Flexible Transition Timing in Discrete-Time Multistate Life Tables using Markov Chains with Rewards

Daniel C. Schneider[*1], Mikko Myrskylä[1,2,3], Alyson van Raalte[1],
Hal Caswell[4], Silke van Daalen[4]

[1]Max Planck Institute for Demographic Research

[2]London School of Economics and Political Science

[3]University of Helsinki

[4]University of Amsterdam

October 31, 2019

Abstract

A standard tool in demographic research are continuous-time multistate life tables (MSLTs). More recently, similar methods have been advanced that are cast in discrete time. While these models use a discrete time grid, it is often useful to calculate derived magnitudes, like state occupation times, under assumptions that posit that transitions take place at other times, such as mid-period. We propose to utilize Markov chains with rewards as an intuitive way of modelling the timing of transitions. This has the following advantages: It allows a flexible modelling of transition timing; carries a manageable theoretical overhead; and is easy to apply. We illustrate the usefulness of rewards-based MSLTs with SHARE data for the estimation of working life expectancy using different retirement transition timings. We also demonstrate that, for the single-state case, the rewards-based MSLTs match traditional life table methods exactly. We provide R and Stata code for the method proposed.

*schneider@demogr.mpg.de

Introduction

The life course can be thought of as a realization of a stochastic process whereby individuals are subject to a risk 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 the cohort state expectancies. In a single state model, with that state being mortality, this aggregate would be the life expectancy. Key to this machinery, which can be thought of as a more general life table framework, is to understand its linkages with Markov theory. Any life table can be written as a Markov chain. Life tables follow the Markov assumption which posits that the probability of making a particular state transition only depends on the previous state, but not on the entire history of states occupied. This holds for all flavors of the traditional life table, such as single-decrement, multiple decrement, cause-deleted, and increment-decrement (multistate). We regard all of these models, whether formulated as a single decrement 'spreadsheet' life table, or with survival and transition expressed as a Markov chain transition matrix, as being part of the multistate life table (MSLT) family.

MSLTs have traditionally been modeled in continuous-time, using transition rates, which we refer to as "continuous-time MSLTs". Alternatively, since many of the underlying data sources - longitudinal studies and retrospective studies with intermittent observations - suggests modeling in discrete time, a new branch of "discrete-time MSLTs" have been developed. Rather than using rates as input, these are modeled with transition probabilities.

Any model that deals with unobserved transition times must make assumptions in that regard. Transitions in discrete-time MSLTs may, strictly speaking, only occur 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 in the age interval 20-25? Do labor 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 case of single-state life tables, (van Raalte and Caswell, 2013) suggested deducting half of the age interval from the calculated MSLT expectancy.[1] This approach is applicable in situations in which age intervals do not vary and there is only one non-absorbing state. It effectively assumes mid-

[1] A similar argument had also been made, implicitly or explicitly and in somewhat different contexts, by (Guilkey and Rindfuss, 1987; Sonnenberg and Wong, 1993).

interval transitions and therefore works well if the data roughly match this assumption, as it may, for example, for the timing of deaths. However, it can actually increase the bias if the assumption is not met. (Dudel, 2018) 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. We will refer to this procedure as "initial period deduction".[2] The initial period deduction has two limitations. First, it only applies to a regular age-grid, which does not, for example, cover the demographic 5-year grid that has the irregular childhood age intervals [0-1) and [1-5). Secondly, 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 article, we recur to the method of Markov Chain with Rewards (MCWR), first proposed in the MSLT context by (Caswell, 2011), in order to replicate the existing timing options for discrete-time MSLTs, address their shortcomings, and significantly expand the choice set. Continuous-time MSLTs embody all traditional life table techniques as special cases. For the discrete case, the same level of generality can be achieved once rewards are included in the calculations. The basic idea behind the usage of rewards is to combine probabilities of reaching certain states with probabilities of moving out of these states, and the flexible assignment of time rewards to origin and destination states. A more thorough exposition follows in the next section. Being discrete-time MSLTs, rewards-based MSLTs inherit all of their advantages (and shortcomings) from them. Their distinguishing feature is that they improve upon timing options in order to deliver more accurate results.

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, whose 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, Aging and Retirement in Europe (SHARE) on working life histories to estimate state expectancies for employment and retirement. We show that, due to the nature of retirement timings, the rewards-based MSLT is the only method that can accurately exploit information on state entry/exit timings. We again analyze the bias of the discrete-time MSLTs. Finally, we provide Stata and R code that implements the calculations to facilitate replication and broader use of the method.

[2] The continuous-time MSLT equivalent to this is the linear l(x) ("linear integration") assumption (Schoen, 1988, p. 78).

Background

Multistate life tables

The origins of the life table, which is probably the most well-known demographic tool, date back to 17th century (Graunt, 1964; Halley, 1693). Maybe somewhat surprisingly, generalizations to multiple alive states were not conceived until the 20th 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, 2013). Within multistate life tables (MSLTs), the key life table outcome, life expectancy, is generalized to state expectancies. For example, using MSLTs one can calculate the expected life time spent in marriage; the expected life time being unemployed; the expected pain-free life time; and so forth.

Both the traditional single-state life table[3] and the multistate tables are expressed in continuous time.[4] In addition, multistate tables typically make use of Markov theory and express the ideas using matrix notation. Being set 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, and the number of time units spent in a state in a certain age interval, and finally, state and overall life expectancies. 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 models, is based on assumptions on properties of the hazard. Semi-parametric models harness fully-fledged (i.e. precisely dated) life history data in order 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. What all of these models have in common is that they require a larger theoretical exposition as well as knowledge of matrix algebra and differential and integral calculus. Moreover, they are frequently computationally burdensome. They are therefore not always easily accessible to researchers.

More recently, another strand of demographic and epidemiological literature has emerged whose model formulations share many ideas with the above-mentioned MSLTs (Millimet et al., 2003; Hunter and Caswell, 2005). The starting point for the recent models, however, are not transition rates, but transition *probabilities*. While there are many textbooks on continuous-time MSLTs available (Hougaard, 2012; Willekens, 2014; van den Hout, 2017; Cook and Lawless, 2018), discrete-time MSLTs do not require precisely that: textbook-length treatment. It is one of their appealing features that they are easy to understand, communicate and apply. At the very simplest level, transition probabilities can be based on

[3] The single state is "alive".

[4] In some expositions this may be hidden behind notation that looks discrete. The underlying ideas, however, are unanimously cast in continuous time.

transition counts. Slightly more sophisticated is the usage of a multinomial logit model, but this is still part of the skill set of any quantitative social scientist. For the Markov calculations that use the probabilities as an input, only knowledge of matrix multiplication is required, one of the most basic operations on matrices. Matrix formulas for obtaining state expectancies are outlined in the methods section below, and explained in greater detail in the supplementary appendix. Moreover, aside from the bootstrap procedure used to obtain standard errors, the computational cost is typically very low.

There are modifications of the above procedure, such as the IMaCh method of (Lièvre et al., 2003), which can account for interval censoring (i.e. for the lack of information on transitions between observational points); the SPACE method of (Cai et al., 2010), which can compute a large number of statistics harnessing simulation techniques; and (Lynch and Brown, 2005), who take a Bayesian approach. While this paper will ignore these extensions, its results are at least partially applicable to these methods too.

The choice between a continuous-time or discrete-time model frequently boils down to an assessment of whether the continuous-time model can substantially make better use of the information, and whether some 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 usage 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 of interview dates within a few weeks may pose no problem, but the situation is different for a survey whose wave spacing varies between 2 and 5 years. In case of doubt, it is always a possibility to cross-check discrete-time results against continuous-time ones. The important point to grasp is that, if simplifying assumptions seem innocuous, discrete-time models are an attractive choice because they are easy to understand, apply, and communicate.

Markov Chains with Rewards

The model is based on an age-stage Markov chain whose state space is comprised of the Cartesian product of $\bar{x}$ age classes and $\bar{s}$ stages. Since the term "stage" implies some sort of sequencing, which is not appropriate for many applications, we opt for slightly ambiguous terminology and use "state" in place of "stage". Later on, in the retirement example, states are "employed", "unemployed/out of labor 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 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 easily generalize to irregularly spaced age intervals. The Markov assumption is that the process is memoryless, i.e. that the probability of a particular transition depends only on the current state, but not on the history of previous state visits.

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 the appendix) will be more suitable for our purposes. Let $\boldsymbol{U}$ denote the transition matrix of the process:

$$\boldsymbol{U} = \begin{pmatrix} \boldsymbol{p_{11}} & \cdots & \boldsymbol{p_{1\bar{s}}} \\ \vdots & \ddots & \vdots \\ \boldsymbol{p_{\bar{s}1}} & \cdots & \boldsymbol{p_{\bar{s}\bar{s}}} \end{pmatrix}$$

where $\boldsymbol{p_{ij}}$ denotes a $\bar{x} \times \bar{x}$ matrix with non-zero elements on the first subdiagonal only:

$$\boldsymbol{p_{ij}} = \begin{pmatrix} 0 & 0 & \cdots & 0 & 0 \\ p_{ij,2} & 0 & \cdots & 0 & 0 \\ 0 & p_{ij,3} & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & p_{ij,\bar{x}} & 0 \end{pmatrix}$$

Element $p_{ij,k}$ denotes the probability of moving from state s_j to state s_i when entering age class k. An important matrix in Markov chain theory and in MSLTs is the "fundamental matrix", which is calculated as

$$\boldsymbol{F} = (\boldsymbol{I_{\bar{x}\bar{s}}} - \boldsymbol{U})^{-1}$$

where $\boldsymbol{I_{\bar{x}\bar{s}}}$ denotes an $\bar{x}\bar{s} \times \bar{x}\bar{s}$ identity matrix and the power of minus one stands for the matrix inverse. 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 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 of subjects starting at age 50, an element of $\boldsymbol{F}$ may indicate that a person is expected to be 0.8 years in the employed state at age 60 if her initial age-state was being employed at age 50. Put differently, this element says that for a large number of individuals, 80% of those who are initially employed at age 50 will be employed at age 60. The remaining 20% are either in the absorbing state "dead" or in one of the alternative non-absorbing states (unemployed/out of labor 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.

Weighing these magnitudes by the initial state distribution yields state expectancies independent of the initial state.

The above procedure yields state expectancies based on the assumption of end-of-period transitions. The initial period deduction method uses $\widetilde{\boldsymbol{F}} = \boldsymbol{F} - \frac{n}{2}\boldsymbol{I}_{\overline{xs}}$ instead of $\boldsymbol{F}$, which corresponds to the assumption of mid-period transitions. In order to add more flexibility with respect to transition timing, we introduce, for each state s_m, rewards matrices $\boldsymbol{R_m}$. Their structure is very similar to $\boldsymbol{U}$ and its submatrices $\boldsymbol{p_{ij}}$. Two matrices $\boldsymbol{R_i}$ and $\boldsymbol{R_j}$ specify, for each transition, two time rewards, one towards the state of origin and one towards the destination state. The rewards-based calculation of the expected time spent in state s links the information on the expected time spent in a certain state at a certain age embodied in the matrix $\boldsymbol{F}$ with the rewards towards a particular state when moving out of the age-state, embodied in the elementwise matrix product $\boldsymbol{P} \circ \boldsymbol{R_m}$. To continue with the above example, if there is a probability of retiring of 0.1 during age 60 when being employed at the 60th birthday, and on average those who retire work for 0.3 years before retiring, the transition from employment to retirement during age 60 contributes $0.8 \times 0.1 \times 0.3 = 0.024$ years to the working life expectancy. If the only additional state were 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 the retirement life expectancy. Likewise, the transition to staying employed 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 exposition in this section has been brief and geared towards an intuitive understanding. The supplementary appendix contains a thorough description of the full methodology and formulas involved.

Empirical Illustrations

We illustrate two interesting aspects of the rewards methodology with empirical applications. The first one touches upon 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 MSLTs incorporate additional information, thereby refining 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 Panel A and B, left-vertical-axes, use data from the Human Mortality Database (University of California, Berkeley and Max Planck Institute for Demographic Research, Germany, n.d.) with a one-year and five-year age spacing, respectively, to illustrate the magnitude and variation of error in the MSLT with end-of-period transitions when compared to life table calculations following standard methodology [ref HMD methods protocol]. As is standard in demographic tabulations, the five-year grid is not exactly regular as it employs the usual childhood intervals of ages 0 and 1-4. The magnitude and direction of bias in the life expectancy estimates for this example is mostly determined 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). Therefore, the end-of-period transition assumption that death occurs at age 1 introduces a strong upward bias. Hypothetically, the maximum bias for this interval is 1. The overall impact of the end-of-period transition assumption from infancy on life expectancy estimates depends on both the average age at death among those who died, as well as the magnitude of infant mortality. For higher life expectancy estimates, although infant deaths occur on average closer to the birth (Andreev and Kingkade), which leads to a larger bias per death, the higher survival probabilities reduce the total bias on 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. The result is a downward sloping data cloud for 1-year age grids, approaching a bias of roughly half of the age interval (0.5) from above as expectancies increase. The upward sloping data cloud of the 5-year spacing, which approaches a bias of 2.5 as expectancies increase, is explained by similar reasoning, taking into account that the first age interval is only one year wide, whereas higher age intervals are five years wide.

The right-hand side vertical axes in both panels 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 numbers 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 is available in the HMD. Furthermore, survivors of an age class are assigned a reward of $n = 1$ in the single-age case and $n_i \in \{1,4,5\}$ in the 5-year interval case. Section 1.2 in the supplementary appendix shows that the two approaches are fully numerically equivalent.

Working and Retirement 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 life and retirement expectancies at age 50 using the Survey of Ageing, Retirement and Health in Europe (SHARE, REF) and the accompanying Job

Episodes Panel (JEP, REF). Increasing life expectancy and increased recognition of the fact that actual retirement decisions differ from mandatory retirement ages has spurred research into how long people actually work, and how long they can enjoy retirement (Loichinger et al. 2016; Lorenti et al. 2018; Leinonen et al. 2015). Although retirement decisions are often only partially related to mandatory retirement ages, in many settings the amount of benefits depends on age such that after certain birthdays (for example, the 65th) the benefits may jump. Such incentives may result in transitions that occur early on within age intervals.

Our main data source is the Job Episodes Panel (JEP) version 6.0.0 [Antonova et al 2014, 2017; (Brugiavini et al., 2013); Orso et al. 2017; DOI: 10.6103/SHARE.jep.600]. JEP, in turn, is mainly derived from wave 3 of the Survey of Health, Ageing, and Retirement in Europe (SHARE) [Börsch-Supan 2018] and complemented by information from waves 1 and 2 (DOIs: 10.6103/SHARE.w1.600, 10.6103/SHARE.w2.600, 10.6103/SHARE.w3.600). SHARE is a representative longitudinal survey started in 2002 containing data on more than 120,000 individuals aged 50 and over. It is conducted for 16 EU member states and mainly funded by the European Commission, with additional funding from the German Ministry of Education and Research, the Max Planck Society for the Advancement of Science, the U.S. National Institute on Aging and from 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 was collected during 2008-09 for almost 30,000 subjects of 14 SHARE countries (Austria, Belgium, Czech Republic, Denmark, France, Germany, Greece, Ireland, Italy, Netherlands, Poland, Spain, Sweden, Switzerland). JEP is a retrospective long panel data set based on SHARELIFE information that features full working life histories for each SHARELIFE respondent. We supplement JEP by 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 who have no record beyond age 50, or who were 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 one 1,519 (women in Greece).

We use the above data and a multinomial logit model to estimate transition probabilities among employment states for each age. Estimation is stratified by sex and country. Since the JEP is entirely based on retrospective interviews, it does not contain mortality information. Therefore, we use data from the Human Mortality Database in order to calculate country-specific probabilities of dying by sex and single age over the period 1985-2004, a time span which covers [TODO: XX]% of our observations. We apply the resulting mortality conditions to all states and to all of 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. We also slightly reclassify employment states to resolve conflicting state information for a smaller fraction of observations. All analyses are unweighted.

Wave 2 data is used to calculate the average exact (fractional) age at retirement for each age, separately by sex and country. This information is not included in JEP and only available for a subset of individuals, which is not a problem in the present context since we only need 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 one minus that fraction as the reward towards retirement life expectancy.

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 early- to mid-60s - which is where the bulk of retirement transitions occur - often take place close to the beginning of the interval. This suggests that incorporating the transition timing information in the calculation of working and retirement expectancies is important.

Figure 3 shows total and state rewards-based life expectancies at age 50 for all and for individual SHARE countries, as well as the bias of other methods. Panel A contains numbers for men and Panel B for women. In each panel, the top graph displays total and component life expectancies calculated using rewards-based MSLTs. The rewards approach accurately assigns for each transition the time that individuals on average spend in the origin and destination states and delivers error free results. The middle and bottom graph in each panel compare other discrete-time methods to this gold standard. The middle graph depicts the difference in working life expectancy numbers between the rewards method and end-of-period calculations (diamonds), as well as the difference between rewards and the initial period deduction method, i.e. the method that subtracted half of the age interval from the diagonal of the fundamental matrix (triangles). The bottom graph of 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 expectancies (bottom graphs in each panel) the direction of the error is switched: The unadjusted end-of-period numbers are about 0.2 years short of the rewards numbers, i.e. they understate the length of retirement. Note that these numbers are not simply negatives of the ones for WLE 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, and not to zero. Secondly, the results for unemployed life expectancy (not shown) capture the remainder. Two features of the initial deduction-based retirement numbers catch the eye: They are almost identical to the unadjusted end-of-period numbers, and they sometimes even increase 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 the fraction of retired people at age 50 is very small, the deduction from retirement life expectancy is small, so the unadjusted and the adjusted numbers are close. The second feature is a simple consequence of the fact the end-of-period estimates are lower than the rewards ones, and any deduction further diminishes its values and increases the bias.

Conclusions

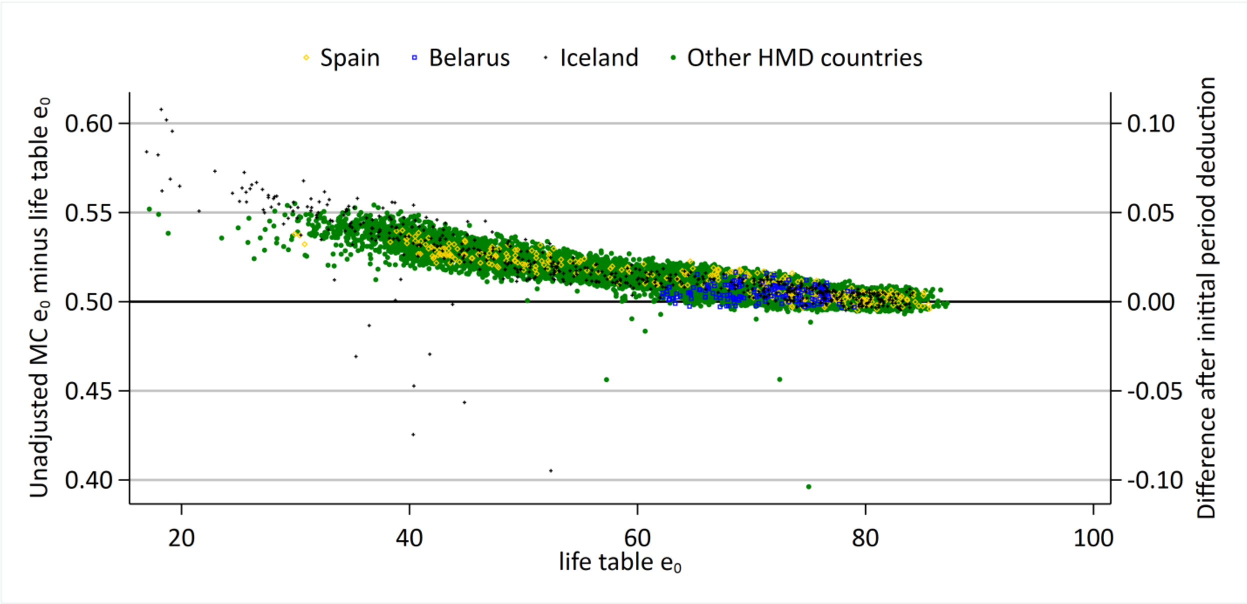
Discrete-time MSLT approaches to calculating state-specific expectancies are broadly used in demography. The standard approach implies an assumption of transitions taking place at the end of the interval. This results in a discrepancy between the Markov chain estimates and life table estimates of state expectancies. Using two high-quality data sets on life expectancy, working life expectancy and retirement expectancy, we have shown that the error can be non-negligible. The error, however, can be completely removed using the Markov chain with rewards approach, in which the researcher has full control over the timing of transitions. In the simplest case of one non-absorbing state, the Markov chain with rewards estimates are the same as with the standard life table calculations traditionally used by demographers. For 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 oftentimes more difficult to apply. Here the proposed rewards-based discrete-time method improves the accuracy of the state expectancy calculations over existing discrete-time methods. To remove the error in discrete-time MSLT state expectancies, we encourage users to incorporate rewards in their analysis. To facilitate replication and broader use of the rewards approach, this paper is supplemented by R and Stata code that implements the calculations.[5]

[5] [ADD LINK TO CODE]

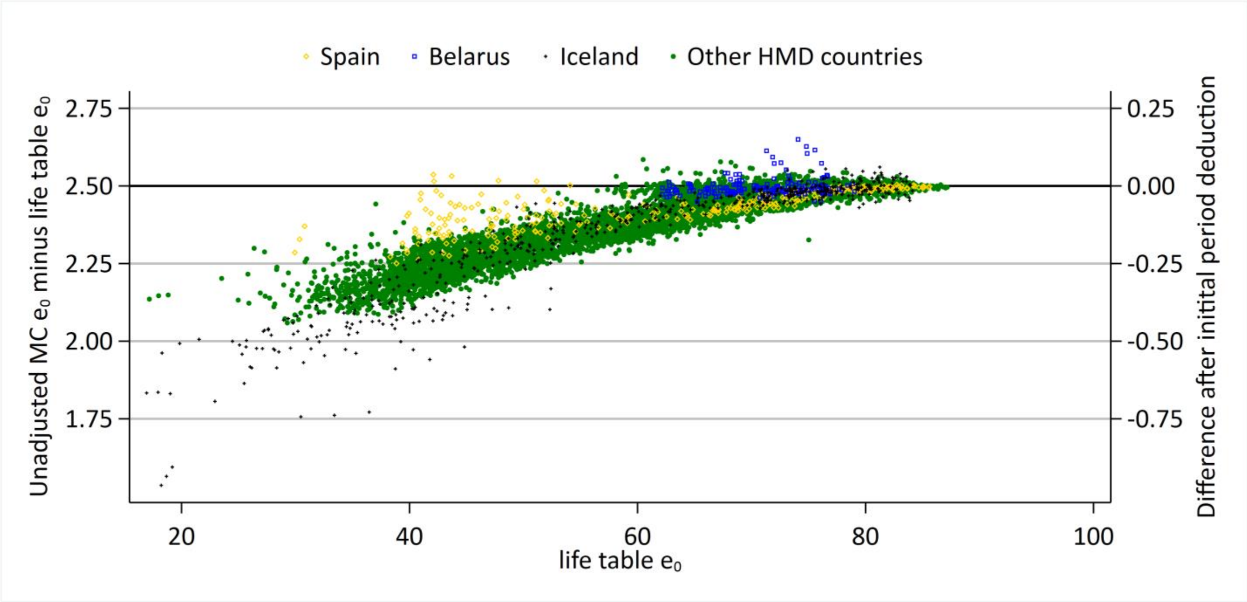
Figures and Tables

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

Panel A: 1-year age interval

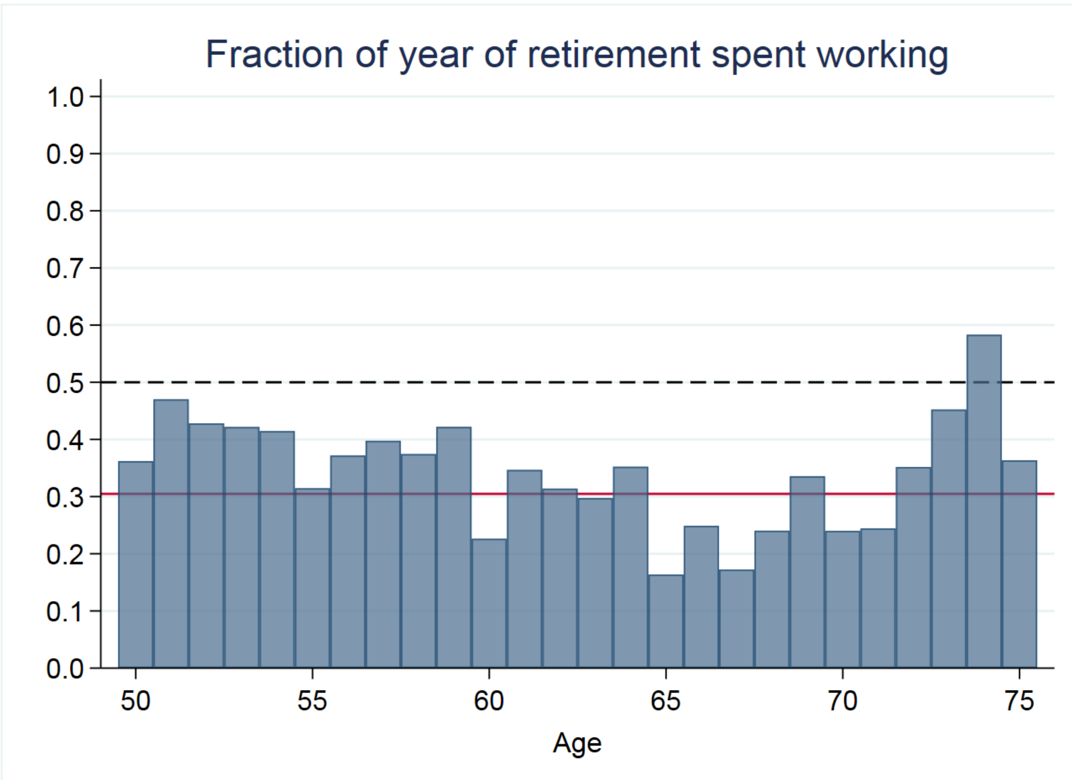


Panel B: 5-year age interval



Note: Data taken from the Human Mortality Database.

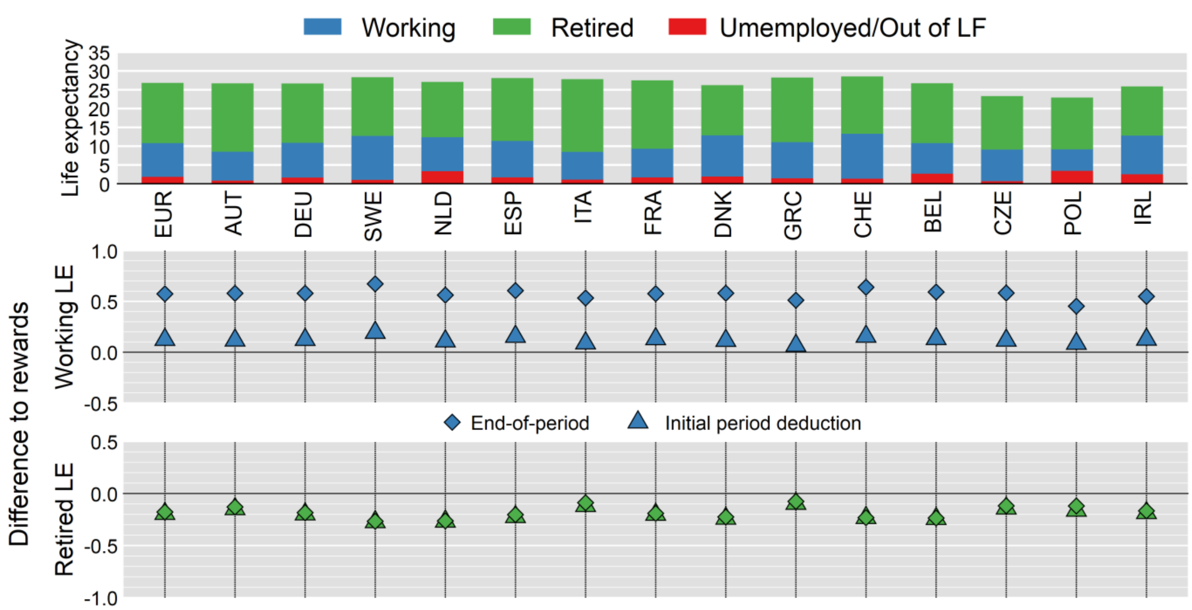
Figure 2. Fraction of the year of retirement spent working, by single age.



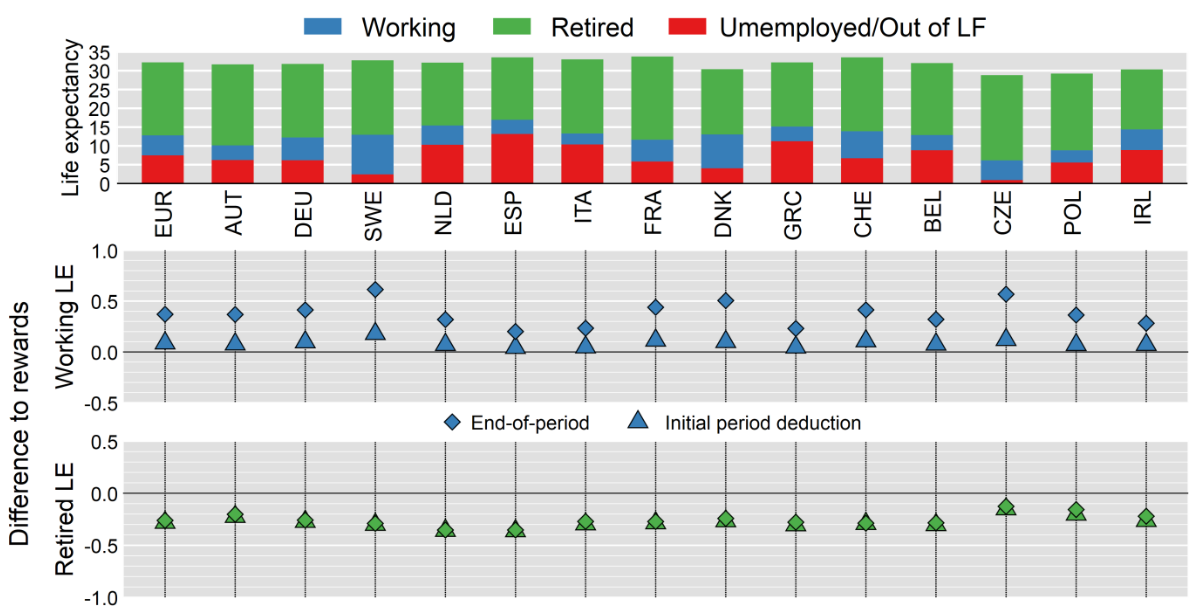
Note: Based on data for all countries from SHARE wave 2. The black dashed line corresponds to the mid-interval reward of 0.5 to each of the states of origin and destination. The red solid line depicts the overall average calculated over all retirement events in the data.

Figure 3. Working life expectancies at age 50 across SHARE countries calculated using rewards, and differences thereof with respect to other methods.

Panel A: Men



Panel B: Women



Note: Based on data for all countries from SHARE wave 2 and the job episodes panel from SHARELIFE. For each panel, the top graph shows total and component life expectancies. The middle graph depicts the difference in working life expectancy numbers 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.

References

Brugiavini, A., Cavapozzi, D., Pasini, G., Trevisan, E., 2013. Working life histories from SHARELIFE: a retrospective panel. SHARE-ERIC, Working Paper Series 11–2013, 14.

Cai, L., Hayward, M., Saito, Y., Lubitz, J., Hagedorn, A., Crimmins, E., 2010. Estimation of multi-state life table functions and their variability from complex survey data using the SPACE Program. Demogr. Res. 22, 129–158. https://doi.org/10.4054/DemRes.2010.22.6

Caswell, H., 2011. Beyond R0: Demographic Models for Variability of Lifetime Reproductive Output. PLoS ONE 6, e20809. https://doi.org/10.1371/journal.pone.0020809

Cook, R.J., Lawless, J.F., 2018. Multistate Models for the Analysis of Life History Data, Monographs on Statistics and Applied Probability. Chapman and Hall/CRC, New York, NY.

Du Pasquier, L.-G., 1912. Mathematische Theorie der Invaliditätsversicherung. Mitteilungen Ver. Schweitzerischer Versicherungsmathematiker 7, 1–7.

Dudel, C., 2018. Expanding the Markov Chain Toolbox: Distributions of Occupation Times and Waiting Times. Sociol. Methods Res. 004912411878254. https://doi.org/10.1177/0049124118782541

Graunt, J., 1964. Natural and Political Observations Mentioned in a Following Index, and Made Upon the Bills of Mortality. J. Inst. Actuar., (reprint: original text published in 1662) 90, 1–61.

Guilkey, D.K., Rindfuss, R.R., 1987. Logistic Regression Multivariate Life Tables: A Communicable Approach. Sociol. Methods Res. 16, 276–300. https://doi.org/10.1177/0049124187016002004

Halley, E., 1693. An Estimate of the Degrees of the Mortality of Mankind. Philos. Trans. 17, 596–610, 653–6.

Hougaard, P., 2012. Analysis of multivariate survival data. Springer Science & Business Media.

Hunter, C.M., Caswell, H., 2005. The use of the vec-permutation matrix in spatial matrix population models. Ecol. Model. 188, 15–21. https://doi.org/10.1016/j.ecolmodel.2005.05.002

Land, K.C., Rogers, A., 2013. 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.

Lièvre, A., Brouard, N., Heathcote, C., 2003. The Estimation of Health Expectancies from Cross-Longitudinal Surveys. Math. Popul. Stud. 10, 211–248. https://doi.org/10.1080/713644739

Lynch, S.M., Brown, J.S., 2005. A New Approach to Estimating Life Tables with Covariates and Constructing Interval Estimates of Life Table Quantities 49.

Millimet, D.L., Nieswiadomy, M., Ryu, H., Slottje, D., 2003. Estimating worklife expectancy: an econometric approach. J. Econom. 113, 83–113. https://doi.org/10.1016/S0304-4076(02)00168-9

Schoen, R., 1975. Constructing Increment-Decrement Life Tables. Demography 12, 313. https://doi.org/10.2307/2060768

Schoen, R., Land, K.C., 1979. A General Algorithm for Estimating a Markov-Generated Increment-Decrement Life Table with Applications to Marital-Status Patterns. J. Am. Stat. Assoc. 74, 17.

Sonnenberg, F.A., Wong, J.B., 1993. Commentary: Fine-tuning Life-espectancy Calculations Using Markov Processes: To the Editor. Med. Decis. Making 13, 170–172. https://doi.org/10.1177/0272989X9301300215
University of California, Berkeley, Max Planck Institute for Demographic Research, Germany, n.d. Human Mortality Database.
van den Hout, A., 2017. Multi-State Survival Models for Interval-Censored Data, Monographs on Statistics and Applied Probability. CRC Press.
van Raalte, A.A., Caswell, H., 2013. Perturbation Analysis of Indices of Lifespan Variability. Demography 50, 1615–1640. https://doi.org/10.1007/s13524-013-0223-3
Willekens, F.J., 2014. Multistate analysis of life histories with R. Springer Berlin Heidelberg, New York, NY.

Appendix Contents

1 Markov Chain Calculations

1.1 General Exposition

1.1.1 Expositional Setup and Notational Conventions

The purpose of this appendix is to present the main formula for Markov chains with rewards and subsequently examine its individual parts in order to gain a deeper understanding of it. Relating the rewards expressions to traditional life table formulas will then be easy. The presentation aims to do this in a way that makes the method and its formulas accessible to readers who only have had little exposure to matrix algebra. The only prerequisite is knowledge of matrix multiplication.

One of the great benefits of the usage of matrices and vectors is the compactness of expressions. However, given the above mentioned goals we opted for a more verbose presentation that uses explicit indexing. Explicitly refering to matrix elements - for example, in order to write out sums of products from matrix multiplications - requires usage of multiple subscripts. This necessarily results in notation that is somewhat tedious. Nevertheless, in the present context this can hardly be avoided, given the aims of the exposition. Readers that are more familiar with matrix algebra may also wish to consult [Caswell and van Daalen, 2019] (manuscript) for a more succinct presentation of the formulas involved.

In Markov models that incorporate both age and status, there are two sensible orderings of elements of matrices and vectors: age-within-status or status-within-age. Here we opt for the age-within-status ordering. The reason is that summations for conditional state expectancies neatly correspond to columns of matrix (2). Needless to say, for the application of the main rewards formula it does not matter which ordering is used.

We assume a single absorbing state (death). The case of multiple absorbing states is not difficult to deduce. We opt for this simplification in order to not further complicate notation, which is already somewhat unwieldy.

Matrices and vectors are denoted in boldface. To distinguish vectors from matrices, they have an arrow decoration.

The following table lists the symbols used:

symbol	meaning
$\bar{x}$	number of age classes
k	age class index, $k = 1, , \dots, \bar{x}$
x_k	beginning age of age class k

x_1	minimum (baseline) age in model. x_1 is the age of the initial state. x_2 is the age at which the first transition takes place
Ξ	set of age classes: $\{x_k \vert k = 1, \ldots, \bar{x}\}$
n	length of age interval
$\bar{s}$	number of *transient* states
S	set of transient states, $S = \{s_i \vert i = 1, \ldots, \bar{s}\}$
i, j	state indexes used for transitions
m	state index, used for rewards
d	the (single) dead state
$\tilde{S}$	set of all states, i.e. transient states and dead: $\tilde{S} = S \cup \{d\}$
$\mathbf{P}$	transition matrix of all states
$\mathbf{U}$	submatrix of $\mathbf{P}$ consisting of transient states only
$\mathbf{F}$	fundamental matrix
$\mathbf{R}_m$	rewards matrix for state s_m
g_j	fraction of population in state s_j at baseline age
e_{x_k}	life expectancy at age x_k
$e_{x_k \vert j}$	life expectancy at age x_k, conditional on initial state s_j
$e^i_{x_k}$	expected lifetime spent in state s_i, starting at age x_k
$e^i_{x_k \vert j}$	expected lifetime spent in state s_i, starting at age x_k, and conditional on initial state s_j
$\otimes$	Kronecker product
$\circ$	Hadamard product
I_c	identity matrix of dimension $c \times c$

Pairs of state indices show the index of the destination state first, followed by the index of the state of origin. For example $p_{s_i s_j}$ indicates the probability of moving from state j to state i. Markov matrices are sometimes stated as column-stochastic matrices (columns sum to one) and sometimes as row-stochastic matrices (rows sum to one). We use column-stochastic matrices. The first index i always refers to rows, and the second index j to columns.

To ease notation, we make use of the following notational shorthands: When an expression has multiple nested subscripts, we frequently only use m, i, j in place of s_m, s_i, s_j, respectively, and k in place of x_k. For example, we write p_{ij} instead of $p_{s_i s_j}$. To give another example, as we will see in section 1.1.3, the scalar $r_{s_m, s_i s_j, x_k}$ denotes the reward towards state s_m when transitioning from state s_j to s_i at age x_k. We will simply write $r_{m,ij,k}$ for it. The consistent usage of all index symbols (m, i, j, k, ...) hopefully aides in keeping expressions transparent. Finally, simplified notation is also used for set membership: Instead of writing $s_j \in S$, we write $j \in S$, and, similarly, $k \in \Xi$.

1.1.2 Standard Calculations (No Rewards)

In general, for the age-within-status ordering all matrices can be thought of as being mainly composed of $\bar{s}^2$ blocks or submatrices, each of dimension $\bar{x} \times \bar{x}$. For example, the transition matrix $\mathbf{P}$ is

$$\mathbf{P} = \begin{pmatrix} \mathbf{p}_{11} & \cdots & \mathbf{p}_{1\bar{s}} & \mathbf{0} \\ \vdots & \ddots & \vdots & \vdots \\ \mathbf{p}_{\bar{s}1} & \cdots & \mathbf{p}_{\bar{s}\bar{s}} & \mathbf{0} \\ \overrightarrow{\mathbf{p}}_{d1} & \cdots & \overrightarrow{p}_{d\bar{s}} & 1 \end{pmatrix}$$

where each submatrix $\mathbf{p}_{ij}$is a square $\bar{x} \times \bar{x}$ matrix. Each submatrix has nonzero elements only on the first subdiagonal:

$$\mathbf{p}_{ij} = \begin{pmatrix} 0 & 0 & \cdots & 0 & 0 \\ p_{ij,2} & 0 & \cdots & 0 & 0 \\ 0 & p_{ij,3} & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & p_{ij,\bar{x}} & 0 \end{pmatrix}$$

where $p_{ij,k}$ denotes the probability of making the transition from state s_j to state s_i when moving from age class $k-1$ to age class k. Submatrices $\overrightarrow{\mathbf{p}}_{dj}$ are not square matrices but $1 \times \bar{x}$ row vectors, which is indicated by the arrow decoration. If there were multiple absorbing states in the model (e.g. for a multiple decrement application), the vector would turn into a matrix with a corresponding number of rows. For ease of exposition, this generalization is omitted here.

The vectors contain probabilities of dying for each age x_k:

$$\overrightarrow{\mathbf{p}}_{dj} = \begin{pmatrix} p_{dj,2} & \cdots & p_{dj,\bar{x}} & p_{dj,\bar{x}+1} \end{pmatrix}$$

where the last element, $p_{dj,\bar{x}+1}$ is equal to one: By assumption, subjects die during the transition from age class $\bar{x}$ to age class $\bar{x}+1$. The last column of $\mathbf{P}$ is a column of zeroes with only the last element being equal to one, indicating the absorbing nature of death. The structure of the full matrix $\mathbf{P}$, using row and column labels in boldface, is then

$\mathbf{i \backslash j}$		$\cdots$	$\cdots$	$\mathbf{1}$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\bar{\mathbf{s}}$	$\cdots$	$\mathbf{d}$
	$\mathbf{k}$	$\mathbf{1}$	$\cdots$	$\bar{\mathbf{x}}-\mathbf{1}$	$\bar{\mathbf{x}}$	$\cdots$	$\cdots$	$\mathbf{1}$	$\cdots$	$\bar{\mathbf{x}}-\mathbf{1}$	$\bar{\mathbf{x}}$	.
$\vdots$	$\mathbf{1}$											
$\mathbf{1}$	$\mathbf{2}$	p				$\cdots$	$\cdots$	p				
$\vdots$	$\vdots$		$\ddots$						$\ddots$			
$\vdots$	$\bar{\mathbf{x}}$			p						p		
$\vdots$	$\vdots$			$\vdots$		$\ddots$				$\vdots$		
$\vdots$	$\vdots$			$\vdots$			$\ddots$			$\vdots$		
$\vdots$	$\mathbf{1}$											
$\bar{\mathbf{s}}$	$\mathbf{2}$	p						p				
$\vdots$	$\vdots$		$\ddots$			$\cdots$	$\cdots$		$\ddots$			
$\vdots$	$\bar{\mathbf{x}}$			p						p		
$\mathbf{d}$	.	p	p	p	1	$\cdots$	$\cdots$	p	p	p	1	1

with p indicting a nonzero probability, $p \in [0,1]$, and zeroes elsewhere. Each column of $\mathbf{P}$ sums to one.

Premultiplying a $(\bar{x} \cdot \bar{s} + 1) \times 1$ state distribution vector $\mathbf{y}_t$ by $\mathbf{P}$ generates the state vector $\mathbf{y}_{t+1}$, i.e. $\mathbf{P}$ moves counts (or fractions) of individuals from states at time t to expected states in time $t+1$, while simultaneously aging the individuals by one time unit.[1]

[1]This intuitive perspective of advancement by one time unit is only applicable if age classes are regularly spaced.

$\mathbf{U}$ is simply a submatrix of $\mathbf{P}$ that only contains transitions among transient states:

$$\mathbf{U} = \begin{pmatrix} \mathbf{p}_{11} & \cdots & \mathbf{p}_{1\bar{s}} \\ \vdots & \ddots & \vdots \\ \mathbf{p}_{\bar{s}1} & \cdots & \mathbf{p}_{\bar{s}\bar{s}} \end{pmatrix}$$

The fundamental matrix $\mathbf{F}$ is calculated as

$$\mathbf{F} = \left(\mathbf{I}_{\bar{\mathbf{x}} \cdot \bar{\mathbf{s}}} - \mathbf{U}\right)^{-1}$$

where $\mathbf{I}$ denotes the identity matrix. The easiest way to understand the above calculation is by remembering the basic geometric series result for $a \in [0, 1)$

$$\lim \sum_{c=0}^{\infty} a^c = \frac{1}{1-a}$$

The formula for $\mathbf{F}$ is simply the corresponding matrix version:

$$\lim \sum_{c=0}^{\infty} \mathbf{U}^c = \lim \left(\mathbf{I} + \mathbf{U} + \mathbf{U}^2 + \mathbf{U}^3 + \dots\right) = \left(\mathbf{I}_{\bar{\mathbf{x}} \cdot \bar{\mathbf{s}}} - \mathbf{U}\right)^{-1}$$

Postmultiplying any power of $\mathbf{U}^c$ by $\mathbf{U}$ again has the effect of shifting the nonzero elements of $\mathbf{U}^c$ one subdiagonal down within each submatrix $\mathbf{p}_{ij}$. For example, $\mathbf{U} \cdot \mathbf{U} = \mathbf{U}^2$ has the submatrices

$$\mathbf{p}_{ij} = \begin{array}{cc|cccccc} & & \cdots & \cdots & \cdots & \mathbf{s_j} & \cdots & \cdots \\ & \mathbf{k} & \mathbf{1} & \mathbf{2} & \cdots & \bar{\mathbf{x}}-\mathbf{2} & \bar{\mathbf{x}}-\mathbf{1} & \bar{\mathbf{x}} \\ \hline \vdots & \mathbf{1} & & & & & & \\ \vdots & \mathbf{2} & & & & & & \\ \vdots & \mathbf{3} & p & & & & & \\ \mathbf{s_i} & \mathbf{4} & & p & & & & \\ \vdots & \vdots & & & \ddots & & & \\ \vdots & \bar{\boldsymbol{x}} & & & & p & & \end{array}$$

where the nonzero elements now denote transition probabilites over two time periods. $\mathbf{U}^3$ has submatrices with the nonzero elements shifted down by another subdiagonal, and so forth.[2] In general, $\mathbf{U}^c$ has submatrices with nonzero elements on the c-th subdiagonal, denoting c-period transition probabilities. As a consequence, we have

$$\mathbf{U}^c = \mathbf{0}, \ \forall c \geq \bar{x}$$

so the result from the geometric series obtains exactly, not just in the limit:

$$\mathbf{F} = \sum_{c=0}^{\bar{x}-1} \mathbf{U}^c \tag{1}$$

In the preceding formula, each term $\mathbf{U}^c$ adds another subdiagonal of nonzero elements to each submatrix $\mathbf{f}_{ij}$, so that it is easy to see that $\mathbf{F}$ must have the following structure:

[2] This does not mean that the elements of the relevant subdiagonal of $\mathbf{p}_{ij}$ of $\mathbf{U}^{c+1}$ only depend on the elements of the relevant subdiagonal of $\mathbf{p}_{ij}$ of $\mathbf{U}^c$. This is not the case. They depend on the relevant subdiagonals of all matrices $\mathbf{p}_{\cdot j}$ and $\mathbf{p}_{i \cdot}$ of $\mathbf{U}^c$.

$$
\mathbf{F} = \begin{pmatrix} \mathbf{f}_{11} & \cdots & \mathbf{f}_{1\bar{s}} \\ \vdots & \ddots & \vdots \\ \mathbf{f}_{\bar{s}1} & \cdots & \mathbf{f}_{\bar{s}\bar{s}} \end{pmatrix}
$$

$$
= \begin{array}{cc|cccc|cc|cccc}
\mathbf{i}\backslash\mathbf{j} & & \cdots & \cdots & \mathbf{1} & \cdots & \cdots & \cdots & \cdots & \cdots & \bar{\mathbf{s}} & \cdots \\
 & \mathbf{k} & \mathbf{1} & \cdots & \bar{\mathbf{x}}-\mathbf{1} & \bar{\boldsymbol{x}} & \cdots & \cdots & \mathbf{1} & \cdots & \bar{\mathbf{x}}-\mathbf{1} & \bar{\boldsymbol{x}} \\
\hline
\vdots & \mathbf{1} & 1 & & & & & & & & & \\
\mathbf{1} & \mathbf{2} & f & 1 & & & \cdots & \cdots & f & & & \\
\vdots & \vdots & \vdots & \ddots & 1 & & & & \vdots & \ddots & & \\
\vdots & \bar{\boldsymbol{x}} & f & \cdots & f & 1 & & & f & \cdots & f & \\
\hline
\vdots & \vdots & & & \vdots & & \ddots & & & & \vdots & \\
\vdots & \vdots & & & \vdots & & & \ddots & & & \vdots & \\
\hline
\vdots & \mathbf{1} & & & & & & & 1 & & & \\
\bar{\mathbf{s}} & \mathbf{2} & f & & & & & & f & 1 & & \\
\vdots & \vdots & \vdots & \ddots & & & \cdots & \cdots & \vdots & \ddots & 1 & \\
\vdots & \bar{\boldsymbol{x}} & f & \cdots & f & & & & f & \cdots & f & 1
\end{array} \qquad (2)
$$

That is, each submatrix $\mathbf{f}_{ij}$ has nonzero elements, indicated by the symbol f, below the main diagonal, and elements equal to one on the diagonal if $i = j$, and zeroes elsewhere. An element of $\mathbf{F}$ contains the probability of reaching state s_i at age x_k, conditional on starting in state s_j at baseline age x_1.

For a regular age grid, we define $\tilde{\mathbf{F}} = n{\cdot}\mathbf{F}$ and corresponding submatrices $\tilde{\mathbf{f}}_{ij}$.[3] The matrix elements of $\tilde{\mathbf{F}}$ contain the *expected length of stay* in a particular state s_i, conditional on initial state s_j. For our purposes, only the entries of columns that refer to the starting age x_1 are relevant. We identify these entries of $\tilde{\mathbf{F}}$ via $\tilde{f}_{s_i s_j,k}$. Note that, while $p_{ij,k}$ refers to the entries of the first subdiagonal of the matrix $\mathbf{p}_{ij}$, $\tilde{f}_{ij,k}$ refers to elements of the first *column* of $\tilde{\mathbf{f}}_{ij}$. Summing the elements of an entire x_1 column of $\tilde{\mathbf{F}}$ yields life expectancies at age x_1, conditional on intitial state s_j, denoted by $e_{x_1|j}$. Summing up the elements of the x_1 column of a submatrix $\tilde{\mathbf{f}}_{ij}$ only yields a conditional component life expectancy, denoted by $e^i_{x_1|j}$. This is the state expectancy for state s_i, conditional on being in state s_j at baseline age. The formulas are

$$
e_{x_1|j} = \sum_{i \in S} e^i_{x_1|j}
$$

$$
e^i_{x_1|j} = \sum_{k \in \Xi} \tilde{f}_{ij,k}
$$

In order to obtain life expectancies that are not conditional on the initial state, a weighted average is calculated,

[3] For a general, unevenly spaced age grid, define $\mathbf{n} = \begin{pmatrix} n_1 & 0 & \cdots & 0 \\ 0 & n_2 & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & n_{\bar{x}} \end{pmatrix}$ and adjust $\mathbf{F}$ according to $\tilde{\mathbf{F}} = (\mathbf{I}_{\bar{s}} \otimes \mathbf{n})\,\mathbf{F}$, where $\otimes$ denotes the Kronecker product. For two matrices A and B, the Kronecker product is

$$
A \otimes B = \begin{pmatrix} a_{11} \cdot B & a_{12} \cdot B & \cdots \\ a_{21} \cdot B & a_{22} \cdot B & \\ \vdots & & \ddots \end{pmatrix}
$$

using the initial state distribution as weights. Denoting the initial proportion of the population in state s_j by g_j, we have for the component (state-specific) life expectancies

$$e^i_{x_1} = \sum_{j \in S} g_j \cdot e^i_{x_1|j}$$

and likewise for the overall life expectancy

$$e_{x_1} = \sum_{i \in S} e^i_{x_1} = \sum_{i \in S} \sum_{j \in S} g_j \cdot e^i_{x_1|j} = \sum_{j \in S} g_j \cdot \sum_{i \in S} e^i_{x_1|j} = \sum_{j \in S} g_j \cdot e_{x_1|j}$$

The above calculations assume end-of-period transitions and hence overestimate state expectancies since death normally occurs around mid-period, on average. The assumption of *mid-period transitions* (i.e. transitions at ages $x_1 + \frac{1}{2}n$, $x_2 + \frac{1}{2}n$, $x_3 + \frac{1}{2}n$, etc.) can be implemented for an evenly spaced age grid by redefining $\tilde{\mathbf{F}} = n\left(\mathbf{F} - \frac{1}{2}\mathbf{I}_{\bar{s}\bar{x}}\right)$ with subsequent calculations unaltered.

It is worth restating the difference between $\mathbf{F}$ and $\tilde{\mathbf{F}}$. While$\mathbf{F}$ contains probabilities of reaching a state, $\tilde{\mathbf{F}}$ contains the expected lengths of stay in a state, which can be seen as *expected time rewards.* When turning to the rewards framework in the next section, $\mathbf{F}$ is used instead of $\tilde{\mathbf{F}}$. Time rewards enter the calculations via separate, newly defined matrices.

1.1.3 Markov Chain with Rewards and Its Application to Transition Timing

Rewards calculations are based on rewards matrices $\mathbf{R}_m$, which associate transitions with payoffs for state s_m. Any transition can be modelled to yield payoffs towards any state or for any number of states. For example, transitioning from s_j to s_i can result in a reward of $\frac{n}{2}$ to both $e^j_{x_1}$ and $e^i_{x_1}$ (via entries in $\mathbf{R}_j$ and $\mathbf{R}_i$), which is an obvious way of modelling the assumption of mid-interval transitions.

The block structure of $\mathbf{R}_m$ is identical to the one of $\mathbf{P}$:

$$\mathbf{R}_m = \begin{pmatrix} \mathbf{r}_{m,11} & \cdots & \mathbf{r}_{m,1\bar{s}} & \mathbf{0} \\ \vdots & \ddots & \vdots & \vdots \\ \mathbf{r}_{m,\bar{s}1} & \cdots & \mathbf{r}_{m,\bar{s}\bar{s}} & \mathbf{0} \\ \overrightarrow{\mathbf{r}}_{m,d1} & \cdots & \overrightarrow{r}_{m,d\bar{s}} & 0 \end{pmatrix}$$

with

$$\mathbf{r}_{m,ij} = \begin{pmatrix} 0 & 0 & \cdots & 0 & 0 \\ r_{m,ij,2} & 0 & \cdots & 0 & 0 \\ 0 & r_{m,ij,3} & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & r_{m,ij,\bar{x}} & 0 \end{pmatrix}$$

For example, the scalar $r_{m,ij,k}$ quantifies the reward towards state s_m for the transition from state s_j to s_i when turning x_k. $\overrightarrow{\mathbf{r}}_{m,dj}$ are row vectors and reward transitions into death:

$$\overrightarrow{\mathbf{r}}_{m,dj} = \begin{pmatrix} r_{m,dj,2} & \cdots & r_{m,dj,\bar{x}} & r_{m,dj,\bar{x}+1} \end{pmatrix}$$

Note that $r_{m,dj,\bar{x}+1}$ can be a positive magnitude, even though death is certain for the age $\bar{x}$ to age $\bar{x}+1$ transition. The last column of $\mathbf{R}_m$ consists entirely of zeros, reflecting the fact that being dead is never rewarded. It is important to keep in mind that all rewards-related matrices, vectors and scalars quantify rewards towards a specific state, and hence receive all an additional state index m.

A convenient matrix formula for calculating lifetime rewards towards state s_m is

$$\mathbf{e}^{m}_{x_1|\cdot} = \left(\mathbf{I}_{\bar{s}} \otimes \mathbf{e}_{1,\bar{x}}\right) \mathbf{F}'\mathbf{Z} \left(\mathbf{1}_{1,\bar{s}\bar{x}+1} \cdot \left(\mathbf{P} \circ \mathbf{R}_m\right)\right)' \tag{3}$$

where $\mathbf{e}_{1,\bar{x}}$ is a $1 \times \bar{x}$ rowvector with one as the first element and zeroes elsewhere, and $\mathbf{Z}$ is defined as $[\mathbf{I}_{\bar{s}} \;\; \mathbf{0}_{\bar{s}\times 1}]$, and $\mathbf{1}_{1,\bar{s}+1}$ is an $\bar{s}+1$ row vector of ones.[4] The formula links $\mathbf{F}$, which contains probabilities of reaching a particular state, with $(\mathbf{1}_{1,\bar{s}+1} \cdot (\mathbf{P} \circ \mathbf{R}_m))'$, which associates that state with rewards in terms of possible transitions out of that state. The other matrices involved just select or drop elements. It is instructive to have a closer look at the different terms that make up the formula. $(\mathbf{1}_{1,\bar{s}+1} \cdot (\mathbf{P} \circ \mathbf{R}_m))'$ does an elementwise multiplication of $\mathbf{P}$ and $\mathbf{R}_m$, sums each column over the rows, and transposes the result. Let $b_{m,ij,k} = p_{ij,k} \cdot r_{m,ij,k}$ denote the scalars of the elementwise multiplications. $b_{m,ij,k}$ is the expected reward for state m of transition ij at age x_k. $(\mathbf{1}_{1,\bar{s}+1} \cdot (\mathbf{P} \circ \mathbf{R}_m))'$ is then equal to

$$\sum_{i \in \tilde{S}} \left(\begin{array}{c} b_{m,i1,2} \\ \vdots \\ b_{m,i1,\bar{x}} \\ r_{m,d1,\bar{x}+1} \\ \hline \vdots \\ \vdots \\ \hline b_{m,i\bar{s},2} \\ \vdots \\ b_{m,i\bar{s},\bar{x}} \\ r_{m,d\bar{s},\bar{x}+1} \\ \hline 0 \end{array} \right)$$

Each element of the above vector contains the expected reward[5] of *being* in a particular state at a particular age, where this magnitude is calculated as the sum of the expected rewards for all possible transitions out of the state. Premultiplication by $\mathbf{Z}$ just cuts off the last element, which is zero. To understand the premultiplication by $\mathbf{F}'$, consider the product implied by the first column of $\mathbf{F}$ only, which is

$$\begin{aligned} & f_{11,1} \cdot \sum_{i \in \tilde{S}} b_{m,i1,2} + f_{11,2} \cdot \sum_{i \in \tilde{S}} b_{m,i1,3} + \cdots + f_{11,\bar{x}} \cdot r_{m,d1,\bar{x}+1} \\ + & f_{21,1} \cdot \sum_{i \in \tilde{S}} b_{m,i2,2} + f_{21,2} \cdot \sum_{i \in \tilde{S}} b_{m,i2,3} + \cdots + f_{21,\bar{x}} \cdot r_{m,d2,\bar{x}+1} \\ + & \ldots \\ + & f_{\bar{s}1,1} \cdot \sum_{i \in \tilde{S}} b_{m,i\bar{s},2} + f_{\bar{s}1,2} \cdot \sum_{i \in \tilde{S}} b_{m,i\bar{s},3} + \cdots + f_{\bar{s}1,\bar{x}} \cdot r_{m,d\bar{s},\bar{x}+1} \end{aligned}$$

The first column of $\mathbf{F}$ refers to those being in state s_1 at baseline age x_1. Consequently, the above expression

[4] The formula uses the Hadamard product, denoted by $\circ$, which indicates elementwise multiplication: For two matrices A and B, $[A \circ B]_{ij} = a_{ij} \cdot b_{ij}$.

[5] The last element of each state block s_j lacks a $p_{\cdot\cdot,\cdot}$ term since it is equal to one.

only contains elements $f_{\cdot 1,\cdot}$. Looking at the first product term, $f_{11,1}$ denotes the probability of reaching state s_1 at age x_1 when starting out in s_1 at age x_1 (which equals one, as a quick check of equation (2) confirms). This probability is multiplied by the expected reward of being in that state, calculated as the sum of the expected rewards of all out-transitions, weighted by the transition probabilities. This calculation is repeated for all elements $f_{i1,k}$, summing over all ages x_k (within each line) and summing over all possible transients states s_i (across lines). In general, the probabilities of reaching future states are multiplied by the expected rewards for these states, and the sum of these numbers over all possible states that a subject can go through yield $e^m_{x_1|1}$, i.e. the overall reward for state s_m, conditional on initial state s_1.

Lastly, the premultiplication of $\mathbf{F}'$ by $(\mathbf{I}_{\bar{s}} \otimes \mathbf{e}_{1,\bar{x}})$ selects those columns of $\mathbf{F}$ that are relevant, namely the age x_1 columns for all possible initial states. The overall result is

$$\mathbf{e}^m_{x_1|\cdot} = \begin{pmatrix} e^m_{x_1|1} \\ \vdots \\ e^m_{x_1|\bar{s}} \end{pmatrix}$$

The calculation of the unconditional magnitudes $e^m_{x_1}$ and e_{x_1} is then as laid out in the previous section, using initial state proportions for obtaining weighted averages.

It is worth reiterating that the above formula for calculating rewards contains $\mathbf{F}$, not $\tilde{\mathbf{F}}$. Any time rewards that account for age intervals different from one and for irregularly spaced age intervals are modelled via the $\mathbf{R}_m$ matrices.

1.2 Equivalence to Standard Life Table Calculations

A notational note must precede the following elaborations. There is a slight difference between notation used in this paper and conventional demographic notation. The former is geared towards transitions at points in time, whereas the latter refers to intervals. For example, the probability of dying within the age interval x_k to x_{k+1} is ${}_nq_{x_k} = 1 - {}_np_{x_k}$ in demographic notation, but $p_{ds_i,x_{k+1}}$ in the notation used so far in this paper. The important point to note is that the age index differs by one. Since we now focus on life table formulas and check the two methods for numerical equivalence, it will be useful to adhere to demographic notation in this section.

Assume that there is only one transient state (alive) and one absorbing state (death). Then the rewards calculations coincide with standard life table calculations if survival probabilities and the values for Chiang's a (the average time lived within the age interval by those who die in the age interval) are known. Suppose we had survival probabilities and values for Chiang's a, denoted by the conventional demographic symbols ${}_np_x$ and ${}_na_x$, respectively. We allow for irregularly spaced age intervals, denoted n_k, but drop the k-index when n itself is in the left index of a variable. The formula for the life expectancy then is

$$\sum_{k=1}^{\bar{x}} \left(\prod_{h=1}^{k-1} {}_np_{x_h} \right) \left({}_np_{x_k} \cdot n_k + (1 - {}_np_{x_k}) {}_na_{x_k} \right) \tag{4}$$

For each age x_k, we calculate the survival probability up to that age (the product term) and multiply it by the sum of the probability of surviving the age interval times the length of the age interval and the expected life time contribution of those who die in the age interval.[6]

We now show that corresponding rewards calculations in (3) are identical. The relevant matrices are:

[6]It is understood implicitly that the product term equals one for the first age group.

$$\mathbf{P} = \begin{pmatrix} 0 & 0 & \cdots & 0 & 0 \\ {}_np_{x_1} & 0 & \cdots & 0 & 0 \\ \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & \cdots & {}_np_{\bar{x}-1} & 0 & 0 \\ 1 -_n p_{x_1} & \cdots & 1 -_n p_{\bar{x}-1} & 1 & 1 \end{pmatrix}$$

The last two columns say that the probability of dying when being in the highest age interval is one, as is the probability of remaining dead once the dead state has been entered. **U** is the submatrix that results when dropping the last column and the last row of **P**, and

$$\mathbf{R} = \begin{pmatrix} 0 & 0 & \cdots & 0 & 0 \\ n_1 & 0 & \cdots & 0 & 0 \\ \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & \cdots & n_{\bar{x}-1} & 0 & 0 \\ {}_na_{x_1} & \cdots & {}_na_{\bar{x}-1} & {}_na_{\bar{x}} & 0 \end{pmatrix}$$

From the argument leading up to equation (1), it is easily seen that, for the single-state case, **F** is

$$\mathbf{F} = \begin{pmatrix} 1 & 0 & 0 & \cdots \\ {}_np_{x_1} & 1 & 0 & \cdots \\ {}_np_{x_1} \cdot {}_np_{x_2} & {}_np_{x_2} & 1 & \cdots \\ {}_np_{x_1} \cdot {}_np_{x_2} \cdot {}_np_{x_3} & {}_np_{x_2} \cdot {}_np_{x_3} & {}_np_{x_3} & \cdots \\ \vdots & \vdots & \vdots & \cdots \\ \prod_{i=1}^{\bar{x}-1} {}_np_{x_i} & \cdots & \cdots & \cdots \end{pmatrix}$$

If we focus on life expecancy at the first age, only the first column is relevant. Its entries corresponds to the product term in equation (4). This column is isolated by the selection term on the left of equation (3). The second parenthesis in equation (4) corresponds to the elements of the second parenthesis in equation (3). Finally, The matrix product of $(\mathbf{I}_{\bar{s}} \otimes \mathbf{e}_{1,\bar{x}})\,\mathbf{F}'$ and $\mathbf{Z}\,(\mathbf{1}_{1,\bar{s}\bar{x}+1} \cdot (\mathbf{P} \circ \mathbf{R}_m))'$ does the appropriate sum of products, as in equation (4).

As an illustration, the HMD values for the demographic 5-year age spacing for Spain and the year 2000 are 0.996, 0.999, 0.999, ..., 0.032, 0.000 for ${}_np_{x_k}$ and 0.14, 1.62, 2.36, ..., 1.42, 1.33 for ${}_na_{x_k}$. Using these values, each of the equations (3) and (4) yield $e_0 = 79.42321813$, which rounds to $e_0 = 79.42$, the value cited in the HMD.

1.3 Multistate Application: Retirement Life Expectancies

This section adapts the general expressions of the previous section for the retirement application, whose basic parameters are described in the table below:

symbol	meaning	value
$\bar{x}$	number of age classes	61
x_1	minimum (baseline) age in model	50
Ξ	set of age classes	$\{50, \ldots, 110\}$
n	length of age interval	1
$\bar{s}$	number of *transient* states	3
S	set of transient states	$\{w, u, r\}$
$\tilde{S}$	set of all states: $\tilde{S} = S \cup \{d\}$	$\{w, u, r, d\}$

In particular, the model incorporates three transient states: working, unemployed, and retired, denoted as w, u, r, respectively.[7] The transition matrix is

$$\mathbf{P} = \begin{pmatrix} \mathbf{p}_{ww} & \mathbf{p}_{wu} & \mathbf{p}_{wr} & \mathbf{0} \\ \mathbf{p}_{uw} & \mathbf{p}_{uu} & \mathbf{p}_{ur} & \mathbf{0} \\ \mathbf{p}_{rw} & \mathbf{p}_{ru} & \mathbf{p}_{rr} & \mathbf{0} \\ \overrightarrow{\mathbf{p}}_{dw} & \overrightarrow{p}_{du} & \overrightarrow{p}_{dr} & 1 \end{pmatrix}$$

and has the following, detailed structure:

$\mathbf{s_i} \setminus \mathbf{s_j}$		$\cdots$	$\cdots$	**w**	$\cdots$	$\cdots$	$\cdots$	**u**	$\cdots$	$\cdots$	$\cdots$	**r**	$\cdots$	**d**
	$\mathbf{x_k}$	**50**	$\cdots$	**109**	**110**	**50**	$\cdots$	**109**	**110**	**50**	$\cdots$	**109**	**110**	**.**
$\vdots$	**50**													
w	**51**	p				p				p				
$\vdots$	$\vdots$		$\ddots$				$\ddots$				$\ddots$			
$\vdots$	**110**			p				p				p		
$\vdots$	**50**													
u	**51**	p				p				p				
$\vdots$	$\vdots$		$\ddots$				$\ddots$				$\ddots$			
$\vdots$	**110**			p				p				p		
$\vdots$	**50**													
r	**51**	p				p				p				
$\vdots$	$\vdots$		$\ddots$				$\ddots$				$\ddots$			
$\vdots$	**110**			p				p				p		
d	.	p	p	p	1	p	p	p	1	p	p	p	1	1

The calculation of the fundamental matrix $\mathbf{F}$ is based on the submatrix $\mathbf{U}$ of $\mathbf{P}$:

$$\mathbf{U} = \begin{pmatrix} \mathbf{p}_{ww} & \mathbf{p}_{wu} & \mathbf{p}_{wr} \\ \mathbf{p}_{uw} & \mathbf{p}_{uu} & \mathbf{p}_{ur} \\ \mathbf{p}_{rw} & \mathbf{p}_{ru} & \mathbf{p}_{rr} \end{pmatrix}$$

The detailed structure of $\mathbf{F}$ is

$$\mathbf{F} = (\mathbf{I}_{\mathbf{61 \cdot 3}} - \mathbf{U})^{-1} = \begin{pmatrix} \mathbf{f}_{ww} & \mathbf{f}_{wu} & \mathbf{f}_{wr} \\ \mathbf{f}_{uw} & \mathbf{f}_{uu} & \mathbf{f}_{ur} \\ \mathbf{f}_{rw} & \mathbf{f}_{ru} & \mathbf{f}_{rr} \end{pmatrix}$$

[7] The symbol r now has two different meanings: It may refer to a particular state (being retired), or to rewards magnitudes. What is meant will be clear from the context, however.

$$
= \begin{array}{cc|cccc|cccc|cccc}
\mathbf{s_i} \setminus \mathbf{s_j} & & \cdots & \cdots & \mathbf{w} & \cdots & \cdots & \cdots & \mathbf{u} & \cdots & \cdots & \cdots & \mathbf{r} & \cdots \\
 & \mathbf{x_k} & \mathbf{50} & \cdots & \mathbf{109} & \mathbf{110} & \mathbf{50} & \cdots & \mathbf{109} & \mathbf{110} & \mathbf{50} & \cdots & \mathbf{109} & \mathbf{110} \\
\hline
\vdots & \mathbf{50} & 1 & & & & & & & & & & & \\
\mathbf{w} & \mathbf{51} & f & 1 & & & f & & & & f & & & \\
\vdots & \vdots & \vdots & \ddots & 1 & & \vdots & \ddots & & & \vdots & \ddots & & \\
\vdots & \mathbf{110} & f & \cdots & f & 1 & f & \cdots & f & & f & \cdots & f & \\
\hline
\vdots & \mathbf{50} & & & & & 1 & & & & & & & \\
\mathbf{u} & \mathbf{51} & f & & & & f & 1 & & & f & & & \\
\vdots & \vdots & \vdots & \ddots & & & \vdots & \ddots & 1 & & \vdots & \ddots & & \\
\vdots & \mathbf{110} & f & \cdots & f & & f & \cdots & f & 1 & f & \cdots & f & \\
\hline
\vdots & \mathbf{50} & & & & & & & & & 1 & & & \\
\mathbf{r} & \mathbf{51} & f & & & & f & & & & f & 1 & & \\
\vdots & \vdots & \vdots & \ddots & & & \vdots & \ddots & & & \vdots & \ddots & 1 & \\
\vdots & \mathbf{110} & f & \cdots & f & & f & \cdots & f & & f & \cdots & f & 1
\end{array}
$$

Denoting by $f_{ij,k}$ the probability of reaching state s_i at age x_k conditional on starting in state s_j at age 50,[8] the formulas from the section 1.1.2. for life expectancy numbers can be directly applied.

Turning to rewards calculations, the matrix that specifies rewards towards state s_m, $m \in S$, becomes

$$
\mathbf{R}_m = \begin{pmatrix}
\mathbf{r}_{m,ww} & \mathbf{r}_{m,wu} & \mathbf{r}_{m,wr} & \mathbf{0} \\
\mathbf{r}_{m,uw} & \mathbf{r}_{m,uu} & \mathbf{r}_{m,ur} & \mathbf{0} \\
\mathbf{r}_{m,rw} & \mathbf{r}_{m,ru} & \mathbf{r}_{m,rr} & \mathbf{0} \\
\overrightarrow{\mathbf{r}}_{m,dw} & \overrightarrow{r}_{m,du} & \overrightarrow{r}_{m,dr} & 0
\end{pmatrix}
$$

For different timing assumptions, we define different rewards matrices. In the matrices below, non-specified elements are zero. For the end-of-period replication, define

$$
\mathbf{R}_w = \mathbf{R}_u = \mathbf{R}_r = \begin{array}{cccc|cccc|cccc|c}
1 & & & & 1 & & & & 1 & & & & \\
 & \ddots & & & & \ddots & & & & \ddots & & & \\
 & & 1 & & & & 1 & & & & 1 & & \\
\hline
1 & & & & 1 & & & & 1 & & & & \\
 & \ddots & & & & \ddots & & & & \ddots & & & \\
 & & 1 & & & & 1 & & & & 1 & & \\
\hline
1 & & & & 1 & & & & 1 & & & & \\
 & \ddots & & & & \ddots & & & & \ddots & & & \\
 & & 1 & & & & 1 & & & & 1 & & \\
\hline
1 & \cdots & 1 & 1 & 1 & \cdots & 1 & 1 & 1 & \cdots & 1 & 1 & 0
\end{array}
$$

For the mid-period specification, define

[8] Since we have a regular age grid with all age intervals of length one, this is equal to the expected length of stay.

$$
\mathbf{R}_w = \begin{array}{cccc|cccc|cccc|c}
1 & & & & 0.5 & & & & 0.5 & & & & \\
 & \ddots & & & & \ddots & & & & \ddots & & & \\
 & & 1 & & & & 0.5 & & & & 0.5 & & \\
\hline
0.5 & & & & 0 & & & & 0 & & & & \\
 & \ddots & & & & \ddots & & & & \ddots & & & \\
 & & 0.5 & & & & 0 & & & & 0 & & \\
\hline
0.5 & & & & 0 & & & & 0 & & & & \\
 & \ddots & & & & \ddots & & & & \ddots & & & \\
 & & 0.5 & & & & 0 & & & & 0 & & \\
\hline
0.5 & \cdots & 0.5 & 0.5 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0
\end{array}
$$

$$
\mathbf{R}_u = \begin{array}{cccc|cccc|cccc|c}
0 & & & & 0.5 & & & & 0 & & & & \\
 & \ddots & & & & \ddots & & & & \ddots & & & \\
 & & 0 & & & & 0.5 & & & & 0 & & \\
\hline
0.5 & & & & 1 & & & & 0.5 & & & & \\
 & \ddots & & & & \ddots & & & & \ddots & & & \\
 & & 0.5 & & & & 1 & & & & 0.5 & & \\
\hline
0 & & & & 0.5 & & & & 0 & & & & \\
 & \ddots & & & & \ddots & & & & \ddots & & & \\
 & & 0 & & & & 0.5 & & & & 0 & & \\
\hline
0 & \cdots & 0 & 0 & 0.5 & \cdots & 0.5 & 0.5 & 0 & \cdots & 0 & 0 & 0
\end{array}
$$

$$
\mathbf{R}_r = \begin{array}{cccc|cccc|cccc|c}
0 & & & & 0 & & & & 0 & & & & \\
 & \ddots & & & & \ddots & & & & \ddots & & & \\
 & & 0 & & & & 0 & & & & 0 & & \\
\hline
0 & & & & 0 & & & & 0 & & & & \\
 & \ddots & & & & \ddots & & & & \ddots & & & \\
 & & 0 & & & & 0 & & & & 0 & & \\
\hline
0.5 & & & & 0.5 & & & & 1 & & & & \\
 & \ddots & & & & \ddots & & & & \ddots & & & \\
 & & 0.5 & & & & 0.5 & & & & 1 & & \\
\hline
0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0.5 & \cdots & 0.5 & 0.5 & 0
\end{array}
$$

If, in addition to the mid-period transitions, we use exact retirement rewards, the matrices become

$$
\mathbf{R}_w = \begin{array}{cccc|cccc|cccc|c}
1 & & & & 0.5 & & & & 0.5 & & & & \\
& \ddots & & & & \ddots & & & & \ddots & & & \\
& & 1 & & & & 0.5 & & & & 0.5 & & \\ \hline
0.5 & & & & 0 & & & & 0 & & & & \\
& \ddots & & & & \ddots & & & & \ddots & & & \\
& & 0.5 & & & & 0 & & & & 0 & & \\ \hline
\nu_2 & & & & 0 & & & & 0 & & & & \\
& \ddots & & & & \ddots & & & & \ddots & & & \\
& & \nu_{\bar{x}} & & & & 0 & & & & 0 & & \\ \hline
0.5 & \cdots & 0.5 & 0.5 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0
\end{array}
$$

$$
\mathbf{R}_u = \begin{array}{cccc|cccc|cccc|c}
0 & & & & 0.5 & & & & 0 & & & & \\
& \ddots & & & & \ddots & & & & \ddots & & & \\
& & 0 & & & & 0.5 & & & & 0 & & \\ \hline
0.5 & & & & 1 & & & & 0.5 & & & & \\
& \ddots & & & & \ddots & & & & \ddots & & & \\
& & 0.5 & & & & 1 & & & & 0.5 & & \\ \hline
0 & & & & \nu_2 & & & & 0 & & & & \\
& \ddots & & & & \ddots & & & & \ddots & & & \\
& & 0 & & & & \nu_{\bar{x}} & & & & 0 & & \\ \hline
0 & \cdots & 0 & 0 & 0.5 & \cdots & 0.5 & 0.5 & 0 & \cdots & 0 & 0 & 0
\end{array}
$$

$$
\mathbf{R}_r = \begin{array}{cccc|cccc|cccc|c}
0 & & & & 0 & & & & 0 & & & & \\
& \ddots & & & & \ddots & & & & \ddots & & & \\
& & 0 & & & & 0 & & & & 0 & & \\ \hline
0 & & & & 0 & & & & 0 & & & & \\
& \ddots & & & & \ddots & & & & \ddots & & & \\
& & 0 & & & & 0 & & & & 0 & & \\ \hline
1-\nu_2 & & & & 1-\nu_2 & & & & 1 & & & & \\
& \ddots & & & & \ddots & & & & \ddots & & & \\
& & 1-\nu_{\bar{x}} & & & & 1-\nu_{\bar{x}} & & & & 1 & & \\ \hline
0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0.5 & \cdots & 0.5 & 0.5 & 0
\end{array}
$$

where ν_k denotes the fraction of the retirement age-year spent working when moving into retirement at age x_k.

For any of the above scenarios, life expectancies can be calculated according to the formula

$$\mathbf{e}_{50|\cdot}^{m} = \begin{pmatrix} e_{50|w}^{m} \\ e_{50|u}^{m} \\ e_{50|r}^{m} \end{pmatrix} = \left(\mathbf{I}_3 \otimes \mathbf{e}_{1,61}\right) \mathbf{F}' \mathbf{Z} \left(\mathbf{1}_{1,4} \cdot \left(\mathbf{P} \circ \mathbf{R}_m\right)\right)'$$

followed by the usual weighting by initial proportions in order to obtain unconditional state and overall expectancies.

References

[Caswell and van Daalen, 2019] Caswell, H. and van Daalen, S. F. (2019). Healthy longevity (and related measures) from multistate incidence-based health models.