings. We also demonstrate that for the single-state case, the rewards approach matches traditional life-table methods exactly. Finally, we provide code to replicate all results from the paper plus R and Stata packages for general use of the method proposed.

Supplementary material for this article is available at: http://dx.doi.org/10.1080/00324728.2023.2176535

Keywords: life tables; multistate models; Markov chains; working life expectancy; discrete-time event history analysis; Human Mortality Database (HMD); Survey of Health, Ageing and Retirement in Europe (SHARE)

[Submitted August 2021; Final version accepted November 2022]

Introduction

The life course can be thought of as a realization of a stochastic process whereby individuals are subject to risks of moving between states, with the sequence of states and sequence of transitions representing their life histories (Willekens and Putter 2014). Aggregating the occupancy times of individuals in different states subject to these movements results in state expectancies. In a single-state model, with that state being 'alive', this aggregate is the life expectancy. The prime demographic frameworks for analysing such life histories are single-state and—in their generalized version—multistate life tables (MSLTs). The latter provide a truly general framework: they contain all flavours of the traditional life table, including single-decrement, multiple-decrement, and cause-deleted tables, as special cases.

MSLTs have traditionally been modelled in continuous time (Keyfitz and Caswell 2005), frequently using transition rates calculated from occurrence–exposure data as a starting point and using assumptions on the evolution of continuous-time functions, such as within-interval linearity of the survivor function $l(x)$ or constancy of the force of mortality $\mu(x)$ to close the table. We refer to these models as 'continuous-time MSLTs'. However, over the course of the past few decades, longitudinal studies with intermittent observations have become more and more important for empirical research. These data sets often naturally suggest modelling in discrete time, and consequently a new branch of models—'discrete-time MSLTs'—has been developed and applied. For examples of recent applications, see Magnusson Hanson et al. (2018), Chiu (2019), Zaninotto and Steptoe (2019), Farina et al. (2020), and Bardo and Lynch (2021). Transitions in discrete-time MSLTs may, strictly speaking, occur only at the points in time that the model is defined on. Nevertheless, for the calculation of the central magnitudes (e.g. state expectancies) it is useful to reason about their placement at other points in time. For example, on average, are subjects dying mid-interval? Are subjects, on average, marrying closer to the upper end of the 20–24 age interval? Do labour market participants stop working closer to the beginning of their retirement-age year? If left unadjusted, the standard formulas of discrete-time MSLTs deliver numbers that pertain to end-of-interval transitions, which, as some of the preceding examples suggest, is not satisfactory.

As a remedy for the special MSLT case of single-state life tables with age classes of equal size n (typically $n = 1$), van Raalte and Caswell (2013, p. 1620) suggest deducting half of the age interval $\left(\frac{n}{2}\right)$ from the calculated MSLT expectancy. A similar argument has also been made, implicitly or explicitly and in somewhat different contexts, by Guilkey and Rindfuss (1987) and Sonnenberg and Wong (1993). This approach is applicable in situations in which age intervals do not vary and there is only one non-absorbing state (e.g. being alive). It adjusts the MSLT formulas by effectively assuming mid-interval transitions (here, deaths) and therefore works well if the data roughly match this assumption, as they may, for example, for the timing of deaths: if all subjects that die in any (single-year) age interval die halfway into the interval instead of at the end of it, total life expectancy is reduced by half a year. However, this procedure can actually increase the bias if the assumption is not met. Dudel (2021) generalizes the idea of mid-interval transitions to multiple non-absorbing states by deducting half of the age interval from the diagonal of the so-called fundamental matrix, which we explain further later. We refer to this procedure as 'initial period deduction'. The initial period deduction method has two limitations. First, it applies only to a regular age grid and does not, for example, cover the demographic five-year grid that contains the irregular childhood age intervals [0–1) and [1–5). Second, while the addition of the mid-interval assumption is an important step forward, its rigid timing is not suitable for some empirical applications.

In this paper, we demonstrate how the Markov chain with rewards (MCWR) method can flexibly implement transition timing into discrete-time MSLTs. In the multistate context, the rewards method has been used for different purposes. Caswell and van Daalen (2021) provide a general overview of the use cases and techniques for calculation. As an example, they show an application relating to incidence-based healthy longevity, with health status as the (time) 'reward' achieved for making transitions between age stages (i.e. surviving from one age to the next). The lifetime accumulation of such rewards provides the basis for calculating means and variances of time spent at different age and health stage combinations. In their exposition, Caswell and van Daalen (2021) assume either mid- or end-of-interval transitions throughout.

While we defer a more general explanation of the rewards method to our Method section, it is important to note here that our usage of the MCWR method can be seen as implementing a multivariate version of Chiang's a (Chiang 1968, 1972): a (time) reward introduces information on the average time spent in a state by those subjects who transition out of the state. The method then accounts for this information in the calculation of state expectancies by correspondingly allocating time to origin and destination states. In continuous-time models, Chiang's a is a reflection of the model relationship between rates and probabilities or, equivalently, assumptions on some continuous-time model functions. In our context—multistate models that are squarely discrete—introducing specifics on transition timing by means of using external (ad hoc) information (as opposed to assumptions based on functional forms) is not an unattractive option. Compared with the formulations of continuous-time models, this cannot compete in terms of mathematical sophistication, since it remains agnostic about the continuous forces that govern the real world. Still, it does have the advantage of being flexible and easy to implement and compute. A succinct technical exposition of the relevant rewards formulas is given in our Method section. Section 1.1 of the supplementary material provides a detailed treatment.

The contribution of this paper is neither the rewards framework nor the use of external information for a multidimensional Chiang's a, both of which are established in the demographic literature. The novelty of this paper consists in the description of how to combine both features to achieve realistic transition timings in demographic applications based on the increasingly popular discrete-time Markov chain setting, thereby allowing researchers to conduct demographically accurate analyses within that powerful yet accessible framework.

We emphasize that we do not develop methods on how to obtain the magnitude of rewards (i.e. pin down transition times). For traditional life tables, one well-known ad hoc specification is to 'borrow' Chiang's a from another population (Preston et al. 2001, p. 45). In the context of multistate models (which can address a wide range of topics, not just mortality), types and sources of external information can take many more forms. We therefore cannot provide concrete advice or methods for the acquisition of timing information; rather, we describe calculations once the researcher has obtained such values, for example through (separate) estimation from data or via assumptions.

We illustrate the flexibility of rewards-based MSLTs with respect to transition timing with two examples. First, we use Human Mortality Database (HMD) data to show that rewards-based MSLTs coincide numerically with traditional spreadsheet-type life-table calculations in the single-state case. This is in contrast to discrete-time MSLTs based on end-of-interval or mid-interval transitions, where bias can be large. In our second empirical example, we turn to the case of multiple states and use retrospective survey data from the Survey of Health, Ageing and Retirement in Europe (SHARE) on working-life histories to estimate state expectancies for employment and retirement. We show that the rewards-based MSLT can accurately exploit information on state entry/exit timings, even though retirement timings follow a highly unusual pattern. We again analyse the bias in standard discrete-time MSLTs. Finally, we provide Stata and R packages that implement the calculations, to facilitate replication and wider use of the method.

Background: The estimation of multistate life tables

The origins of the life table, which is probably the most well-known demographic tool, date back to the seventeenth century (Graunt [1662] 1964; Halley [1693] 1874). Maybe somewhat surprisingly, generalizations to multiple alive states were not conceived until the twentieth century (Du Pasquier 1912). The subject finally received extensive treatment during the 1970s and early 1980s (Schoen 1975; Schoen and Land 1979; Land and Rogers 1982). Within MSLTs, the key life-table outcome, life expectancy, is generalized to state expectancies. For example, using MSLTs the researcher can calculate the expected lifetime spent in marriage; the expected lifetime being unemployed; the expected pain-free lifetime; and so forth. One fundamental assumption that is part of any life-table method is the Markov assumption, which posits that the probability of making a particular state transition depends only on the current state, not on the entire history of states occupied. For this reason, and since concepts are usually expressed in matrix formulas, MSLTs are also described in the literature as 'multistate Markov models', 'matrix population models', or similar. Estimating an MSLT can be seen as consisting of two steps: First the basic parameters are estimated from the data, and in the second step Markov theory is applied to calculate the desired statistics. In this paper, when talking about a 'model', both steps are implied. For the second step, state expectancies are the most fundamental concept, but an immense number of results related to other statistics also exist. For further exposition and references, see Caswell et al. (2018).

Both the traditional single-state life table (whose single state is 'alive') and traditional multistate life tables are expressed in continuous time. In some expositions this fact may be hidden behind notation that looks discrete, but the underlying ideas—a continuous force of mortality and, by implication, a continuous survivor function and age distribution at death—are always rooted in continuous time. The central model concept is that of instantaneous transition rates, with subsequent model calculations leading to transition probabilities defined on pairs of points in time, the number of time units spent in a state in a certain age interval, and finally, state and overall life expectancies. These traditional MSLTs, or increment–decrement life tables as they were then mostly called, were invariably based on some functional form assumption (e.g. a linear or exponential survivorship function within age groups) to close the model. Different assumptions result in different formulas for the calculation of MSLT magnitudes: most notably in the present context, transition timings (see Schoen [1979] for a comparison of some popular methods).

While the model inputs of the MSLTs developed in the 1970s and 1980s are period transition rates, a related strand of popular models—multistate survival (regression) models—is based on assumptions on properties of the hazard. Semi-parametric models harness fully fledged (i.e. precisely dated) life-history data to estimate schedules of instantaneous rates for further processing; parametric models do the same but must be based on assumptions about the global shape of the hazard. As in traditional MSLTs, assumptions on continuous-time functions pin down the transition timing formulas of the model.

A huge advantage of regression models over the traditional MSLT estimation techniques is that they can accommodate covariates. Different tables can be presented for different values of those covariates based on the same parameter estimates, dispensing with the need for sample splits. Still, there are downsides that all these models share. They require a larger theoretical exposition, as well as knowledge of matrix algebra and differential and integral calculus. Moreover, they are frequently computationally burdensome. For these reasons, they are not always easily accessible to researchers.

More recently, another strand of demographic and epidemiological literature has emerged, starting with Millimet et al. (2003). The model formulations of this new body of work share many ideas with the continuous-time MSLTs already mentioned. They too make use of Markov theory and frequently express their ideas using matrix notation. However, they differ in that they are cast in discrete time, and this makes them considerably simpler. While there are many textbooks available on continuous-time MSLTs (Hougaard 2012; Willekens 2014; van den Hout 2017; Cook and Lawless 2018), discrete-time MSLTs do not require textbook-length treatment. One of their appealing features is that they are easy to understand, communicate, and apply. A substantial simplification is that transition probabilities are estimated directly and do not need to be inferred from estimated rates or estimated instantaneous rates.

At the very simplest level, transition probabilities can be estimated via transition counts. Slightly more sophisticated is the use of multinomial logistic regression, which, like the continuous-time survival models, can accommodate covariates. It is a regression technique that is well accessible to quantitative social scientists. More background on discrete-time regression procedures can be found in Jenkins (1995), Willett and Singer (1995), and Singer and Willett (2003). Applications of these techniques in a multistate (competing-risk) setting are treated in Allison (1982), Steele (2011), and Tutz and Schmid (2016, ch. 8). The estimated regression models allow for easy prediction of transition probabilities. For the Markov calculations that use the probabilities as an input, only knowledge of matrix multiplication, one of the most basic operations on matrices, is required. Matrix formulas for obtaining state expectancies are outlined in the next section and explained in greater detail in the supplementary material (sections 1.1.1 and 1.1.2). Moreover, aside from the bootstrap procedure used to obtain standard errors, the computational cost is typically very low. Noteworthy extensions of this procedure include: the IMaCh method (Lièvre et al. 2003; see also Brouard 2021), which calculates a higher-frequency (embedded) Markov chain; the SPACE method (Cai et al. 2010), which can compute a large number of statistics harnessing simulation techniques; and GSMLT, a Bayesian approach developed by Lynch and Brown (2005). Of these, IMaCh is the only procedure that has potentially attractive properties with respect to transition timing. We will come back to this point in the next section.

A salient feature of discrete-time models is that transition probabilities are obtained directly from an estimated model without any prior calculation of rates. It is important to note that this simplification also bears important costs. First, regularly spaced data are required. Second, the interval-censored data they run on require the assumption of no unobserved transitions between observational points. Third, and of central interest for the present discussion, the estimation procedure is agnostic about transition timing. Given these caveats, the choice between a continuous- or discrete-time model frequently boils down to an assessment of whether the continuous-time model can make substantially better use of the information and whether some of the simplifying assumptions needed for discrete-time models are tenable. Among the points to consider are whether the data provide precisely dated transitions and/or interview/sampling times. If they do, discrete-time methods may induce loss of information. Generally, the assessment of whether slight inaccuracies in the use of the timing of observations and/or additional assumptions are justifiable needs to be done on a case-by-case basis. For example, for an annual survey, ignoring the variation in interview dates within a few weeks may pose no problem, but the situation is different for a survey whose wave spacing varies between two and five years. In cases of doubt, it is always possible to cross-check discrete-time results against continuous-time ones. The important point to grasp is that if the simplifying assumptions seem innocuous, discrete-time models are an attractive choice because they are: (1) easy to understand; (2) easy to apply; and (3) easy to communicate. We now elaborate on each of these points in turn.

First, the discrete-time multistate method is widely accessible to researchers because it does not require specialized knowledge as does, for example, multistate estimation in continuous time that deals with interval censoring. The supplementary material, which contains a mathematical exposition of the relevant discrete-time material (sections 1.1.1 and 1.1.2), is a good illustration of this. Even though the discussion starts from first principles regarding matrix algebra, the relevant sections span merely seven pages. Second, any statistical software that allows matrix expressions can be used to process the relatively few lines of code that are needed to evaluate the relevant expressions. Crucially also in this respect, the computational cost is typically lower than for continuous-time models; this not only affects the research flow but also expands the possibilities for model formulation, most notably with respect to the number of states and number of covariates. Given a suitably sized data set for stable parameter estimation, even large models with, say, 10 states and 10 covariates are often calculated within minutes using standard techniques, such as multinomial logistic regression. The ensuing calculation of state expectancies then is only a matter of seconds. The only aspect that needs a considerable amount of computer time is the bootstrap that must be run to construct confidence intervals. However, alternatives based on asymptotic theory are possible, already partially spelled out in the literature (Lièvre et al. 2003), and under active development by the authors of this paper. Lastly, the relative simplicity of discrete-time multistate models makes it easier for researchers to target their papers towards a wider audience. The complete method can be stated in relatively little space and can often be fully accommodated in a paper's methods section. At most, a short appendix is required, but no references to external textbooks are necessary in principle. It should be remembered that all these advantages rest on certain features of the data and corresponding assumptions. These requirements must be met in order for the method to be applicable. In cases where they are likely (or known) to be violated, the researcher needs to turn to alternative methods.

Being discrete-time MSLTs, rewards-based MSLTs inherit all their advantages (and shortcomings). The distinguishing feature of rewards-based MSLTs compared with other discrete-time MSLTs is that they improve on timing options in order to deliver more accurate results. The specification of rewards can be seen as the discrete-time analogue of the functional form assumptions in continuous-time models. We turn to the description of rewards-based MSLTs next.

The method of Markov chains with rewards

An absorbing Markov chain describes the trajectories of individuals through the states of the chain, eventually arriving at the absorbing state of death. An MCWR associates a fixed or stochastic numeric value (the reward) with each possible transition from one state into another. In his original development of the method, Howard (1960, p. 17) emphasizes that rewards can be anything (dollars, voltage levels, units of production, etc.). The theory was introduced into ecology and demography by Caswell (2011) in the context of lifetime reproduction. Here, rewards count towards the number of offspring produced. The theory has since been applied to various other topics, including the fertility transition (van Daalen and Caswell 2015), lifetime income and expenditures (Caswell and Kluge 2015), healthy longevity and disability-adjusted life years (Caswell and Zarulli 2018), episodes of disability (Dudel and Myrskylä 2020), and periods of malnourishment (Owoeye et al. 2020). Note that in many demographic applications, including the ones in this paper, 'reward' can be read as 'time reward': for example, time that accrues to the expected time spent in a particular state. The analysis of an individual's accumulated rewards then leads to the analysis of state expectancies and longevity. The complete MCWR theory for demographic applications, including simple expressions for all the statistical moments and for sensitivity analysis, is given in van Daalen and Caswell (2017) and Caswell and van Daalen (2021).

The model of our exposition is based on an age–stage Markov chain whose state space comprises the Cartesian product of ω age classes and τ transient stages. For simplicity, we assume that subjects that are in transient stages will eventually transition into a single absorbing state (death), but a generalization to multiple absorbing states (e.g. multiple causes of death) is not difficult. As is common in the literature, we opt for slightly ambiguous terminology and sometimes use 'state' in place of 'stage'. Later on, in the retirement example, the states are employed, unemployed/out of labour force, retired, and dead, with the first three being transient states and death being an absorbing state. The terminological ambiguity arises since a state can now refer either to employment status or to the Markov state, which is the employment status at a certain age. However, what is meant will be clear from the context. This section assumes regularly spaced age intervals of size n, but results can easily be generalized to irregularly spaced age intervals. The Markov assumption is that the process is memoryless, that is, the probability of a particular transition depends only on the current state, not on the history of previous states occupied.

Matrices and vectors of a Markov model that encompasses both age and state can be organized by age within state or by state within age. Both presentations are equivalent in the substantive sense. We opt for the age-within-state ordering in the following exposition because the ordering of elements in the fundamental matrix (see section 1.1.2, supplementary material) is more suitable for our purposes. Let $\boldsymbol{P}$ denote the transition matrix of the process, which we partition according to transient states and the single absorbing state:

$$\boldsymbol{P} = \begin{pmatrix} \boldsymbol{U} & 0 \\ \boldsymbol{p'}_{\boldsymbol{d}} & 1 \end{pmatrix} \tag{1}$$

The row vector $\boldsymbol{p'}_{\boldsymbol{d}}$ contains probabilities of dying:

$$\boldsymbol{p'}_{\boldsymbol{d}} = \begin{pmatrix} \boldsymbol{p'}_{d1} & \cdots & \boldsymbol{p'}_{d\tau} \end{pmatrix} \tag{2}$$

where each subvector in the expression corresponds to a particular transient state and contains death probabilities for all ages. The submatrix $\boldsymbol{U}$ of $\boldsymbol{P}$ contains transitions among the τ transient states:

$$\boldsymbol{U} = \begin{pmatrix} \boldsymbol{P}_{11} & \cdots & \boldsymbol{P}_{1\tau} \\ \vdots & \ddots & \vdots \\ \boldsymbol{P}_{\tau 1} & \cdots & \boldsymbol{P}_{\tau\tau} \end{pmatrix} \tag{3}$$

where $\boldsymbol{P}_{ij}$ denotes a $\omega \times \omega$ matrix (corresponding to the number of age classes) with non-zero elements on the first subdiagonal only:

$$\boldsymbol{P}_{ij} = \begin{pmatrix} 0 & 0 & \cdots & 0 & 0 \\ p_{ij,2} & 0 & \cdots & 0 & 0 \\ 0 & p_{ij,3} & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & p_{ij,\,\omega} & 0 \end{pmatrix} \tag{4}$$

Element $p_{ij,\,k}$ denotes the probability of moving from state j to state i when entering age class k, which runs from $k = 2$ to $k = \omega$; k can also be seen as indexing the beginning age of each age interval, which can be denoted by x_k. Differently from frequent demographic notation, we define $k = 1$ to correspond to the baseline age, so the first transition takes place at $k = 2$. Also note that we use column-stochastic matrices, whose columns sum to one: column indices (j) refer to origin states and row indices (i) refer to destination states. Many papers use the alternative set-up of row-stochastic matrices, with the consequence that all expressions appear transposed. It is important here not to let the differing conventions become a source of confusion. The supplementary material (section 1) explains these issues in greater detail.

An important matrix in Markov chain theory and in MSLTs is the fundamental matrix, which is calculated as:

$$\boldsymbol{F} = (\boldsymbol{I}_{\omega\tau} - \boldsymbol{U})^{-1} \tag{5}$$

where $\boldsymbol{I}_{\omega\tau}$ denotes an $\omega\tau \times \omega\tau$ identity matrix and the power of minus one stands for the matrix inverse. The special structure of age–stage-classified models, in which each age–stage combination can be reached at most once due to the ageing process, implies that the fundamental matrix contains for each initial age–state the probabilities of reaching any later age–state. When multiplied by n, the length of the (regular) age interval, this is equivalent to the expected length of stay in a particular age–state, given an initial age–state. For example, for annual data on subjects starting at age 50, an element of $\boldsymbol{F}$ may indicate that a person's probability of being in the employed state at age 60 is 0.8 if their initial age–state was employed at age 50. Put differently, this matrix element says that for a large number of individuals, 80 per cent of those who are initially employed at age 50 will be employed at age 60. The remaining 20 per cent will be either in the absorbing state 'dead' or in one of the alternative non-absorbing states (unemployed/out of labour force; retirement) at age 60. Multiplying $\boldsymbol{F}$ by the length of the age interval (here, $n = 1$) results in a matrix whose elements indicate the expected length of stay. Finally, summing up the appropriate elements of $n \cdot \boldsymbol{F}$ will give state expectancies at age 50, given an initial state. A weighted average of these magnitudes yields state expectancies independent of the initial state. Many papers use empirical proportions of study participants at or around the baseline age of the model. An alternative is to use stable prevalence weights implied by the transition probabilities (Rogers 1975; Brouard 2019). Generalizations to an irregular-age grid are easily accommodated, as explained mathematically in the supplementary material (section 1) and demonstrated in our example application in the next section.

The procedure just described yields state expectancies based on the assumption of end-of-period transitions. Instead of $\boldsymbol{F}$, the initial period deduction method uses $\tilde{\boldsymbol{F}} = \boldsymbol{F} - \frac{n}{2}\boldsymbol{I}_{\omega\tau}$, which corresponds to the assumption of mid-period transitions. To add more flexibility with respect to transition timing, we introduce for each state (m) rewards matrices $\boldsymbol{R}_{\boldsymbol{m}}$. Their structure in terms of zero and (potentially) non-zero elements is identical to the full transition matrix $\boldsymbol{P}$. For any transition, say from state j to state i, and at any age, it is easy to make corresponding entries in the rewards matrices $\boldsymbol{R}_j$ and $\boldsymbol{R}_i$ (i.e. $m = j$ for $\boldsymbol{R}_j$ and $m = i$ for $\boldsymbol{R}_i$) that attribute time rewards to states j and i according to time spent in the origin state and time spent in the destination state for that particular transition. The rewards-based calculation of the expected time spent in any state m links the information on the probability of reaching a certain state at a certain age (as indicated by the matrix $\boldsymbol{F}$) with the rewards towards a particular state when moving out of the age–state, as embodied in the elementwise matrix product $\boldsymbol{P} \odot \boldsymbol{R}_m$, in which each element of $\boldsymbol{P}$ is multiplied with the corresponding (single) element in $\boldsymbol{R}_m$. To continue with the previous numerical example that assumed a probability of 0.8 of being employed when turning 60, if additionally there was a probability of 0.1 of making the transition from employment to retirement during age 60, and if on average those who retired worked for 0.3 years before retiring, the transition from employment to retirement during age 60 would contribute $0.8 \times 0.1 \times 0.3 = 0.024$ years to working life expectancy. If the only additional state was death and there was no mortality during age 60, the transition would contribute $0.8 \times 0.1 \times 0.7 = 0.056$ years to retirement life expectancy. Likewise, the transition to staying employed during age 60 would contribute $0.8 \times 0.9 \times 1 = 0.72$ years to working life expectancy and $0.8 \times 0.9 \times 0 = 0$ years to retirement life expectancy.

The mathematical exposition in this section has been brief and geared towards an intuitive understanding. The reader is invited to consult the supplementary material (section 1), which contains a much more thorough description of the full method and the formulas involved. Furthermore, section 1.3 of the supplementary material shows the rewards matrices pertaining to the multistate application given in the next section. Readers may find these matrices helpful for understanding the details of the calculations once they have familiarized themselves with the multistate application in the main text.

As mentioned in the previous section, the nonzero elements of the model matrices are typically estimated from data. For example, a popular choice for estimating the $\boldsymbol{P}_{ij}$ are predicted probabilities from multinomial logit models. Likewise, to use the rewards method, the data must typically also allow for estimating the relevant rewards parameters. The empirical analysis on retirement transitions (in the next section) is an example of this. It goes without saying that parameters, in addition to being estimated, can also be merely assumed, for example for counterfactual analyses.

This section has presented MCWR as an enhancement of discrete-time methods with respect to transition timing. As mentioned, one discrete-time method—IMaCh—already has related capabilities. IMaCh stands for 'Interpolated Markov Chain'. By mapping the observational times onto a higher-frequency grid, the IMaCh method can, given a fine enough grid: (1) account for transition timing; (2) dispense with the assumption of no unobserved transitions taking place between two observational points; and (3) accommodate irregularly spaced observations.

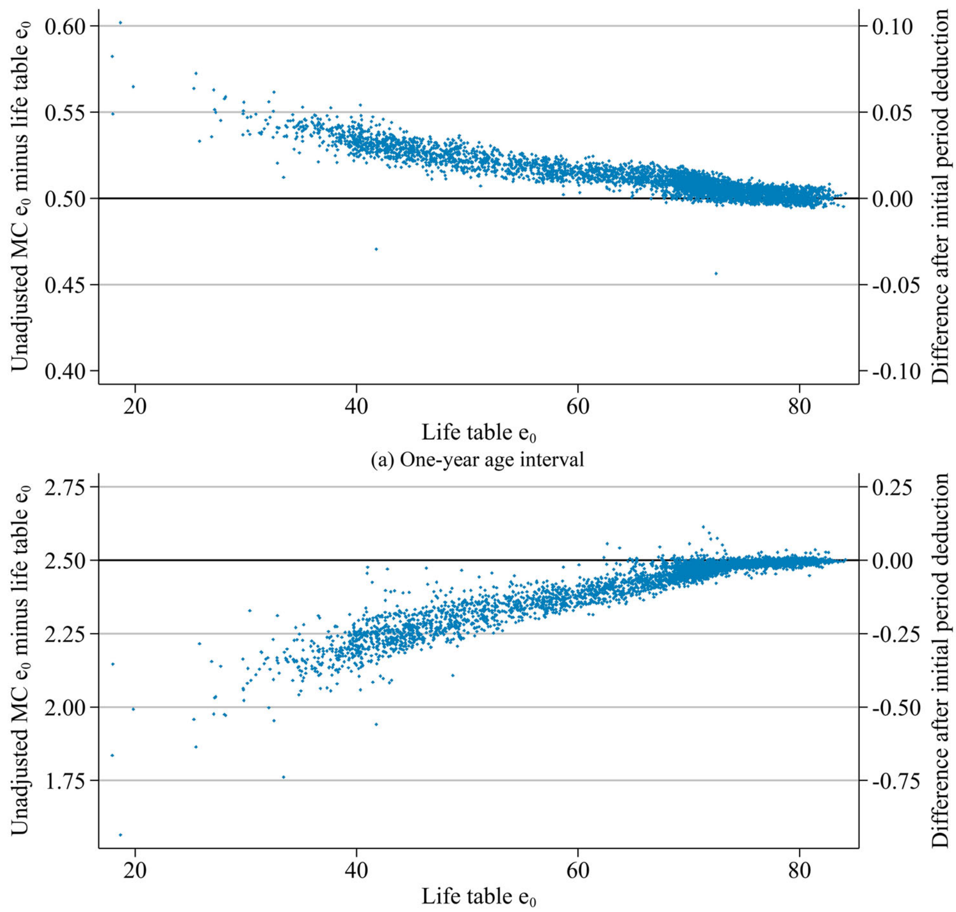
Its scope for applications, however, also has limitations. IMaCh is computationally costly, to the extent that for many applications it curtails the state space or the desired set of regressors. It is also not yet available as a package in any of the major statistics software bundles. Moreover, using a monthly data set on limitations of activities of daily living for measuring the relative performance of IMaCh against a regular interval-censored discrete-time model, Wolf and Gill (2009) do not find that the former substantially outperforms the latter, at least with respect to the calculation of state expectancies. Nevertheless, for smaller models IMaCh may well be the discrete-time method of choice, and it is to be hoped that computational improvements will broaden its applicability in the future.

Empirical illustrations

We illustrate two interesting aspects of the rewards method with empirical applications. The first touches on the issue of traditional (single-state) life tables being a special case of MSLTs, which in the discrete case is not strictly true. The rewards method is shown to be suitable for bridging the numerical gap. The second application is a proper multistate one and shows how the rewards-based MSLTs incorporate additional information, thereby refining the estimates. The example chosen uses information on retirement transitions to improve estimates of working life expectancy.

Life expectancy in the Human Mortality Database

Figure 1 uses data from the HMD (University of California, Berkeley, and Max Planck Institute for Demographic Research, Germany, n.d.; Wilmoth et al. 2020) with one-year (panel (a)) and five-year (panel (b)) age spacing, respectively, to illustrate the magnitude and variation of error in the MSLT with end-of-period transitions when compared with life-table calculations following the usual method (see left-hand y-axis). As is standard in demographic tabulations, the five-year grid is not exactly regular as it includes the usual childhood intervals of ages



(a) One-year age interval

(b) Five-year age interval

Figure 1 Difference between uncorrected Markov chain estimates of life expectancy and life-table-based life expectancy

Notes: Each data point is based on life-table data for both sexes combined for a particular country and year. The figures include data for all countries and years recorded in the HMD. Life expectancy is life expectancy at birth (e_0). MC refers to Markov chain.

Source: Human Mortality Database (HMD).

0 and 1–4. The magnitude and direction of bias in the life expectancy estimates for this example is determined mostly by infant mortality. In the first year, deaths occur shortly after birth on average, that is, at the very left end of the interval and not at mid-interval (age 0.5 years). Therefore, the end-of-period transition assumption that death occurs at age one introduces a strong upward bias. Hypothetically, the maximum bias for this interval is one year. The overall impact of the end-of-period transition assumption for infancy on life expectancy estimates depends on both the average age at death among those who die and the magnitude of infant mortality. For the numerically higher life expectancy estimates in Figure 1, infant deaths occur on average closer to birth (Andreev and Kingkade 2015), which leads to a larger bias per death, but this is overcompensated by the higher survival probabilities of the first age interval. The overall effect is a reduction in the bias in life expectancy estimates. At ages older than zero, subjects die closer to mid-interval on average. Here the end-of-period transition assumption introduces a bias close to 0.5 for intervals in the one-year age grid (panel (a)). The result is a downward-sloping data cloud, approaching from above a bias of roughly half of the age interval (0.5) as expectancies increase. The upward-sloping data cloud of the five-year spacing (panel (b)) is explained by similar reasoning: the first age interval (age 0) can contribute a maximum bias of one year. The second age interval (ages 1–4) contributes on average a bias of roughly two years, whereas higher age intervals contribute on average roughly 2.5 years. Consequently, as deaths shift to older age intervals, the bias approaches 2.5 from below.

The right-hand y-axes in both panels of Figure 1 depict the remaining bias after initial period deduction. Note that initial period deduction is, strictly speaking, not applicable to the unevenly spaced age grid (panel (b)), but in the case of a single transitory state simply amounts to deducting a fixed number from the total life expectancy, which seems like a useful shortcut. Initial period deduction removes a sizeable fraction of the error present in the unadjusted numbers but does not solve the problem completely.

A solution is provided by the rewards-based approach, which yields identical values to standard life-table calculations. The corresponding rewards specification consists of Chiang's a, that is, the average time lived by those who die in the age interval, commonly denoted by ${}_n a_x$. This information—based on assumptions—is available in the HMD. Furthermore, survivors of an age class are assigned a reward of $n = 1$ in the single-age interval case or $n_i \in \{1, 4, 5\}$ in the five-year interval case. Section 1.2 in the supplementary material shows that the two approaches are fully numerically equivalent.

Working and retirement life expectancies in SHARE

To illustrate the error in MSLTs in a situation of multiple non-absorbing states when conventional transition timing assumptions are not met, we calculate working and retirement life expectancies at age 50 using the SHARE data and the accompanying Job Episodes Panel (JEP). Increasing life expectancy and increased recognition of the fact that actual retirement decisions differ from mandatory retirement ages have spurred research into how long people actually work and how long they can enjoy retirement (Loichinger and Weber 2016; Leinonen et al. 2018; Lorenti et al. 2019). Although retirement decisions are often only partially related to mandatory retirement ages, in many settings the level of benefits depends on age, such that after certain birthdays (e.g. the 65th) there is a step increase in benefits. Such incentives may result in transitions that occur early on within age intervals.

Our main data source is the JEP version 6.0.0 (Brugiavini et al. 2013; Antonova et al. 2014; Orso et al. 2017). The JEP, in turn, is derived mainly from Wave 3 of SHARE (Börsch-Supan et al. 2013) and complemented by information from Waves 1 and 2. SHARE is a representative longitudinal survey that started in 2002 and contains data on more than 120,000 individuals aged 50+. It is conducted for 16 EU member states and funded mainly by the European Commission, with additional funding from the German Federal Ministry of Education and Research, the Max Planck Society for the Advancement of Science, the US National Institute on Aging, and various national funding sources. Wave 3 of SHARE, known as SHARELIFE, is special in that it focuses on collecting data on individuals' life histories and contains modules with retrospective information. Data were collected during 2008–09 for almost 30,000 subjects in 14 SHARE countries (Austria, Belgium, Czech Republic, Denmark, France, Germany, Greece, Ireland, Italy, the Netherlands, Poland, Spain, Sweden, and Switzerland). The JEP is a retrospective long-panel data set based on SHARELIFE information and features full working-life histories for each SHARELIFE respondent. We supplement the JEP with basic demographic information taken from Wave 2 of SHARE.

The accuracy of work histories is at the level of one state (i.e. employed, out of labour force/ unemployed, retired) recorded per age-year. Precise start and end dates of state spells are not given. The JEP contains such work histories for 28,492 subjects. After dropping subjects with no record beyond age 50 or who are already retired at age 50 or whose state information we deem too inaccurate, 26,554 subjects remain. The smallest subsample contains 350 subjects (men in Ireland) and the largest 1,519 (women in Greece). Although the longest life history in our data set goes back to 1957, the bulk of the data points (more than 85 per cent) fall in the time range 1985–2008.

We use these data and a multinomial logit model to estimate transition probabilities between employment states for each age. Estimation is stratified by sex and country. Since the JEP is based entirely on retrospective interviews, it does not contain mortality information. Therefore, we use data from the HMD to calculate country-specific probabilities of dying by sex and single age over the period 1985–2004. We apply the resulting mortality conditions to all states in all our analyses. We include age dummies for each single age from 50 to 70 in the regression, and it is for these ages that we calculate transition probabilities. After age 70, retirement is assumed. About 3.7 per cent of observations in the JEP data set show conflicting information between retirement status and other employment states (Brugiavini et al. 2013, p. 9). In such cases we use the first retirement date given. As a consequence, no transitions from retirement back to another employment state occur in the data. We also slightly reclassify employment states to resolve conflicting state information for a smaller proportion of observations. All analyses are unweighted.

Wave 2 data on the year and month of retirement (SHARE variables ep328_ and ep329_) are used to calculate the average (fractional) age at retirement for each age, separately by sex and country. This information is not included in the JEP and is available only for a subset of individuals (39 per cent of the subjects in the estimation sample); this is not a problem in the present context since we need only mean estimates. We use this fraction of the retirement age as a reward towards the state directly preceding retirement (working or unemployed) for this age, and we use one minus that fraction as the reward towards retirement life expectancy. In terms of the rewards matrices in the Method section, these values enter the matrices $\boldsymbol{R}_j$ and $\boldsymbol{R}_i$, respectively.

Figure 2 shows for work–retirement transitions the average fraction of the year that is spent working. It can be seen that retirement transitions during the ages from the early to mid-60s—where the bulk of retirement transitions occur—often take place close to the beginning of the interval. This suggests that incorporating transition timing information in the calculation of working and retirement life expectancies is important.

Figure 3 shows rewards-based state and total life expectancies at age 50 for all SHARE countries and for individual countries, as well as the bias of other methods. Panels (a) and (b) show values for men and for women, respectively. In each panel, the top graph displays total and component life expectancies calculated using rewards-based MSLTs. For each transition the rewards approach accurately assigns the time that individuals spend on average in the origin and destination states. The middle and bottom graphs in each panel compare these results with those from other discrete-time methods. The middle graphs depict the difference in working life expectancy values between the rewards method and end-of-period calculations (diamonds), as well as the difference between the rewards method and the initial period deduction method, that is, the method that subtracts half of the age interval from the diagonal of the fundamental matrix (triangles). The bottom graph in each panel does the same for retirement life expectancy.

For both men and women, the end-of-period approach (diamonds) overestimates working life expectancy in most cases, and the magnitude of the error is often more than half a year, despite the length of the age interval being only one year. The initial period deductions approach (triangles) delivers estimates that are much closer to the rewards approach, eliminating most of the bias.

For retirement life expectancies (bottom graphs in each panel), the direction of the error is switched: the unadjusted end-of-period values are about 0.2 years short of the rewards values (i.e. they understate the length of retirement). Note that these numbers are not simply negatives of those for working life expectancies for two reasons. First, the methodological discrepancy between the life expectancy estimates amount to 0.5, not zero, so the component life expectancy differences add up to that, not to zero. Second, the results for unemployment life expectancy (not shown) capture the remainder. Two features of the initial-period-deduction-based retirement differences catch the eye: they are

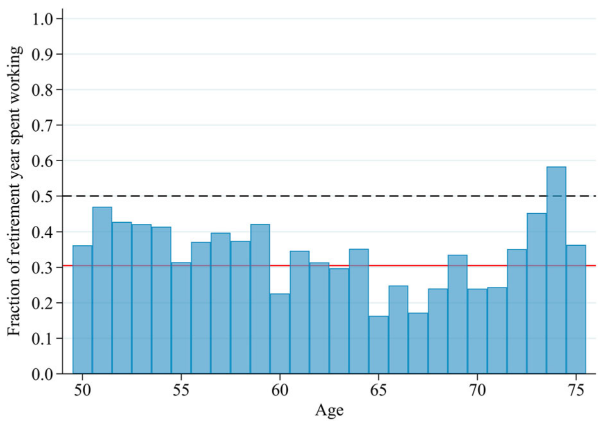


Figure 2 Mean fraction of the year of retirement spent working, by single age, European countries, 2006
Notes: Based on data for all countries from SHARE, Wave 2. The horizontal dashed line corresponds to the mid-interval reward of 0.5 to each of the states of origin and destination. The horizontal solid line depicts the overall average calculated over all retirement events in the data.
Source: Survey of Health, Ageing and Retirement in Europe (SHARE), Wave 2 (2006).

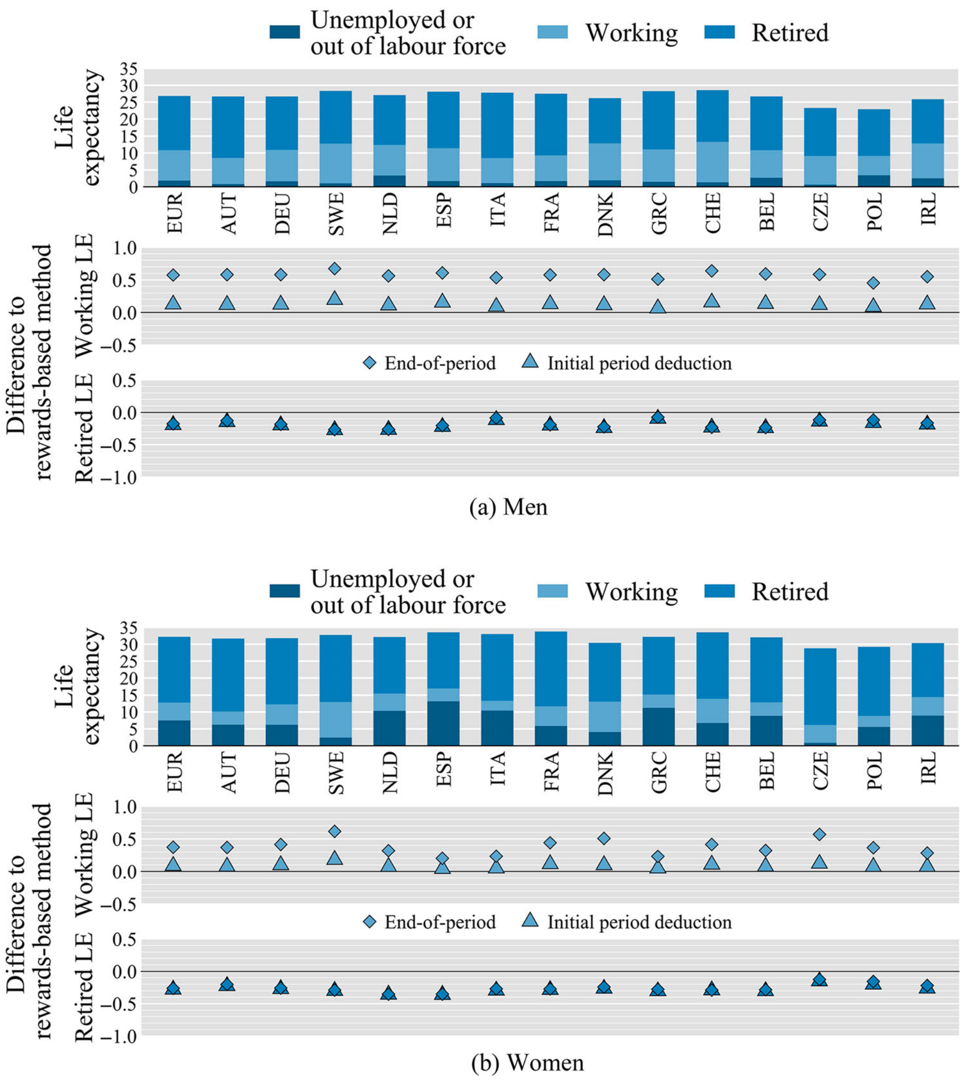


Figure 3 Working and retirement life expectancies at age 50 for men and women across SHARE countries: calculated using MCWR method and differences thereof with respect to end-of-period calculations and initial period deduction

Notes: Based on data for all countries from SHARE Wave 2 and the JEP from SHARELIFE. For each panel, the top graph shows total and component life expectancies. The middle graph depicts the difference in working life expectancy (LE) values between the rewards method and end-of-period calculations (diamonds), as well as the difference between rewards and the initial period deduction method (triangles). The bottom graph of each panel does the same for retirement life expectancy. All expectancies are in years.

Source: Survey of Health, Ageing and Retirement in Europe (SHARE), Waves 2 and 3.

almost identical to the unadjusted end-of-period values, and this method sometimes even slightly increases the bias. The first feature is explained by the fact that the deduction-based method weighs the deduction from the initial period by the population distribution over the initial states. Since at age 50 the fraction of retired people is very small, the deduction from retirement life expectancy is small, and so the unadjusted and adjusted values are close. The second feature is a simple consequence of the fact that the end-of-period estimates are lower than the rewards-based ones, and any deduction further diminishes their values and increases the bias.

Conclusion

Discrete-time MSLT approaches to calculating state-specific expectancies are widely used in demography. Unlike their continuous-time MSLT forebears, which solved the transition timing problem by assuming various functional forms (mostly of survivorship within age groups), discrete-time MSLT approaches have needed to rely on oversimplified assumptions on transition timing between age intervals. The standard approach implies an assumption of transitions taking place at the end of the interval. This results in a discrepancy between Markov chain estimates and life-table estimates of state expectancies. Using two high-quality data sets on life expectancy, working life expectancy, and retirement life expectancy, we have shown that the error can be non-negligible. However, it can be completely removed using the MCWR approach, in which the researcher has full control over the timing of transitions, in the sense that model calculations can accommodate any (separately obtained) values of transition times.

In the simplest case of one non-absorbing state, the MCWR estimates are the same as the standard life-table calculations traditionally used by demographers. In the case of multiple non-absorbing states, discrete-time MSLTs currently offer only a very limited set of timing choices, and continuous-time models are often more difficult to apply. Here, the proposed rewards-based discrete-time method improves the accuracy of state expectancy calculations compared with existing discrete-time methods. We encourage users to incorporate rewards in their analyses to remove the error in discrete-time MSLT state expectancies. To facilitate replication and wider use of the rewards approach, this paper is supplemented by R and Stata packages that implement and guide calculations (see Note 3 and supplementary material, section 2).

Notes and acknowledgements

1 All three authors are based at the Max Planck Institute for Demographic Research, Rostock, Germany. Mikko Myrskylä is also at the Centre for Social Data Science, University of Helsinki, Finland.

2 Please direct all correspondence to Daniel C. Schneider, Max Planck Institute for Demographic Research, Konrad-Zuse-Str. 1, 18057 Rostock, Germany; or by E-mail: schneider@demogr.mpg.de.

3 R and Stata packages accompany this research paper. The package name for either software package is mcwr (Markov chains with rewards).
The R package can be installed from the CRAN repository for R versions 3.6.0 or higher by issuing in R:

```
> install.packages("mcwr")
```

The Stata package can be installed from the SSC repository for Stata versions 12 or higher by issuing in Stata:

```
. ssc install mcwr
```

4 A Stata replication script for all results of this paper can be found at: https://osf.io/68tkb/?view_only=cac73ab86b274d2c9a6f408965fab168. The source data, which are available free of charge, must be obtained separately. Instructions for this are given in the replication script.

5 Mikko Myrskylä was supported by the Strategic Research Council (SRC), FLUX consortium, decision numbers: 345130 and 345131. Alyson van Raalte was funded by the European Research Council, grant number 716323.

6 Author contributions and consent: All authors contributed to the study conception and design. Material preparation, data collection, and analysis were performed by Daniel Schneider. The first draft of the manuscript was written by Daniel Schneider and Mikko Myrskylä. All authors contributed substantially to subsequent revisions. All authors read and approved the final manuscript.

7 We thank Hal Caswell and Silke van Daalen for helpful discussions and critical comments.

8 Data availability: All data used in this paper are available free of charge. Our main data source is the JEP version 6.0.0 (DOI: 10.6103/SHARE.jep.600), which is derived mainly from Wave 3 of SHARE and complemented by information from Waves 1 and 2 (DOI: 10.6103/SHARE.w1.600, DOI: 10.6103/SHARE.w2.600, and DOI: 10.6103/SHARE.w3.600).

9 Alyson van Raalte can be found on Twitter: (@AlysonVanRaalte).

Disclosure statement

No potential conflict of interest was reported by the authors.

ORCID

Daniel C. Schneider http://orcid.org/0000-0003-1615-2482
Mikko Myrskylä http://orcid.org/0000-0003-4995-027X
Alyson van Raalte http://orcid.org/0000-0002-0676-8921

References

Allison, Paul D. 1982. Discrete-time methods for the analysis of event histories, *Sociological Methodology* 13: 61. https://doi.org/10.2307/270718

Andreev, Evgeny M., and W. Ward Kingkade. 2015. Average age at death in infancy and infant mortality level: Reconsidering the Coale-Demeny formulas at current levels of low mortality, *Demographic Research* 33: 363–390. https://doi.org/10.4054/DemRes.2015.33.13

Antonova, Liudmila, Luis Aranda, Giacomo Pasini, and Elisabetta Trevisan. 2014. Migration, family history and pension: The second release of the SHARE Job Episodes Panel, *SHARE Working Paper Series*: 18–2014.

Bardo, Anthony R., and Scott M. Lynch. 2021. Cognitively intact and happy life expectancy in the United States, *The Journals of Gerontology: Series B* 76(2): 242–251. https://doi.org/10.1093/geronb/gbz080

Börsch-Supan, Axel, Martina Brandt, Christian Hunkler, Thorsten Kneip, Julie Korbmacher, Frederic Malter, Barbara Schaan, et al. 2013. Data resource profile: The Survey of Health, Ageing and Retirement in Europe (SHARE), *International Journal of Epidemiology* 42(4): 992–1001. https://doi.org/10.1093/ije/dyt088

Brouard, Nicolas. 2019. Theory and applications of backward probabilities and prevalences in cross-longitudinal surveys, in A. S. R. S Rao and C. R. Rao (eds), *Handbook of Statistics. Volume 40: Integrated Population Biology and Modeling, Part B*. Elsevier, pp. 435–486.

Brouard, Nicolas. 2021. *Computing Health Expectancies using IMaCh: A Maximum Likelihood Computer Program using Interpolation of Markov Chains* [computer program], version 0.99r23. Available: https://euroreves.ined.fr/imach/

Brugiavini, Agar, Danilo Cavapozzi, Giacomo Pasini, and Elisabetta Trevisan. 2013. Working Life Histories from SHARELIFE: A Retrospective Panel, *SHARE Working Paper Series* 11–2013.

Cai, Liming, Mark Hayward, Yasuhiko Saito, James Lubitz, Aaron Hagedorn, and Eileen Crimmins. 2010. Estimation of multi-state life table functions and their variability from complex survey data using the SPACE program, *Demographic Research* 22: 129–158. https://doi.org/10.4054/DemRes.2010.22.6

Caswell, Hal. 2011. Beyond R0: Demographic models for variability of lifetime reproductive output, *PLoS ONE* 6(6): e20809. https://doi.org/10.1371/journal.pone.0020809

Caswell, Hal and Fanny Annemarie Kluge. 2015. Demography and the statistics of lifetime economic transfers under individual stochasticity, *Demographic Research* 32: 563–588. https://doi.org/10.4054/DemRes.2015.32.19

Caswell, Hal, Charlotte de Vries, Nienke Hartemink, Gregory Roth, and Silke F. van Daalen. 2018. Age × stage-classified demographic analysis: A comprehensive approach, *Ecological Monographs* 88(4): 560–584. https://doi.org/10.1002/ecm.1306

Caswell, Hal and Virginia Zarulli. 2018. Matrix methods in health demography: A new approach to the stochastic analysis of healthy longevity and DALYs, *Population Health Metrics* 16: 8. https://doi.org/10.1186/s12963-018-0165-5

Caswell, Hal and Silke van Daalen. 2021. Healthy longevity from incidence-based models: More kinds of health than stars in the sky, *Demographic Research* 45: 397–452. https://doi.org/10.4054/DemRes.2021.45.13

Chiang, Chin Long. 1968. *Introduction to Stochastic Processes in Biostatistics*. New York, NY: Wiley.

Chiang, Chin Long. 1972. On constructing current life tables, *Journal of the American Statistical Association* 67(339): 538–541. https://doi.org/10.1080/01621459.1972.10481245

Chiu, Chi-Tsun. 2019. Living arrangements and disability-free life expectancy in the United States, *PLoS ONE* 14 (2): e0211894. https://doi.org/10.1371/journal.pone.0211894

Cook, Richard J. and Jerald F. Lawless. 2018. *Multistate Models for the Analysis of Life History Data*. New York, NY: Chapman and Hall/CRC.

van Daalen, Silke and Hal Caswell. 2015. Lifetime reproduction and the second demographic transition: Stochasticity and individual variation, *Demographic Research* 33: 561–588. https://doi.org/10.4054/DemRes.2015.33.20

van Daalen, Silke F. and Hal Caswell. 2017. Lifetime reproductive output: Individual stochasticity, variance, and sensitivity analysis, *Theoretical Ecology* 10(3): 355–374. https://doi.org/10.1007/s12080-017-0335-2

Dudel, Christian and Mikko Myrskylä. 2020. Estimating the number and length of episodes in disability using a Markov chain approach, *Population Health Metrics* 18(1): 15. https://doi.org/10.1186/s12963-020-00217-0

Dudel, Christian. 2021. Expanding the Markov chain toolbox: Distributions of occupation times and waiting times, *Sociological Methods & Research* 50(1): 401–428. https://doi.org/10.1177/0049124118782541

Du Pasquier, Louis-Gustave. 1913. Mathematische Theorie der Invaliditätsversicherung [Mathematical theory of disability insurance], *Zeitschrift für Schweizerische Statistik* 49: 161–209. https://www.sgvs.ch/papers/sjesBackIssues/1913_PDF/1913-I-13.pdf

Farina, Mateo P., Mark D. Hayward, Jung Ki Kim, and Eileen M. Crimmins. 2020. Racial and educational disparities in dementia and dementia-free life expectancy, *The Journals of Gerontology: Series B* 75(7): e105–e112. https://doi.org/10.1093/geronb/gbz046

Graunt, John. [1662] 1964. Natural and political observations made upon the bills of mortality, *Journal of the Institute of Actuaries* 90: 1–61. https://doi.org/10.1017/S002026810001564X

Guilkey, David K. and Ronald R. Rindfuss. 1987. Logistic regression multivariate life tables: A communicable approach, *Sociological Methods & Research* 16(2): 276–300. https://doi.org/10.1177/0049124187016002004

Halley, Edmond. [1693] 1874. An estimate of the degrees of the mortality of mankind, drawn from curious tables of the births and funerals at the city of Breslaw; With an attempt to ascertain the price of annuities upon lives, *Journal of the Institute of Actuaries and Assurance Magazine* 18(4): 251–262. https://doi.org/10.1017/S2046167400044712

Hougaard, Philip. 2012. *Analysis of Multivariate Survival Data*. New York: Springer Science & Business Media.

van den Hout, Ardo. 2017. *Multi-State Survival Models for Interval-Censored Data*. New York: CRC Press.

Howard, Ronald A. 1960. *Dynamic Programming and Markov Processes.* New York, NY: MIT, Wiley.

Jenkins, Stephen P. 1995. Easy estimation methods for discrete-time duration models, *Oxford Bulletin of Economics and Statistics* 57(1): 129–136. https://doi.org/10.1111/j.1468-0084.1995.tb00031.x

Keyfitz, Nathan and Hal Caswell. 2005. *Applied Mathematical Demography*. 3rd ed. New York, NY: Springer.

Land, Kenneth C. and Andrei Rogers. 1982. *Multidimensional Mathematical Demography.* Proceedings of the Conference on Multidimensional Mathematical Demography Held at the University of Maryland, College Park, Maryland, March 23–25, 1981, Sponsored by the National Science Foundation, Elsevier.

Leinonen, Taina, Pekka Martikainen, and Mikko Myrskylä. 2018. Working life and retirement expectancies at age 50 by social class: Period and cohort trends and projections for Finland, *The Journals of Gerontology: Series B* 73(2): 302–313. https://doi.org/10.1093/geronb/gbv104

Lièvre, Agnès, Nicolas Brouard, and Christopher Heathcote. 2003. The estimation of health expectancies from cross-longitudinal surveys, *Mathematical Population Studies* 10 (4): 211–248. https://doi.org/10.1080/713644739

Loichinger, Elke and Daniela Weber. 2016. Trends in working life expectancy in Europe, *Journal of Aging and Health* 28(7): 1194–1213. https://doi.org/10.1177/0898264316656509

Lorenti, Angelo, Christian Dudel, and Mikko Myrskylä. 2019. The legacy of the Great Recession in Italy: A wider geographical, gender, and generational gap in working life expectancy, *Social Indicators Research* 142: 283–303. https://doi.org/10.1007/s11205-018-1910-7

Lynch, Scott M. and J. Scott Brown. 2005. A new approach to estimating life tables with covariates and constructing interval estimates of life table quantities, *Sociological Methodology* 35(1): 177–225. https://doi.org/10.1111/j.0081-1750.2006.00168.x

Magnusson Hanson, Linda L., Hugo Westerlund, Holendro S. Chungkham, Jussi Vahtera, Naja H. Rod, Kristina Alexanderson, Marcel Goldberg, et al. 2018. Job strain and loss of healthy life years between ages 50 and 75 by sex and occupational position: Analyses of 64 934 individuals from four prospective cohort studies, *Occupational and Environmental Medicine* 75(7): 486–493. https://doi.org/10.1136/oemed-2017-104644

Millimet, Daniel L., Michael Nieswiadomy, Hang Ryu, and Daniel Slottje. 2003. Estimating worklife expectancy: An econometric approach, *Journal of Econometrics* 113(1): 83–113. https://doi.org/10.1016/S0304-4076(02)00168-9

Orso, C. E., Agar Brugiavini, Danilo Cavapozzi, and Giacomo Pasini. 2017. SHARE Job Episodes Panel. Release Version: 6.0.0 SHARE-ERIC. Dataset. https://doi.org/10.6103/SHARE.jep.600

Owoeye, Seyifunmi M., Bamidele M. Oseni, and Ezra Gayawan. 2020. Estimating lifetime malnourished period and its statistics based on the concept of Markov chain with reward, *Heliyon* 6(5): e04073. https://doi.org/10.1016/j.heliyon.2020.e04073

Preston, Samuel H., Patrick Heuveline, and Michel Guillot. 2001. *Demography: Measuring and Modeling Population Processes.* Oxford: Blackwell.

van Raalte, Alyson A. and Hal Caswell. 2013. Perturbation analysis of indices of lifespan variability, *Demography* 50(5): 1615–1640. https://doi.org/10.1007/s13524-013-0223-3

Rogers, Andrei. 1975. *Introduction to Multiregional Mathematical Demography.* New York, NY: Wiley.

Schoen, Robert. 1975. Constructing increment-decrement life tables, *Demography* 12(2): 313–324. https://doi.org/10.2307/2060768

Schoen, Robert. 1979. Calculating increment-decrement life tables by estimating mean durations at transfer from observed rates, *Mathematical Biosciences* 47(3–4): 255–269. https://doi.org/10.1016/0025-5564(79)90041-5

Schoen, Robert and Kenneth C. Land. 1979. A general algorithm for estimating a Markov-generated increment-decrement life table with applications to marital-status patterns, *Journal of the American Statistical Association* 74(368): 761–776. https://doi.org/10.1080/01621459.1979.10481029

Singer, Judith D. and John B. Willett. 2003. *Applied Longitudinal Data Analysis.* New York, NY: Oxford University Press.

Sonnenberg, Frank A. and John B. Wong. 1993. Commentary: Fine-tuning life-expectancy calculations using Markov processes: To the editor, *Medical Decision Making* 13(2): 170–172. https://doi.org/10.1177/0272989X9301300215

Steele, Fiona. 2011. Multilevel discrete-time event history models with applications to the analysis of recurrent employment transitions, *Australian & New Zealand Journal of Statistics* 53(1): 1–20. https://doi.org/10.1111/j.1467-842X.2011.00604.x

Tutz, Gerhard and Matthias Schmid. 2016. *Modeling Discrete Time-to-Event Data*. Cham: Springer International Publishing.

University of California, Berkeley, and Max Planck Institute for Demographic Research. n.d. Human Mortality Database. Available: www.mortality.org (accessed: August 2018).

Willekens, Frans J. 2014. *Multistate Analysis of Life Histories with R*. New York, NY: Springer Berlin Heidelberg.

Willekens, Frans J. and Hein Putter. 2014. Software for multistate analysis, *Demographic Research* 31: 381–420. https://doi.org/10.4054/DemRes.2014.31.14

Willett, John B. and Judith D. Singer. 1995. It's déjà vu all over again: Using multiple-spell discrete-time survival analysis, *Journal of Educational and Behavioral Statistics* 20(1): 41–67. https://doi.org/10.3102/10769986020001041

Wilmoth, J. R., K. Andreev, D. Jdanov, D. A. Glei, and Tim Riffe. 2020. Methods protocol for the human mortality database, Version 6. Available: www.mortality.org

Wolf, Douglas A. and Thomas M. Gill. 2009. Modeling transition rates using panel current-status data: How serious is the bias?, *Demography* 46(2): 371–386. https://doi.org/10.1353/dem.0.0057

Zaninotto, Paola and Andrew Steptoe. 2019. Association between subjective well-being and living longer without disability or illness, *JAMA Network Open* 2(7): e196870. https://doi.org/10.1001/jamanetworkopen.2019.6870