Measuring working-life expectancies: multistate vector regression approach vs. prevalence-based life table method

Markku Nurminen
Department of Public Health, University of Helsinki, Finland
MarkStat Consultancy, Helsinki; Finland

Abstract

BACKGROUND
Demographic aging is ensued by many societal challenges. Extending working life has been proposed as a key measure to adjust to the involved socio-economic problems. Still measurement of the length of working careers is not yet a standard practice.

OBJECTIVE
This article re-examines traditionally used and modern statistical methods for the analysis of life tables with a view to measure accurately the future employment time.

DATA AND METHODS
This paper discusses the advantages and limitations of the two alternative methods (Sullivan, 1971, and Davis et al., 2001), and compares the performance of their working-life expectancy estimates in a population-based study that were used to analyze aggregate sequential labor force survey data from Finland in 2000-2010.
RESULTS

We provide cogent methodological arguments substantiated by empirical findings to evince the better usefulness of the preferred multistate vector regression approach over the commonly used actuarial current life table technique.

CONCLUSIONS

The multistate modeling and estimation methodology presented enables a superior statistical analysis of stochastic processes in working life and deals with a real problem encountered foremost in developed countries with aging populations.

1. Introduction

Population aging confronts all developed countries and a variety of strategies have been undertaken to address it. In Finland (Nurminen et al., 2005), like in many other countries (Dhillon and Ladusingh, 2013), extending the time spent in employment over the life course has been put forward as a key measure to adjust to the increase of the longevity of the population. However, measurement of the length of working careers is not easy and not yet a standard practice. Thus it is important to apply accurate measures to quantify the duration of future employment time. This paper first recalls the definitions of alternative expectancy measures and re-examines the practices of statistical methods employed for measuring working life expectancy. Secondly it compares the estimates of working-life expectancies calculated using two competing approaches, the regular period life table technique (Sullivan 1971) and the multistate vector regression approach (Davis et al. 2001). We then explain the reasons for discrepancies between the results, derived empirically from an application to a recent population-based study from Finland (Nurminen 2012).
Finally, although many of the issues raised here are specific to Finland, nonetheless, the questions concerning the development of time spent in employment and its measurement are especially relevant for several developed countries.

2. Alternative measures of working-life expectancy

This section considers the attributes of working-life expectancy in comparison with some other approaches of measuring the length of working life as well as recalls the definitions and historical trends of the alternative measures.

The estimation of the duration of future employment time is not a simple matter. In reviewing the alternative employment activity measures, Hytti (2009) discussed the relative advantages and limitations of the retirement exit age (i.e. the average age of withdrawal from the labor market) versus active-life expectancy. Hytti pointed out that labor market exit age acts rapidly and to the correct direction of the changes in the transitions to retirement. However, the exit age measure does this ignoring the cumulative experience up to the present time. By comparison, the expectancy was said to react slowly to the changes in the participation of labor market and in the usage of pension scheme. But the expectancy measure — which can be regarded as a far-sighted feature — is also influenced by the behavior of the studied population in the preceding years. Another advantage is that expectancy shows whether or not the development tends toward the targets set in the official employment and pension policies.

For an individual at a particular age, working-life expectancy is the expected number of working years remaining in one's life (e.g. Nurminen 2008). While this is a hypothetical construct that cannot be directly measured, it is an intuitive and broadly accessible concept. As such it can provide the means for summarizing and
comparing the labor market status of surveyed populations as well as for monitoring the time trends in employment statistics.

This definition is interpreted statistically as the expected (average) value of the distribution of the length of working-life in the population, which is consistently estimated by the sample mean. It is the future duration or occupation time in the employed state conditional on an individual's initial age. Note that 'future' time is ingrained in the concept of expectation. However, in the context of a birth cohort's follow-up it refers to the aging of its members within the range of a calendar year period that was used as the population-time data base for the estimation of the multistate regression model. Therefore, the working-life expectancy should not be interpreted as a predicted value. Forecasting beyond the observed data base is treacherous, e.g., because the premises of the underlying regression model may change.

Summary measures for the working years can provide useful indicators for evaluating labor force potential plus for evaluating the need for policy adjustments. Although the calculation of the working-life expectancy tables is complicated, their wide use demonstrates that they nevertheless are relevant and comprehensible.

Approaches to measure the length of working life vary according to what object exactly is being targeted and how it is formally estimated. Because of the abundant early retirement in the past, discussion on extending working lives has focused largely on pension systems (Kannisto 2012). Increasing life expectancy also suggests that a natural way to extend working lives is to push the final exit from working life further in the life course. However, over the past several years there has been a growing understanding in Finland and in other parts of Europe that it is useful to see the question of extending working lives in a context of the entire life span.
(Hytti and Nio 2004, Vogler-Ludwig 2009, Hytti and Valaste 2009). The duration of working life is influenced not only by retirement but also by other age-related activities over the course of life, such as participation to education and child care, spells of unemployment or long term sickness absence.

Measures of the participation to working life over the life span may be founded on the experiences of actual cohorts. The time cohort members spent in working life may be measured retrospectively, taking persons alive at the end of the working-age years and calculating the time they have participated to working life over the course of their life (e.g. Lehto 2012). Alternatively, an actual cohort may be longitudinally followed prospectively from birth up to over the working-age years taking into account the mortality of the cohort. Both methods estimate the length of working life of a cohort as it has realized in past circumstances over the history of the cohort. They do not, however, provide us with information on how the length of working life is evolving in time and under prevailing circumstances.

Expectancies aim to capture the present state of a population. They estimate the average time spent alive in a defined state (alive, employed, not retired, etc.) computed from age-specific probabilities or risks over a period of time. Partial life expectancies up to a given age were first used in health metrics to calculate the 'healthy life expectancy' of the total life expectancy as a response to the intriguing qualitative question of whether the increase in life expectancy adds to 'healthy' or 'un-healthy' years. Or, to put the question quantitatively: What share of the increase in life expectancy is taken in a state of disability? For a demographic-epidemiologic application, see Davis et al. (2002b).
3. Development of model-based expectancies for working-life tables

To place the recently developed regression modeling methodologies using working-life tables in the general context of demographic life tables, we recall the distinctions between of the different approaches to the analysis of labor force participation.

Applications of working population health indicators (Nurminen 2004) such as active life expectancy have been numerous (e.g. Lièvre et al. 2003, Strulik and Werner 2012). They have been applied also in Finland for employment patterns by the Social Insurance Institution (Hytti and Nio 2004) and for retirement by the Finnish Centre for Pensions (Kannisto 2006). Working-life applications of partial life expectancy have gained growing interest over the past few decades as an analogous issue of how participation to work life develops in relation to total life expectancy (Nurminen 2008, Hytti 2009). As a summary statistic of age-specific participation to working life, working-life expectancy measures the average length of working life remaining for an individual at a given point in time. It is usually gauged in units of expected years remaining in working-life that are meaningful to ordinary laypersons. The term 'partial life expectancy' refers to the average number of years remaining between exact ages $x$ and $z$ (where $15 \leq x < z < 65$, the limiting age) by persons alive at an exact age $x$. Summary measures of the working-life years can provide useful indicators for evaluating labor force potential.

A study for the EU Commission sought to investigate the working-life indicator which should complement the monitoring instruments of the European Strategy by focusing on the entire life cycle of active persons and persons in employment (Vogler-Ludwig 2009). The study recommended using the working-life expectancy as one of the core labor market indicators at European and national level. Recently the Employment Committee decided that the working-life expectancy will replace the
average age of withdrawal from labor force (or exit age) indicator. The expectancy indicator will be utilized for monitoring the European employment guidelines. The measure has already been included in the Joint Assessment Framework indicator package by experts who peruse the targets and trends in employment set for the EU's growth strategy.

Working-life expectancies have been applied using various definitions of working-life participation. In one of the early studies in the field, Hytti (1998) measured active life time as opposed to the time spent in retirement. Several studies have defined participation to working life as equal to labor force participation (e.g. Hytti and Nio 2004, Hytti and Valaste 2009 [labor market expectancy], Vogler-Ludwig 2009 [duration of active working life]). The problem with this definition is that time spent gainfully employed and time spent unemployed are merged. By pooling the employed and unemployed individuals these expectancies obscure the age-pattern of working-life participation because employment and unemployment are highly age-specific phenomena but in a mutually reversed manner. Some studies have, thus, defined the expectancy of labor market time as the expectancy of time spent employed (Hytti and Valaste 2009, Vogler-Ludwig 2009). Yet another way to measure the participation to working life is in terms of working time. This has been carried out by calculating the expected lifetime duration of working time in hours (Vogler-Ludwig 2009) or by dividing employment time to full-time and part-time employment (Hytti and Valaste 2009), which would also include those working only part of the year due to the seasonal nature of their job. Thus, depending on person's education, level of skill, and type of job, a non-trivial fraction of one's working life might be spent not working.
We prefer to define the working-life expectancy exclusively in terms of participation to gainful employment. However, to get an even clearer picture, we complement working-life expectancy with expectancies estimated for time spent in unemployment and for inactive time. By explicitly distinguishing the multiple relevant labor market states and estimating their expectancies we aim to gain a better understanding of the interplay between the working and non-working life over the entire life span. While we recognize the significance of working time and, in particular, the role of part-time work in labor market policies aimed at lengthening life-time employment and adding flexibility to different phases of life, measuring working-time expectancy is beyond the scope of present paper. In theory, the multi-state vector regression approach could be applied to partially inform on this question. This could be achieved by splitting the 'employed' state into several states based on an individual's usual number of weekly working hours during the observation period. For example, we could for a model with state space: employed for at least 35 hours a week, employed for between 15 and 34 hours a week, employed for between one and 14 hours a week, unemployed, inactive.

Working-life expectancy may be estimated using either marginal probabilities (prevalence rates) or transition probabilities (incidence rates). The actuarial prevalence-based working-life expectancy describes the prevailing work life-participation rate in a cross-sectional sample of a population (Sullivan 1971). This rate incorporates workers' past labor market experiences as these manifest in the age-specific frequencies of being occupied in a labor market state during a given short time period. Prevalence-based life tables are based on external information on the proportion of given health dimension on the other hand, for instance disability, in each state. The Sullivan method relies on widely-available data (period life table and
age-specific cross-sectional prevalence of disability). But it makes the 'stationarity' assumption, i.e. it assumes that observed cross-sectional prevalence of disability/mortality is equal to that of the hypothetical ('synthetic') cohort comprised of data from actual cohorts that are present at different ages, in a specified year. Contemporary work treating marginal probabilities by use of the multistate vector regression modeling approach for estimating working-life expectancies was developed in Davis et al. (2001). Incidence-based expectancy is calculated from longitudinal cohort data that are required to estimate transitions between various states (Davis et al. 2002a, Davis et al. 2007).

While ordinary current life table technique is set up on a calculation of employment probabilities based on a period life table (Sullivan 1971), we prefer to base the analysis on a cohort life table modeling approach (Davis et al. 2001). A limitation of the former type indicators appears to be that they are descriptions of the whole life cycle rather than specific phases of working life (like combined periods of unemployment). Moreover, they describe the current state of working life participation over all ages, rather than providing estimates for the years spanning the entire working life. Yet the limitations pertain only to the period life table technique, not to the multistate cohort life table approach.

Considering the advantages and limitations of the two comparative methods, our stand is that, while the period life table expectancy is the most common measure in a readily useable form, it is an estimate pertaining to a particular point in time. In contrast, regression-based cohort life table expectancy, which models and projects future labor force participation, is a more appropriate statistic for planning and policy development objectives. Cohort life table expectancy can be theoretically founded on large-sample, weighted least squares theory, and therefore allows stochastic data
analysis and inference (inter alia, with respect to significance tests, interval estimates, interaction effects, time trends, and projections).

The Markov property is not required for the estimation of the working-life expectancies using marginal probabilities (Davis et al. 2001), whereas the estimation that uses transition probabilities was done under the Markov assumption (Davis et al. 2002a). The failure of the Markov condition would mean that the estimates are not statistically efficient but the method is still useable. When the Markov property does not hold standard errors can be obtained using the method of Liang and Zeger (1986). The multinomial regression approach is suited to the analysis of discrete-time aggregated data that are usually produced by official statistical agencies. Brunsdon and Smith (1998) had used the same logistic transformation of the marginal probabilities (i.e. the logarithmic transformation of the ratios of probabilities) as Davis et al. (2001). However, they used ARIMA (autoregressive integrated moving average) modeling, rather than weighted least squares, to estimate the model parameters.

An important contribution to estimating work life expectancy was made in a methodological paper in the United States by Millimet et al. (2003). In the US study, Bureau of Labor Statistics work life tables were subdivided simultaneously by host of factors such as gender, race, and education, not only by a singular characteristic at a time. Pooling together multiple years of data, rather than using a single wave of the Current Population Survey, ensured that the estimates of working-life expectancy are not overly sensitive to the particular economic conditions that existed the year the data were collected. The data-analytic approach was to apply an econometric model, instead of a simple relative frequency calculation. The modeling strategy allows one to draw greatly more information about persons' working-life behavior and also
permits much more detailed working-life expectancy tables to be constructed. The US study was the first one to recognize explicitly the fact that because working-life tables are generated from survey data, sample variation may be significant.

The Millimet et al. (2003) study resembles the presented research in many respects. Their model, like that of Davis et al. (2001), explicitly incorporated three labor force states: employed, unemployed and inactive (out of the labor force). However, Millimet et al. estimated their multinomial model on three subsets of data for the working-life states. Similarly Davis et al. use a logistic transform to estimate probabilities, but the estimation of the multistate model parameters is done for the three states together by weighted least squares for a large-sample form of vector regression equation.

The major difference and the novelty of the method of Davis et al. (2001), compared to the related method of Millimet et al. (2003), is that it first proves the asymptotic normality of the empirical log-ratios. The next step is the estimation of the parameterized true log-ratios by way of generalized estimation equations. It is only possible to proceed in this way because the method deals with a large number of individuals. Millimet et al. did not exploit the large number of individuals and they used a standard package for maximizing the likelihood function. In a sense the method of Davis et al. (2001) is not logistic regression since it ends up with weighted least squares as opposed to solving non-linear likelihood equations by Newton-Raphson or some other numerical devise. That is why Davis et al. refer to their approach as a large-sample version of multiple regression.

Working-life expectancies are formally defined in terms of working-life table probabilities, thus they have a direct probabilistic interpretation. The working-life expectancy measure may be expanded to multistate working-life expectancies by
building it on the probability of occupancy in a given state amongst the multiple states of work ability or labor market activity (Nurminen et al. 2004a, Nurminen and Nurminen 2005). One can also construct a model for transition probabilities between labor or health states simultaneously and use the multistate life table (also known as increment-decrement life table) method (Nurminen et al. 2004b, Davis et al. 2007).

Davis’ multistate vector regression approach has the distinct property that it allows the summary to be multivariate. In other words, the parameter of the outcome state (i.e. the logarithm of the ratio of the probability of a given state to the probability of the referent state) is expressed in terms of multiple covariates of a temporal, spatial and socioeconomic nature. Given available data, the covariates may also represent alternative labor market or pension policies. An analysis of multistate working-life expectancies by using a multivariate regression model enables one to capture the joint impact of several simultaneously contributing causes of early retirement (e.g. via work disability). Nurminen (2012) utilized these attributes of the multistate regression method in order to gain a reliable picture of the multidimensional changes in Finnish working-life expectancies during the years 2000-2010.

Forecasting working life expectancies based on marginal probabilities for occupancy in labor states and setting prediction intervals around the estimates can also be done using multistate regression modeling approach (Nurminen 2012). Recently, Majer et al. (2013) presented a theoretical framework for a multistate life table model that projects transition probabilities by the Lee-Carter method, and illustrate how it can be used to forecast future health expectancy.
4. Comparing multistate and prevalence-based life table methods

For illustration we present results from a national population-based study conducted in Finland in which life table techniques were used for estimating expectancies from official statistics which are readily available (Nurminen 2012).

The methodological interest in this paper was in the appropriate usage of inferential tools for discrete time stochastic processes in practical research applications. The novel approach developed by Davis et al. (2001) for estimating working-life expectancies differs from the traditional Sullivan (1971) method in many fundamental facets. Although the advanced method uses data from the life tables and the annual Labour Force Surveys of Statistics Finland (OSF 2011), it estimates the working-life expectancies jointly for multiple years throughout the entire study period. An alternative approach that has been applied to the analysis is to carry out separate estimations for a series of survey or census years and then fit a curve to describe trends, as was done in Hytti and Nio (2004) in their monitoring of cross-sectional employment activity data over a number of years. The present analysis spanned eleven years (2000-2010) and a large number of individuals, so the results are not so sensitive to economic conditions as a survey that would rely on only single year of data.

Nurminen (2012) based the analysis of panel or cohort data on a large-sample regression model fitted to a multistate life table, instead of a simple relative frequency calculation from the average demographic experiences of artificial cohorts (constructed using an arbitrary radix for the number of survivors at each given age). This stochastic inferential approach allows one to draw probabilistic inferences on several work life characteristics and also permits much more detailed working-life tables to be estimated, for example stratified by socioeconomic factors. The study
modeled the state probabilities as a function of polynomials of age and year, and Gross Domestic Product. The set of variates describing demographic and economic conditions faced by persons can be expanded, but not at will. This modeling approach enables one to circumvent the problem of small cell sizes encountered in a modest disaggregation of data.

The traditional prevalence life table technique (Sullivan 1971) is limited when applied to intrinsically dynamic processes with multiple decrements, like the labor force process. The life table calculated from prevalence rates cannot provide the occurrence/exposure rates in a continuous time frame. If labor force participation rates change over time, these trends are incorporated more accurately in the multistate life table method than in the prevalence-based technique. The latter method is very sensitive to particular fluctuations in labor force activities. Calculations could therefore overstate the labor force involvement at times of expansion and understate it during a recessionary period (Richards 2000).

The multistate regression methods were developed to overcome the limitations of the traditional prevalence techniques. The states are defined to be multiple, some of which are transient (or recurrent) while others are assumed non-transient. In the quoted study (Nurminen 2012), the customary life table was enhanced by explicitly defining a three-state employment state space: (1) 'employed' (permanently employed, employed for fixed-term, and self-employed); (2) 'unemployed'; (3) 'outside the labor force' (students, conscripts, disability and old-age pensioners, etc.). In the context of this analysis, the sum of the working-life expectancy and the expectations of time spent in the states 'unemployed' and 'economically inactive' is equal to the partial life expectancy between exact ages 15 and 65 years, as normally defined by demographers. This decomposition is different to the two-state system
which estimates the duration of 'active working life' by classifying persons as 'active' (in the labor force) or 'inactive' (out of the labor force) (Hytti and Nio 2004). The relative frequency approach relies on age group or other subgroup comparisons and calculates the (marginal or transition) probabilities from average behaviors of the sample at each age. The tabular analysis of further disaggregated data (e.g. by allowing various modes of exit from the labor force) would necessarily turn out to be cumbersome or impossible without resorting to modeling. The regression analysis of panel or cohort data is applicable when the numbers are reasonably large.

The results evinced that the duration of working lives in Finland have extended positively for both genders in the 2000s (Table 1). For a 15-year-old man the expected length of working career up to age 64 in 2010 was 34.6 years, while for women it followed at 34.0 years. The favorable trends are forecast to continue up to 2015 under the provision of economic equilibrium (Nurminen 2012).

The working-life expectancy measure has been characterized as being sensitive to volatile labor market variations in a report of the working group for lengthening working careers (Prime Minister’s Office 2011). For example, in 2008 the Finnish working-life expectancy at age 15 was 34.6 years but it decreased due to the rapid decline in employment in the recession year 2009 by one whole year. Actually, this expectancy was computed using the traditional (Sullivan 1971) technique on a year-by-year basis. The multistate vector regression (Davis et al. 2001) approach to expectancy, which is based on fitting a smooth model over the studied interval, herein 2000-2010, does not overestimate the effect of such changes on the total length of working career (Figure 1). In this study (Nurminen 2012), the model yielded the following estimates of male working-life expectancies for the years 2008, 2009,
and 2010: 34.5, 34.2, 34.6. The drop from 2008 to 2009 was only 0.3 years and the counteractive rise from 2009 to 2010 was 0.4 years.

**Table 1:** Working-life expectancies, in years, by gender for the age span of 15-74 years in the study period 2000-2010. 

Comparison of the cohort-type and current life table estimates computed by the methods of Davis et al. (2001) and Sullivan (1971), respectively. 

<table>
<thead>
<tr>
<th></th>
<th>Males</th>
<th>Year</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Method</td>
<td>2000</td>
<td>2001</td>
<td>2002</td>
<td>2003</td>
<td>2004</td>
<td>2005</td>
<td>2006</td>
<td>2007</td>
<td>2008</td>
<td>2009</td>
</tr>
<tr>
<td></td>
<td>Davis et al.</td>
<td>34.2</td>
<td>34.1</td>
<td>34.2</td>
<td>34.2</td>
<td>34.3</td>
<td>35.3</td>
<td>35.4</td>
<td>35.5</td>
<td>35.5</td>
<td>35.2</td>
</tr>
<tr>
<td></td>
<td>Sullivan</td>
<td>32.4</td>
<td>32.8</td>
<td>32.5</td>
<td>32.5</td>
<td>32.7</td>
<td>33.1</td>
<td>33.8</td>
<td>34.5</td>
<td>35.2</td>
<td>33.5</td>
</tr>
<tr>
<td></td>
<td>Difference</td>
<td>1.8</td>
<td>1.3</td>
<td>1.7</td>
<td>1.7</td>
<td>1.6</td>
<td>2.2</td>
<td>1.6</td>
<td>1.0</td>
<td>0.3</td>
<td>1.7</td>
</tr>
<tr>
<td></td>
<td>Females</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Method</td>
<td>2000</td>
<td>2001</td>
<td>2002</td>
<td>2003</td>
<td>2004</td>
<td>2005</td>
<td>2006</td>
<td>2007</td>
<td>2008</td>
<td>2009</td>
</tr>
<tr>
<td></td>
<td>Davis et al.</td>
<td>31.5</td>
<td>31.8</td>
<td>32.1</td>
<td>32.4</td>
<td>32.7</td>
<td>33.0</td>
<td>33.4</td>
<td>33.7</td>
<td>34.1</td>
<td>34.4</td>
</tr>
<tr>
<td></td>
<td>Sullivan</td>
<td>30.6</td>
<td>31.1</td>
<td>31.6</td>
<td>31.5</td>
<td>31.5</td>
<td>32.2</td>
<td>32.8</td>
<td>33.5</td>
<td>34.0</td>
<td>33.7</td>
</tr>
<tr>
<td></td>
<td>Difference</td>
<td>0.9</td>
<td>0.7</td>
<td>0.5</td>
<td>0.9</td>
<td>1.2</td>
<td>0.8</td>
<td>0.6</td>
<td>0.2</td>
<td>0.1</td>
<td>0.7</td>
</tr>
</tbody>
</table>

* Source: Nurminen (2012), Table 9.2.1.

§ The age group 65-74 years added to the working-life expectancy computed for the age interval 15-64 years 1.1 years for men and 0.8 years for women in 2010.

# In its strictest form, *cohort* (working-) *life table* records the actual experience of a particular group of individuals (the worker cohort) from a specific age to (the final retirement from the labor force or) death. The *current or period* (working-) *life table* considers the experience of a given population during one short period of time, for example, the resident population of Finland in 2010.
Figure 1:* Current life table and model-based estimates of working-life expectancy for the age span of 15 to 74 years by gender in 2000-2011. §

* Source: Nurminen (2012), Figure 9.2.1.

§ The current life table estimates were computed using the Sullivan (1971) method and the model-based estimates with the method of Davis et al. (2001). The values for the current year 2011 are preliminary estimates.

Finally, because working-life tables are generated from survey data, sampling variation may be important (e.g. due to population dynamics, economic fluctuations, interview methods), especially in small samples. Although the Finnish official research institutes acknowledge this fact, they do not provide standard error estimates for their active working life expectancies (Kannisto 2006). Under stationary conditions (i.e. independence of an initial health state), a new 'equilibrium' estimate of the prevalence rate and its approximate variance has been developed by Diehr et al. (2007). In the Davis et al. (2001) approach, standard errors (and covariances) can be found by using the delta method based on the loss function or alternatively estimated regression coefficients. We point out that standard errors, based on either
by Monte Carlo sampling from the estimated asymptotic normal distribution of the
delta method or Monte Carlo simulation, only reflect the inaccuracy in model
estimation from observed prevalence rates; these rate estimates have an inherent
variance because they are obtained from a sample survey (rather than a census).
That is, our standard error calculation does not incorporate the variance in the
estimated rates.

To contrast results obtained with these two fundamentally different methods Table
1 presents working-life expectancy estimates obtained for the 15-74 span of age for
the years 2000-2010. The regression model-based estimates indicate lesser
variation than the corresponding prevalence estimates for both genders. The current
working-life table estimates designate large fluctuations especially for males around
the economic upturn year 2008. Both series of estimates show lower figures for
males in the economic recession year of 2009 compared to the neighboring years.
The recession affected foremost men's employment. Preliminary statistics pointed to
the fact that the recovery was delayed until 2011 among women. A finding is that
that the absolute difference between the two sets of expectancies has risen and then
fallen over time.

Yet these estimates are parallel taking into consideration the basic
methodological differences in the estimation approaches: viz. modeling versus
tabulation. The two methods differ in that the multistate regression estimates are
derived from cohort life table data for rates of change between the states (of being
either (un)employed in the labor force or being outside it, or finally leaving the
population through death), whereas the current life table estimates do not reflect
dynamic changes but instead are computed directly from annual labor force
participation rates during the period of sampling. Theoretically, the two methods
should give identical results if the populations were stable and age-specific transitions did not alter from year-to-year. Current life tables formed in a recessionary period, during which labor force exits increase, present a bleak picture of working-life involvement. Conversely, those calculated during a subsequent period of recovery tend to exaggerate labor force attachment (Smith 1982).

Uncertainty of the expectancy figures is inherent in any statistical estimation. Debating whether the actual working-life expectancy in a given year and a certain sub-population is exactly 35.5 or 35.2, or which method is somewhat less accurate, is missing the essence, unless one wants to look at a 'snapshot' rather than a trend to gain a wider perspective. Almost any period chosen for observation provides somewhat inaccurate (biased) working-life expectancies for at least some demographic groups. Nevertheless, the discrepancies between the working-life expectancies estimated using the two different methods are important and large, e.g., for year 2010: 2.0 y for males and 1.5 y for females in excess for the model-based estimates versus prevalence.

5. Discussion

Working-life expectancy is a period or cohort measure, depending on whether cross-sectional or longitudinal data are available. Usually the latter data are not readily accessible or are prohibitively expensive. Real people age and die as members of cohorts through successive periods subject to ever-changing rates. Period life expectancy is constructed as an entirely synthetic measure, referring to a fictitious cohort living its whole life according to the rates of a single period. Thus working-life expectancies, defined in this study abstractly through imaginary cohorts, can be paralleled with the experience of real cohorts, and makes them more tangible and
easier to interpret. Davis et al. (2001) remarked, "Sullivan's method contains both period and cohort considerations but predominantly the former. The cohort element arises typically from surveys designed to estimate prevalences of health [or labor force] states whereas the period component is due to the use of standard life tables."

The major novelty of the multistate methodology is in its way to reconstruct the relevant elements of the longitudinal stochastic process that generated the working-life table datasets from sequential cross-sectional surveys. This was made possible by estimating the marginal probabilities using a weighted least squares procedure under a multistate vector regression model, and this in turn led to estimates of cohort expectancies. The goal of Davis et al.'s approach is not to reconstruct the full increment-decrement system, which they recognize cannot be recovered with their methodology, but to estimate unconditional health expectancies, as in the Sullivan method. They use successive cross-sections not to estimate inter-survey period conditions, but to track actual cohorts as they appear in these cross-sections and estimate corresponding cohort health expectancies, with the aim of recovering the multi-state system prevalent during the inter-survey period. Thus the Davis et al. approach is very different to the traditional Sullivan technique that does not yield a cohort measure from cross-sectional data.

Mathers (2000) has noted, "The problems with Sullivan's method arises not because it uses prevalence and mortality data averaged over all health states, but because the data it uses are dependent on past conditions in the population [p 190]". Therefore, the Sullivan method for calculating period expectancy does not either produce a 'pure' cross-sectional indicator derived from the current prevalence rates summarizing the experience of a population at a point in time. The period expectancy computed from standard life tables, for example at the age of 25 for
Finnish men in 2010, is a pure cross-sectional indicator only in the sense that it gives the expectation of remaining employment time for persons who experience at each age of their life the risk of moving outside the workforce observed for that age for Finnish men in 2010.

There has been some debate in the literature about the magnitude of the bias in the Sullivan method; see Guillot and Yu (2009) for more references. Certain authors have concluded on the basis of actual data that the differences between observed and equilibrium proportions healthy are significant, and that health expectancy estimates based on the Sullivan method should not be used for conclusions on compression or expansion of morbidity (Barendregt et al. 1997, Lievre et al. 2003). On the other hand, a simulation study indicated that Sullivan’s method provides acceptable estimates of the period working-life expectancy if the changes in transition rates over a reasonably long term are smooth and fairly regular (Mathers and Robine 1997). In particular, the repeated application of the method can provide good estimates of trends in health and work life expectancy. However, there is no guarantee that the assumption of the Sullivan method holds in real situations. Moreover, age-specific transition rates and conditional health expectancies cannot be estimated with the Sullivan method. But as Davis et al. (2001) pointed out, “By definition the Sullivan method as described cannot supply these estimates [of cohort expectancies], except in so far that a period measure is a surrogate for the analogous cohort quantity [p 1099].”

As has been insightfully recognized by Goldstein and Wachter (2006), period and cohort expectancies can be interpreted as a time-delayed (lagged) measure of the experience of real cohorts in populations undergoing steady demographic changes such as mortality improvement and aging. They showed the correspondence
between period life expectancy at birth and the life expectancy of particular cohort. Their approach was to look at the relationship between period and cohort life expectancy in terms of two measures: (1) the lag of years by which the equivalent period and cohort life expectancies are observed, and (2) the gap in life expectancy that a cohort gains by experiencing improving rather than steady mortality.

For countries with mortality rates close to those of the present day Finland, today's period life expectancy at birth summarizes the expected longevity of people born about 40 to 50 years ago who are or would be now in the prime of middle age. Current life expectancy at age 65 matches expected survival beyond 65 for people who celebrated their 65th birthday with a lag of some 15 years ago (cf. Goldstein, 2006, Myrskylä, 2010).

The presented study (Nurminen 2012) found that the marked difference (gap) between period (the Sullivan method) and cohort (the Davis et al. method) working-life expectancies at a moment in time rose and shrank over the last decennium (the 2000s). The likely reason for this observable fact is the volatility of market forces.

Compelling empirical evidence of this relationship was previously provided by the prospective follow-up study of an actual cohort of active Finnish municipal workers at three successive surveys (Nurminen et al. 2004a,b). The data were analyzed both as aggregate cross-sectional data for estimating marginal probabilities and as an individually-linked longitudinal discrete-time aggregate data for the estimation of transition probabilities between four work/heath states. The index expectancy state was defined as currently/continued having excellent or good work ability. A comparison of the working-life expectancies gave the following estimates for a 45-year-old man up to retirement age of 62.5 years in 1981: 7.3 years (transition probabilities), 5.5 years (marginal probabilities). The gap favored, by 1.8 years, the
working-life expectancy based on transition probabilities, calculated conditional on having been in the index state at the preceding time point.

6. Conclusions
To summarize the differences between the applied regression estimates that simulate the cohort expectancies (Davis et al. 2001) and those obtained by the current life table technique (Sullivan 1971) we underscore two points. First, the Finnish study (Nurminen 2012) used estimates of the multistate probabilities based on a parametric model rather than non-parametric working life table estimates which do not take into account the decline in mortality over time. Second, the estimated state occupation times pertained to a particular birth (age) cohort, instead of the study a particular point in given period of calendar time (year). We conclude that cohort measures are preferable because they are more relevant to persons now living and to planners of future health services.

Acknowledgments
I wish to thank Dr. Brett A. Davis, Australian Institute of Health and Welfare, Canberra, Australia, for valuable comments on the multistate regression approach.

References


