Journal of Forensic Economics 17(2), 2004, pp.167-183

The Markov Assumption for Worklife Expectancy

Edward M. Foster and Gary R. Skoog*

I. Introduction

A. Objective for this paper

This paper is intended to provide a general understanding of:

- how worklife expectancy is calculated,
- the nature and role of the assumption used by the Bureau of Labor Statistics that worklife transitions are governed by a Markov process,
- how worklife expectancy is used,
- criticisms of the Markov assumption, and
- alternatives that at least some forensic economists use.

B. What is worklife expectancy and how is it determined?

The worklife expectancy at age x (WLE_x) is the average number of years that a person in a given cohort will spend either working or actively looking for work during the remainder of his or her life. In tables calculated by the U. S. Department of Labor's Bureau of Labor Statistics (BLS), cohorts have been distinguished by sex, age, and either race (white v. all other) or for three groupings by level of education (see Smith, 1986, p. 1). The BLS did not have sufficiently large samples to permit classification by finer educational groupings or for combined race and education, using their estimation method.

BLS tables are constructed from information collected in the *Current Population Survey* (CPS), "a nationwide monthly household survey conducted by the Bureau of the Census on behalf of the BLS" (Smith, 1986). By matching interviews with the same person a year apart, the BLS knows whether each individual in the sample is active or inactive on each of the two dates. Of those who are classified as active in the first interview and who remain in the survey a year later, some remain active for the year, and others exit from the labor force; some people disappear from the sample, though, because they have moved, refuse re-interviews, could not be matched or die during the year. Of those who start the year as inactive and remain in the survey a year later, some remain inactive and others enter the labor force; those who disappear from the sample have either moved, refused the second year's interviews or have died.

*Respectively, Department of Economics, University of Minnesota, Minneapolis, MN; Department of Economics, DePaul University and Legal Econometrics, Inc., Glenview, IL. Thanks to Jim Ciecka and Kurt Krueger for helpful comments; the authors are responsible for remaining errors.

The worklife tables are built on the assumption that the numbers who remain active or inactive and those who enter and exit the labor force reflect the probabilities of making those transitions. The CPS data cannot be used to estimate the probability of death because the CPS does not track those who disappear to determine who died; however, reliable mortality statistics are available from the US Life Tables (Arias, 2002).

The worklife tables also assume that probabilities of transition observed in one year will remain unchanged in the future. That enables the BLS to assume that the transition probabilities observed for a 50-year-old this year will apply to a current 30-year-old 20 years hence when he or she reaches 50. The assumed stability of labor force transition probabilities over time parallels the assumed stability of rates of mortality over time embedded in the U. S. Life Tables. Both survival rates and female labor force participation have increased over time, so that it is important to use the most recent valid tables for them. Male labor force participation has declined and risen since the BLS tables were published, returning to its early 1980's levels (Ciecka, Donley, et al., 1999-2000). To be useful for projection, transition probabilities and participation rates should be "typical" with respect to the business cycle, so that the latest time period is not necessarily the best if it reflects a peak or trough. Although the Markov model requires data on individuals one year apart, there is no reason that one cannot create a large dataset of such years spanning the business cycle.

C. What is a (first order) Markov process?

A "process" in general simply means a collection of variables whose values evolve over time, and the rules under which they evolve. A *stochastic* process evolves with some random element–a dice roll, say, or random numbers generated according to some other rule–helping to determine that evolution. The BLS worklife table keeps track of only three values that the variables can take on: active (in the labor force), inactive, or dead, and uses discrete time units of a year to track movements.

A (first order) *Markov* process is a stochastic process with only a one-period memory, in the sense that transition probabilities from its state at time t to its state at time $t+1$ depend only on the current state, not on past history. In the language of statistics, the present state is a *sufficient statistic* for the prediction of the transition probabilities between this period's observation and the next.

The state of a stochastic process typically varies over *time*, but in calculating *WLE*, the variable is *age*. Varying with age superficially seems the same as varying with time, but an increase in age changes the *WLE* transition probabilities. So Markov models of *WLE* are more complex than models with transition probabilities that are constant (homogeneous) over time, discussed in standard texts. Two classic references for the homogeneous case are Cox and Miller (1965, pp. 1-248) and Feller (1968, Ch. 15-16).

D. Is the Markov process realistic for calculating worklife expectancy?

Chaos theory has taught us that some events are truly random: even knowing the value of a causal variable to 10 places of accuracy may not allow one to predict the outcome–the outcome may be determined by the difference of one digit in the hundredth, or the thousandth, place, with tiny changes in the causal variables leading to large changes in outcome (the "butterfly effect"). However, social scientists generally start from the self-confident position that using the motivations explored by physiology and the social sciences, peoples' behavior can be explained, if only we know enough about their background. From that viewpoint, probability models of behavior are a confession of ignorance. It is not that the behavior lacks underlying physical or psychological causes; rather that the observer cannot see them.

The Markov assumption for a worklife table means we assume that the probability that someone in a given cohort will be active next year depends only on whether she was active or inactive this year. It doesn't matter that she had 20 previous years as an active member of the labor force; her chances of being inactive next year are the same as someone who had been in and out of the labor force 10 times in the same period.

If the purpose of the table is to summarize the average behavior of an entire cohort, the Markov assumption seems perfectly appropriate. However it may not be if the purpose of the table is to forecast the behavior of a particular individual. The probability of staying in the labor force is not the same kind of random process as the probability of getting another head in a series of coin tosses. In the first case many well-understood socio-economic, physical, and psychological circumstances unique to the individual could go a long way to explaining the decision she makes; since the BLS did not have the capacity to take all of those factors into account, it treated them as "random." The Markov model of labor force participation developed by the BLS treats all members of a particular cohort (based on age, sex, and education) as though they were balls in one of two urns at the beginning of each year. A certain number of balls chosen at random will be removed from the active urn to the inactive, and a few more will be removed to the dead urn. At the same time some will be chosen at random from the inactive urn for transfer to the active urn, and others to the dead urn. In each case, the model says that every individual in a given urn has the same chance to be chosen for transfer as every other one.

In reality, there is not just one urn of active members of the cohort, there are many, and the chance of being chosen for transfer to the inactive state depends on which urn you are in. The individuals in these urns vary by health, by occupation, by motivation to work and a host of other variables that affect the probability of leaving the labor force. Moreover, some individuals are switched from urn to urn within the year within the given state: they switch marital states, health levels, or change their levels of motivation. This heterogeneity, unobserved by the analyst, will bias the resulting worklife projection if the individual in question is not representative of the class being treated as homogeneous.

E. Why use the Markov assumption?

One uses the Markov assumption because the individual is judged to be reasonably representative of the class, or for lack of anything better. To be applied to real life, a model is constrained to use only the information that is available.

In practice, even if you know a lot more about the people in the sample than is used in *WLE* tables, there won't be enough of them to be able to divide them into groups by detailed combinations of race, level of education, occupation, health habits, etc., to get reliable estimates of transition probabilities for these fine groupings.[1] Alternatives to the Markov model are discussed in Part V, below.

II. The Elementary Mathematics of Worklife Expectancy Tables

A. Definition of life expectancy

Standard actuarial calculations are performed for a *stationary population*: a mental construct based on the assumptions, first that the number of births is constant from year to year and second, for each age the probability of death is constant over time, so that the age distribution of the population is constant over time. A standard simplifying assumption is that on the average those who die between integer age x and $x + 1$ will die at age $x + 0.5$. Where ℓ_x represents the number in a stationary population living at exact age x, the average number of people alive over the 12 months after the xth birthday is $(\ell_x + \ell_{x+1})/2$. This is also the total number of years lived during that year by the ℓ_x people alive at the beginning of the year, which we record as:[2]

(1) $$L_x = (\ell_x + \ell_{x+1})/2.$$

Add the total number of years to be lived over all future ages by those aged x today, divide by the number in the cohort at age x, and the result is the life expectancy: the average number of years yet to be lived by a person who is exactly x years old today. In symbols,

(2) $$e_x \equiv \sum_{t=x}^{\varpi} L_t \Big/ \ell_x \equiv T_x / \ell_x$$

Here ω (omega) represents the oldest age to which anyone lives, and T_x represents the sum of L_t, for $t = x$ to ω. The equal sign with three bars, or identity symbol, may be read "is identically equal to," or "equals by defini-

---

[1]Richards and Abele (1999) have presented a *WLE* table based on the Markov model that combines race, Hispanic origin, and education; see their Table 4. They also present *WLE* tables using the method used by BLS prior to 1982 that give much finer detail including smoking status and occupation. A discussion of the relative advantages of the older methodology and the Markov model is beyond the scope of this paper; see Smith (1982a), Richards (2000), and Skoog and Ciecka (2004).

[2] ℓ_{x+1} live through the year, and $\ell_x - \ell_{x+1}$ die during the year, living for 1/2 year on average. So $L_x = \ell_{x+1} + (\ell_x - \ell_{x+1})/2$, which can be re-arranged to give equation (1) in the text.

tion." The life expectancy tables for the U.S. population, published annually by the National Center for Health Statistics, provide columns for (among others) ℓ_x, L_x, and T_x in addition to the resulting life expectancy, e_x (see Arias, 2004, Table 1). The cohort under consideration could be all people of age *x*, all males, all females, or subgroups classified by race, education, or any other characteristic(s) considered relevant provided they may reasonably be assumed to remain the same as the individual ages.

The link between numbers living at the beginning and end of the year depends on the probability of death during the year after attaining age x, q_x. The link is given by:

(3) $$\ell_{x+1} = \ell_x \cdot \left(1 - q_x\right).$$

B. Definition of worklife expectancy

Actuaries embroider symbols with subscripts and superscripts, before and after the primary symbols. The people living at exact age *x* may be classified into those who are active, ${}^{a}\ell_{x}$, and those who are inactive, ${}^{i}\ell_{x}$. In the worklife table context it is customary to write ${}^{\bullet}\ell_{x}$ to represent what is denoted in the life table context by ℓ_x:

(4) $$\ell_x \equiv {}^{\bullet}\ell_x \equiv {}^{a}\ell_x + {}^{i}\ell_x .$$

The pre-superscript dot signals that all states for living members of the system, active and inactive, are included. In the BLS tables, "active" means "in the (civilian) labor force," whether employed or unemployed. "Inactive" means "not in the labor force," that is to say neither working nor actively looking for work.

Worklife expectancy at age *x* is then defined as the average number of years that a cohort of age *x* will be in the labor force, before they either retire permanently or die. Thus a first point to notice is that worklife expectancy does not tell us how many years a person can be expected to *work* in the future, but rather how many years he or she can be expected to either *work or look for work*.[3] A second point is that no distinction is made between full-time and part-time work. A person who was "economically active" –that is working or looking for work, full- or part-time– at some time during the week before the interview (the reference week) is classified as in the labor force.[4]

Those in the labor force are on the average healthier than the rest of the population, so rates of mortality undoubtedly differ between the two groups. However no mortality table records these differences, so the worklife expec-

---

[3]It would be feasible to measure employed v. unemployed or inactive rather than active v. inactive, and below we refer to studies that do so (see Flinn and Heckman, 1981 and Millimet, et al., 2002). The BLS made a judgment that those who are unemployed but actively looking for work would behave more like the employed than like those who are not in the labor force.

[4]To be classified as unemployed, one must have had no employment in the reference week but been available for work, except for temporary illness, and searched for work sometime during the four previous weeks. Persons waiting to be recalled from layoff need not have been searching.

tancy tables assume that the active and inactive populations face the same probability of dying.

Those who are active at both the beginning and end of the year, called *actives*, are assumed to be active all year. Those who are active at the beginning and inactive at the end of the year, called *exits or separations*, are assumed (on the average) to be active for half a year; the same assumption is made for those who are inactive at the beginning but active at the end, called *entrants or accessions*. Those who are inactive at both the beginning and the end of the year, called *inactives*, are assumed to be inactive all year.

We assume that entrants and exits change their status, on the average, in the middle of the year, so just as in equation (1), the total number of active years lived during the year by the ${}^{\bullet}\ell_x$ individuals alive at the beginning of the year is given by:

(5) $${}^{\bullet}L_x^a = \left({}^a\ell_x + {}^a\ell_{x+1} \right) \big/ 2 .$$

Add the total number of active years to be lived over all future ages by those aged *x* today, divide by the number in the pool at age *x*, and the result is the worklife expectancy. In symbols,

(6) $${}^{\cdot}e_x^a \equiv \sum_{t=0}^{\omega} {}^{\cdot}L_{x+t}^a \Big/ {}^{\cdot}\ell_x .$$

The dots in the pre-superscripts remind us that the definition is for the average worklife expectancy of an individual at age *x* without regard to current labor-force status. After the development through (12) below, this may be shown to be a weighted average of the worklife expectancies for those who start as active at age *x*, and those who start as inactive.

If a forensic economist is applying worklife expectancy to an individual, (6) is the appropriate concept to use for a child not yet old enough to be in the labor force, or for an adult whose labor-force status is unknown. But it is not appropriate to apply ${}^{\cdot}e_x^a$ to an individual whose labor-force status is known.

The appropriate worklife expectancy concept when labor-force status at age *x* is known is ${}^a e_x^a$, if active, or ${}^i e_x^a$, if inactive–the symbols are defined at (9) and (12). Worklife expectancy for those in the labor force will be higher than for those who are not. Calculations appear complex, because so many possible transitions must be accounted for, and their probabilities assessed.

Six year-to-year transition probabilities are needed. Two give the probability of dying while active and inactive; these are assumed to be equal to q_x, the probability that an active or inactive individual will die during the year. The other four transition probabilities, for those who remain alive, are:

(7) $${}^1p_x^2 :$$

the probability that an individual in state 1 at age *x* will survive and be in state 2 at age *x*+1, where states 1 and 2 can each stand for *a*, active, or *i*, inactive. These are related to the transition probabilities conditional upon survival, ${}^{1}\pi_x^2$, by:

$$ {}^{1}p_x^2 = (1-q_x)\,{}^{1}\pi_x^2 . $$

Transition probabilities are estimated from the linked Current Population Survey data of years t and t+1 as follows (the "hat" in $\hat{\pi}$ indicates that it is a statistical estimate of π):

$$ \text{8(a)} \qquad {}^{i}\hat{\pi}_x^i = \frac{\text{inactives}}{\text{inactives+entrants}};\ {}^{i}\hat{\pi}_x^a = \frac{\text{entrants}}{\text{inactives+entrants}}; $$

$$ \text{8(b)} \qquad {}^{a}\hat{\pi}_x^i = \frac{\text{exits}}{\text{actives+exits}};\ {}^{a}\hat{\pi}_x^a = \frac{\text{actives}}{\text{actives+exits}}. $$

The denominators of (8a) represent all of the inactives at age x *who were sampled and matched in the following year*, while the denominators of (8b) represent all of the actives at age x *who were sampled and matched in the following year*. These estimated probabilities, conditional on the initial state, obviously sum to 1 for each row, and mimic the population counterparts they seek to estimate. The matching process in the CPS sample requires that each subject remain in the sample, not dying and not moving from his initial residence. In the next (un-numbered) equation, the four groups (inactives, entrants, exits, and actives) refer to the cohort who had attained age x in the first year of the linked observations and remained in the sample and was matched at age $x + 1$.

Using the relation specified in (7) between ${}^{1}p_x^2$ and ${}^{1}\pi_x^2$, the estimated transition probabilities incorporating mortality are:

$$ {}^{i}\hat{p}_x^i = (1-q_x)\,{}^{i}\hat{\pi}_x^i;\ {}^{i}\hat{p}_x^a = (1-q_x)\,{}^{i}\hat{\pi}_x^a; $$
$$ {}^{a}\hat{p}_x^i = (1-q_x)\,{}^{a}\hat{\pi}_x^i;\ {}^{a}\hat{p}_x^a = (1-q_x)\,{}^{a}\hat{\pi}_x^a. $$

Corresponding to the four transition probabilities there are four expectations,

$$ \text{(9)} \qquad {}^{1}e_x^2 : $$

the expected number of years that an individual in state 1 at age x will spend in state 2 over the remainder of life, where states 1 and 2 can each stand for a, active, or i, inactive.

The relation between these expectations and the expectation of life at age x is given by:

$$ \text{(10)} \qquad {}^{1}e_x^a + {}^{1}e_x^i = e_x , $$

where 1 can stand for a or i: All of the remaining life expectancy will be spent in either the active or the inactive state.

Starting at age x we may let the radix ${}^{a}\ell_x$ be arbitrarily assigned (say, 100,000), while ${}^{i}\ell_x$ is set to 0. Then the following (matrix) equation will generate the number of actives and inactives at the next age.

$$\begin{pmatrix} {}^{a}\ell_{x+1} \\ {}^{i}\ell_{x+1} \end{pmatrix} = \begin{pmatrix} {}^{a}p_x^{a} & {}^{i}p_x^{a} \\ {}^{a}p_x^{i} & {}^{i}p_x^{i} \end{pmatrix} \begin{pmatrix} {}^{a}\ell_x \\ {}^{i}\ell_x \end{pmatrix} \tag{11}$$

Repeating this equation with x+1 substituted for x will generate the actives and inactives at age x+2, etc. Applying (5) and (6) with this initial condition, but with the "dots" on the upper left superscripts replaced by a for active, will give the first column in the expectancy matrix below. Repeating this procedure with ${}^{i}\ell_x$ as a nonzero radix and ${}^{a}\ell_x = 0$ will give the second column of the following expectancy matrix E_x:

$$E_x = \begin{pmatrix} {}^{a}e_x^{a} & {}^{i}e_x^{a} \\ {}^{a}e_x^{i} & {}^{i}e_x^{i} \end{pmatrix} = \begin{pmatrix} {}^{a}t_x^{a} / {}^{a}l_x & {}^{i}t_x^{a} / {}^{i}l_x \\ {}^{a}t_x^{i} / {}^{a}l_x & {}^{i}t_x^{i} / {}^{i}l_x \end{pmatrix}. \tag{12}$$

The ${}^{1}t_x^{2}$ notation in (12) refers to the sum of the ${}^{1}L_x^{2}$ to the terminal age of the life table, shown in (2) as T_x; the superscripts 1 and 2 denote the initial state and state being measured, respectively.

Equations (11) give the typical exposition, showing how the process evolves in the "forward" direction. However, by starting at the final age and proceeding "backwards" to earlier ages, Skoog (2002) has shown that the following recursion holds, where P_x is the matrix of transition probabilities shown in (11):

$$E_x = \frac{1}{2}I + \frac{1}{2}P_x + E_{x+1}P_x\,. \tag{13}$$

This may be viewed as a computing algorithm that can be used to calculate the expectations given in (10), tracing through all of the alternative possible transitions.[5] Having obtained E_{x+1}, use equation (13) to obtain E_x. (14a) and (14b) give two of the four equations in (13), which together allow calculation of worklife expectancy for active and inactive persons at any age.

$$14(a) \qquad {}^{a}e_x^{a} = \underbrace{0.5 + 0.5\,{}^{a}p_x^{a}}_{\text{First year's contribution to } e} + \underbrace{{}^{a}p_x^{a}\,{}^{a}e_{x+1}^{a} + {}^{a}p_x^{i}\,{}^{i}e_{x+1}^{a}}_{\text{Future years' contribution to } e}$$

$$14(b) \qquad {}^{i}e_x^{a} = \underbrace{0.0 + 0.5\,{}^{i}p_x^{a}}_{\text{First year's contribution to } e} + \underbrace{{}^{i}p_x^{a}\,{}^{a}e_{x+1}^{a} + {}^{i}p_x^{i}\,{}^{i}e_{x+1}^{a}}_{\text{Future years' contribution to } e}$$

The first term in (14a) shows that everyone who starts the year as active generates at least 0.5 years of activity (worklife expectancy), because those

---

[5]For the oldest age, call it ω, $E_\omega = 0$, which allows the calculation of $E_{\omega\text{-}1}$ in (13) to start the recursion.

who change states do so on the average in the middle of the year. In addition, those who remain active, represented by ${}^{a}p_{x}^{a}$, generate an additional 0.5 years of activity. Then, in the future, those who remain active all year, ${}^{a}p_{x}^{a}$, will have the additional worklife expectancy of those who are active at the beginning of the next age, ${}^{a}e_{x+1}^{a}$, while those who become inactive, ${}^{a}p_{x}^{i}$, will have the additional worklife expectancy of those who are inactive at the beginning of the next age, ${}^{i}e_{x+1}^{a}$. The logic for (14b) is identical except that we start with those who are inactive, who will generate no years of activity *unless* they are among the fraction ${}^{i}p_{x}^{a}$, which become active during the year, in which case they will generate 0.5 years of activity.

The other two equations generated by expanding (13) give the expected number of inactive years for each of the two groups, those who are currently active and inactive.

C. Appropriateness of the Markov model for projecting the behavior of a single individual

Labor economists have studied workers' degree of attachment to the labor force, and know that it varies systematically with age, education or race, and sex, as recorded in the BLS worklife tables. However it also varies with wealth, health, marital status, number of dependants, wage rates, alcohol and other drug use patterns, marginal tax rates, and occupation.[6] The BLS tables do not track dependence on those variables because some of the information is not collected, because even for the information that is collected, sample sizes are inadequate to estimate the effects of all those variables using the BLS methodology, and because some of the statuses may change.

We should acknowledge, too, that even if these problems were overcome, the most careful economic study would not fully explain all of the variation we see in individuals' patterns of labor force participation. For given values of all the variables listed above, there will still be individual variations in behavior: some people are lazy, and others are industrious. We might describe those variations ("heterogeneity") as due to unobserved explanatory variables or as due to chance, but in either case, as in labor economics generally, econometric fits are not perfect and there is unexplained variation.

There are also regime changes: individuals become disabled or healed, reform their lives or suddenly throw away their opportunities, and make abrupt shifts in behavior that could not be predicted from the past.

There are yet two more issues if the *WLE* information is to be applied to project lost earnings due to death or disability.

First, some departures from the labor force are voluntary; if the appropriate legal standard of loss is earning capacity, such voluntary departures,[7] at least those before "final" retirement, arguably would not be appropriately

[6]The basic reference for an economic theory of labor force participation is Gary Becker (1965). Also see recent labor economics textbooks, e.g., Ehrenberg and Smith (2003, Ch. 6, 7) or Kaufman and Hotchkiss (2000, Ch. 3).

[7]Does a person who "voluntarily" drops out of the labor force because he is "discouraged" or "burned out" have the "capacity" to participate? Must "earning capacity" be reasonably certain to occur or to be realized?

considered, in which case the *WLE* table would not measure the appropriate worklife capacity concept.[8] This consideration is particularly important for women, most of whom leave the labor force at least temporarily for childbirth and many of whom leave for longer periods while raising families. However men, too, undertake voluntary departures from the labor force. Some such voluntary departures are for education intended to increase earning capacity, typically early in ones career. However, "capacity" in many jurisdictions is accompanied by the words "reasonably likely to occur," which brings us back very close, if not exactly, to worklife expectancy.

Second, the loss of earnings due to death or disability is not accurately measured by the loss of paychecks received from future employers. The *WLE* for a 30-year-old who suffers death or permanent disability will reflect the possibility that but for this injury there might have been other injuries in the future. But such future injuries would have in some cases led to workers' compensation payments, disability insurance payments, or compensation by tortfeasors or their insurers. Only a fraction of expected future involuntary departures from the labor force would have led to comparable proportional losses of earnings.[9]

As forensic economists we can never hope to forecast all future regime changes, to understand all of the variations of character and personal circumstances that lead to labor force participation that is above or below the average, or to know what fraction of involuntary departures from the labor force would lead to complete loss of income. But we can, in the case of experienced workers, observe something of their past behavior, and judge whether or not the future is likely to be average, in deciding whether or not to use the Markov-based table. If we decide to use such a table, implementation issues still arise, discussed in section IV, after a brief discussion of the sampling base for the tables.

III. The Sample

The population from which the Current Population Survey is drawn to generate BLS worklife expectancy tables is the civilian non-institutional resident population: those persons who are resident in the U.S. and not in the armed forces, in prison, nursing homes, or other long-term care facilities. Conceptually the probabilities of transition recorded in the worklife tables are really conditional probabilities, based on the assumption that one has not moved overseas, become incarcerated, enlisted in the army, or moved into a nursing home during the year (see U.S. Bureau of the Census and U.S. Bureau of Labor Statistics, 2002).

This conceptual restriction is not significant in describing the aggregate behavior of a population, since the number of Americans living abroad, in the military, or in institutions is relatively stable in most years. However there can be considerable turnover so that transition probabilities can be significant

---

[8]See Richards and Abele (1999), Ch.7, for a discussion of alternatives to *WLE* to measure lost earning capacity.

[9]This point was made in Skoog and Ciecka (1998), p. 249.

even though net numbers in each category do not change. For example, while the number of persons in the military is relatively stable over time, there is considerable turnover concentrated among young people.

In practice, the probabilities reported in the BLS worklife tables are for an even more limited population, because transition probabilities were recorded only for matched samples in which the individual lived in the same household at the beginning and end of the year: anyone who changes residence, even though remaining a member of the civilian non-institutional population, was excluded from the sample.

A final observation about the sample is that those recorded as having attained age *x* are not interviewed on their birthdays: on the average, they are interviewed at age x+0.5; to transform the observations to those that would apply at exact age *x*, transition probabilities for those who have attained age x-1 (average age, x-0.5) and x (average age, x + 0.5) are averaged to yield a group whose average age is *x*.

IV. Uses of the Markov Assumption in Worklife Expectancy Calculations

All forensic economists should be aware of these standard results from the actuarial science literature (definitions follow the statement of the results):

- For any positive interest rate, the present value of an annuity certain that extends to the end of life expectancy exceeds the present value of the corresponding life annuity, a_x ; for an interest rate of zero the two are equal; for a negative interest rate the inequality is reversed.
- The same statements are true if “worklife expectancy” is substituted for “life expectancy.”

Here “annuity certain” means a stream of equal periodic payments that continue for a fixed period, not contingent on whether a person is alive or in the labor force. The “corresponding life annuity” means a series of constant payments continuing for as long as the person is alive, whether that exceeds or falls short of the life expectancy. Though the term is not in common use, a “worklife annuity” continues while the person is in the labor force.

The result was clearly stated by Nicholas Bernouilli (1709) in his doctoral thesis at the end of page 31, appeared in English in David Jones (1844, end of paragraph 127), and was proved in the King (1887) textbook at p. 112-113 in 1887. Mathematically elegant proofs were offered by Steffensen, (1918, p. 87 and 1919, p. 277) who proved this relationship that he called “familiar to actuaries.” The result has also been brought to the attention of forensic economists, e.g., in Ben-Zion and Reddall (1985).[10]

If the forensic economist chooses to use an annuity certain rather than a life annuity despite the possible upward bias it gives to the measured loss, the

---

[10]Ben-Zion and Reddall focus on annuities in a divorce setting.

years to which future losses should be attributed must still be settled. Here are some alternatives.

A. Front loading

It may appear straightforward to look up the years of worklife expectancy WLE_x as of the accident date, ${}^{a}e_x^{a}$ if active or ${}^{i}e_x^{a}$ if inactive, and assume that *the next* WLE_x years would have been spent in the labor force. But in fact, on the average, people leave the workforce more than once. Since payments in the distant future are worth less than earlier payments, front loading produces an upper bound to the present value. As this fact has become more widely recognized, the use of front loading has diminished in forensic economic practice.

This shortcut approach also misapprehends the construction of the tables, since all one knows from the tables is that WLE_x represents the *total* number of years in the labor force up to a truncation age of, say, 75 or 80. This is an intrinsic difficulty associated with the tables: one cannot capture in one number the desired detail for the most precise present value calculation. We are led in practice to the alternatives of the next sections.

B. An ad hoc distribution of worklife expectancy

In these approaches, the worklife expectancy is spread over the number of years remaining to a standard retirement age, often 65, 66 or 67. Each year to the retirement age is given a weight. Using age 65, $WLE_x/(65-x)$ produces uniform weights, but other weights, say, proportional to the probability of survival, could be substituted. Assuming that the first 65-WLE_x years are out of the labor force would correspond to *backweighting* and would generally produce a lower bound to the present value estimate.

C. A refinement: weight distribution in personal injury cases

When the plaintiff has survived the period from the injury to the trial, and testifies that he would have participated in the labor force up to the trial date, one may consider allocating the years to trial from WLE_x and choosing the weights as in Section B, but pro-rating only the post-trial worklife expectancy which remains. In more detail, if *y* is the age at the trial date, so that *y-x* represents the years between the accident and the trial for which participation is assumed, then the uniform pro-rata post-trial weights become $\left(WLE_x-\left(y-x\right)\right)/\left(65-y\right)$.[11]

D. The theoretically correct but more computationally intensive method: Decompose

[11]An alternative in these circumstances would be to use $WLE_y/\left(65-y\right)$; however it seems generally agreed that we should take the plaintiff in his circumstances (including life or worklife expectancy, if relevant) at the date of injury, not at the date of trial.

Formulae for the theoretically correct weights implied by the increment decrement model are given in Skoog (2002), together with the recursive structure that makes computation efficient. One needs the underlying transition probabilities to perform this decomposition. As advances are made in forensic economics, computer software will undoubtedly permit this calculation to be done without specific programming effort.

E. Assessment of front loading and uniform loading

Using the theoretically correct decomposition of Section IV-D, Skoog and Ciecka (2004) have explored the biases which come from the front loading of Section IV-A and uniform loading of Section IV-B, in the case where age-earnings profiles are not present. They have produced nomograms (graphs of percentage error against age) at various net discount rates, ranging from 1% to 4%, corresponding to their 24 worklife tables. One may front load or uniform load and visually look up the correction factor. This factor could then be applied directly, or indirectly in conjunction with corresponding and offsetting change of the fringe benefit assumption. The 24 nomograms cannot be simply summarized, except to note that the factors differ markedly depending on whether one began active or inactive.

V. Alternatives to the (First Order) Markov Assumption

A. The use of scenarios

Let us suppose that the economist thinks that available worklife expectancy tables do not describe the prospects for the individual being evaluated because attachment to the labor force seems either unusually high or unusually low. In these cases, although one may not think that the averages given by worklife expectancy are appropriate, there may be no sound statistical basis for an alternative explicit assumption about future worklife.

Macroeconomic forecasters deal routinely with forecasts that depend on contingencies to which they cannot assign probabilities. They then fall back on *scenarios*: alternative coherent stories of what might happen. The scenario describes circumstances that are assumed to hold for the forecast to be valid. For example, one might assume that OPEC will not take action to restrict petroleum supplies.

The Social Security Administration and the U. S. Bureau of the Census use the same approach to forecasting population, labor force and other variables for the future–variables that will depend on the behavior of people, in some cases, not yet born. Scenarios are based on assumptions about birth rates, changes in mortality and other variables that the forecasters are unwilling or unable to project. Business forecasters also specify alternative scenarios as a way of dealing with contingencies for which numerical probabilities cannot be assigned (see Mendell, 1985).

One may treat worklife expectancy in that way. Some forensic economists take a completely agnostic position on age at retirement, listing future losses year by year to some reasonable upper bound age, so that the triers of fact may choose the age they prefer. Guidance in this upper bound age may come from Skoog-Ciecka (2001b) where probability intervals of varying sizes are tabulated.

B. Overwriting transition probabilities

For an active individual with a long history of activity–a strong labor force attachment–it is natural to set ${}^{a}p^{i}_{x+j}$, j=1,2,… to 0 in the equations of motion (11) of the Markov process–at least up to some retirement age. This amounts to acknowledging death as the only reason that the subject would fail to participate in the future, prior to some conventional target retirement age. Holders of sedentary jobs without mandatory retirement, accompanied by good health insurance, a high income and a high psychic income (e.g., college professors or forensic economists) might be candidates for this procedure. This approach is similar to the LPE model[12] (Brookshire and Cobb, 1983) in which "P" is set to 1 (and E set to 1 as has been implicit throughout).

C. Higher order Markov process

Skoog (2003) develops the theory and formulae that generalize the basic increment-decrement model by permitting the transition probabilities to depend not only on last year's state, but on the states two, three, and generally, p years ago.[13] The difficulty in implementing such a model will be to identify large longitudinal data sets that track individuals over more than one year. There is modest hope for such information from The Bureau of the Census' Survey of Income and Program Participation (SIPP), but this survey was not designed to measure labor force participation–individuals answer whether they have been employed and whether they have been unemployed, and they may answer "yes" to both questions. The University of Michigan's Panel Study on Income Dynamics (PSID) data represents another potential source, but it does not have nearly the sample size of the CPS.

D. Microeconometric model of individual behavior

Peracchi and Welch (1994), Ciecka and Donley (1996) and Millimet, et al., (2003) have developed parametric econometric models of transitions for labor force participation. The latter authors follow Flinn and Heckman (1981) and include three separate states for those who survive: employment, unemployment, and inactivity. Their model estimates transition probabilities among these three states drawing on the same Current Population Survey data as has been used by the BLS, the Ciecka, et al., updates (1995; 1997; 1999-2000)

---

[12]L is the probability of living (survival), P the probability of participation in the labor force conditional on being alive, and E the probability of employment conditional on labor force participation.
[13]The first-order Markov process was described above as one in which the transition probabilities depend only on the current state. In a second-order Markov process, the transition probabilities governing the state at age x+1 depend on the current and past states, two periods earlier than x+1. In a third-order process the x to x+1 transition probabilities depend on the states at x, x-1 and x-2, and so on.

and Skoog and Ciecka (2001), making the same Markov assumption that transition probabilities from age x to $x + 1$ depend only on the state at age x. The significant innovation compared to work of the Bureau of Labor Statistics is to use the information provided in the survey much more parsimoniously, so that many more causal variables may be accounted for. Millimet, et al., use estimating equations that include as explanatory variables sex, age, race, marital state, number of children under six, number of children under 18, education, and occupation. Education is classified into three categories (less than high school degree, high school graduate to some college, and college degree or above); occupation is classified into four categories (managerial and professional; technical, sales and administrative; service; and operators and laborers). A drawback to this approach is that some of the explanatory variables, for example marital status and occupation, may not remain fixed in the future.

This fine grid of information would allow one to apply the Markov model to a given individual with more confidence, though the computational demands would be formidable. One would need to generate a future in which the number of children in each age group changed appropriately, year-by-year, then calculate a tailored worklife expectancy appropriate to the individual using the recursive formula described in Section II. Unfortunately the paper of Millimet, et al., does not give the estimating equations that would allow construction of this kind of tailored worklife expectancy; but the method has now been developed, and computing power is cheap. It is possible that in the near future such individualized calculations will be feasible, and may combine parametric tailoring with the Skoog decomposition of total worklife expectancy into parts of each additional year worked.

There will be the usual tension between parametric and non-parametric models–where a few parameters capture the variation in the independent variable well, the estimation is more efficient (produces lower sampling errors), but where the parameterization breaks down, biases are introduced. Apart from the tables produced by Millimet and his colleagues, all existing increment-decrement worklife expectancy tables (the work of Smith, Ciecka, et al., Richards and Abele, and Skoog and Ciecka cited above) have been non-parametric in spirit.

VI. Conclusion

We have seen that the use of a worklife expectancy table unavoidably imposes straightjackets on the projections of future labor force participation of any individual to whom it is applied: that person's future must be projected as though average for the entire cohort even though there may be good reason for believing that the individual is not at all average. Any forensic economist using the tables should be aware of the underlying assumptions, and have some familiarity with the mathematical and statistical basis for their construction.

Alternatives exist for avoiding the use of the tables where they do not seem appropriate. Nevertheless, it is routine to use the statistical assumption that a particular plaintiff is "average" with respect to other aspects of the economic loss due to a claimed tort, for lack of more specific information, and the assumption is perhaps as appropriate here as in other aspects of forensic work.

References

Arias, E., *United States life tables, 2001, National vital statistics reports*, 52(14), Hyattsville, MD: National Center for Health Statistics, 2004.

Becker, G., "A theory of the allocation of time," *Economic Journal*, 1965, 75(299), 493-517.

Ben-Zion, Barry and Ronald L. Reddell, "Life Expectancy and Actuarial Present Value," in *Research in Law and Economics*, JAI Press, 1965, 7, 161-171.

Bernouilli, M. Nicholaus, *The Use of the Art of Conjecturing in Law, Doctoral Dissertation*, 1709, Basel, translated from Latin under the direction of Richard J. Pulskamp, 1999.

Brookshire, M. L. and W. E. Cobb, "The Life-Participation-Employment approach to worklife expectancy in personal injury and wrongful death cases," *For the Defense*, July 1983, 20-26.

Ciecka, J. and T. Donley, et al., "A Logit Model of Labor Force Participation," *Journal of Forensic Economics*, 1996, 9(3), 261-282.

_______, _______, _______, "A Markov Process model of work-life expectancies based on labor market activity in 1992-93," *Journal of Legal Economics*, 1995, 5(3), 17-41.

_______, _______, _______, "A Markov Process model of work-life expectancies based on labor market activity in 1994-95," *Journal of Legal Economics*, 1997, 7(1), 2-25.

_______, _______, _______, "A Markov Process model of work-life expectancies based on labor market activity in 1997-98," *Journal of Legal Economics*, Winter 1999-2000, 9(3), 33-68.

_______, S. Epstein, et al., "Updated estimates of work-life expectancies based upon the increment-decrement model," Journal of Legal Economics, 1995, 5(1), 1-33.

Cox, D. R. and H. D. Miller, *The Theory of Stochastic Processes*, London: Methuen & Co. Ltd., 1965.

Ehrenberg, R. G. and R. S. Smith, *Modern Labor Economics*, 8th Edition, New York: HarperCollins College Publishers, 2002.

Feller, W., *An introduction to probability theory and its applications*, New York: John Wiley and Sons, Inc., 1968.

Flinn, C., and Heckman, J., "Models for Analysis of Labor Force Dynamics," in R. Basmann and G. Rhodes, eds., *Advances in Econometrics*, JAI Press, 1981.

Jones, David, *On the Value of Annuities and Reversionary Payments, With Numerous Tables*, The Society for the Diffusion of Useful Knowledge, London, Robert Baldwin, 1844.

Kaufman, B. E. and J. L. Hotchkiss, *The Economics of Labor Markets*, Fort Worth: The Dryden Press, a Division of Harcourt College Publishers, 2000.

King, George, *Institute of actuaries text book of the principles of interest, life annuities, and assurances, and their practical application, part II, life contingencies* (*including life annuities and assurances*), London: Charles and Edwin Layton, 1887.

Mendell, J. S., *Nonextrapolative methods in business forecasting, Scenarios, vision, and issues management*, Westport, Conn.: Quorum Books, 1985.

Millimet, D. L., M. Nieswiadomy, et al., "Estimating worklife expectancy: An econometric approach," *Journal of Econometrics*, 2003, 113(1), 83-113.

Peracchi, F. and F. Welch, "Trends in Labor Force Transitions of Older Men and Women," *Journal of Labor Economics*, 1994, 12(2), 210-242.

Richards, H., "Worklife expectancies: Increment-decrement less accurate than conventional," *Journal of Forensic Economics*, 2000, 13(3), 271-289.

Richards, H. and J. R. Abele, *Life and worklife expectancies*, Tucson, AZ: Lawyers and Judges Publishing Company, Inc., 1999.

Skoog, G., "Worklife expectancy: Theoretical results," De Paul University manuscript. Paper presented at Allied Social Sciences/AEA Annual Meetings. Atlanta, Georgia, January 4-6, 2002.

_______,"Worklife expectancy for higher order Markov models," De Paul University manuscript. Paper presented at Allied Social Sciences/AEA Annual Meetings, Washington, DC, January 3-5, 2003.

_______, and James Ciecka, "Probability Mass Functions for Labor Market Activity Induced by the Markov (Increment-Decrement) Model of Labor Force Activity," *Economics Letters*, 2002, 77, 425-431.

_______, and _______,"The Markov (Increment-Decrement) Model of Labor Force Activity: New Results Beyond Worklife Expectancies," *Journal of Legal Economics*, 2001a, 11(1), 1-21.

_______, and _______, "The Markov (Increment-Decrement) Model of Labor Force Activity: Extended Tables of Central Tendency, Variation, and Probability Intervals," *Journal of Legal Economics*, 2001b, 11(1), 23-87.

_______, and _______, "Reconsidering and Extending the Conventional/Demographic and LPE Models: The LPd and LPi Restricted Markov Models," *Journal of Forensic Economics*, 2004, 17(1), 47-94.

_______, and _______, "The Allocation of Worklife Expectancy and the Analysis of Front and Uniform Loading," paper presented at the Fall, 2004 American Academy of Economic and Financial Experts meetings, Nashville, TN, 2004.

Smith, S. J., "New worklife estimates reflect changing profile of labor force," *Monthly Labor Review*, 1982a, 105(3), 15-20.

_______, *Table of working life: The increment-decrement model*, Washington, DC: U.S. Department of Labor, Bureau of Labor Statistics, 1982b.

_______, *Worklife estimates: The effects of race and education*, Washington, DC: U.S. Department of Labor, Bureau of Labor Statistics, 1986.

Steffensen, J.F., "On certain inequalities between mean values, and their application to actuarial problems," *Skandinavisk Aktuarietidskrift*, 1918, 1, 82-97.

_______, "On certain inequalities and methods of approximation," *Journal of the Institute of Actuaries*, 1919, 51, 274-297.

U.S. Bureau of the Census and U.S. Bureau of Labor Statistics, Washington, DC: *Design and Methodology*, 2002.