## Worklife Expectancy via Competing Risks/Multiple Decrement Theory with an Application to Railroad Workers

Gary R. Skoog and James E. Ciecka\*

### I. Introduction

A key concept in economics, and arguably the key concept in forensic economics, worklife expectancy, has been treated by actuaries, demographers and forensic economists with different models. The Markov or multiple increment/multiple decrement model has been employed by all three groups, while multiple decrement theory (also known as competing risks in biometrics) represented an earlier approach and a special case, in which transitions into the measured state are disallowed. As discussed in Skoog-Ciecka (2004), in some cases, e.g., railroad worker worklife expectancy Skoog-Ciecka (1998 and 2006), hybrids of these two approaches may prove fruitful, given appropriate but only occasionally available data. In this paper we develop and extend these ideas to allow a look at worklife expectancy in occupations where actuarial data provides longitudinal records of transitions. Beyond worklife expectancy, we also develop probability mass functions (pmf's) which enable us to calculate any other distributional characteristic of time devoted to a specific occupation. We note that occupation-specific worklives provide especially useful information when money earnings and fringe benefits vary by occupation and when people change occupations throughout their worklives. We make an extended application utilizing data for railroad workers and show much lower worklives than previously calculated by others and ourselves.

We have in mind four reasons persons engaged in an occupation exit the occupation: death, disability, retirement, and withdrawal (to another occupation, or out of the labor force). Ideally, we would have such data by year on individuals working in the railroad or other crafts in a region covered by a multiemployer pension plan, e.g., carpenters, ironworkers or laborers in a metropolitan area. We would observe a first year in which contributions are made on their behalf into a pension fund. We would follow them over time and record whether they remain in the occupation and region the next year, or whether they have transitioned from active to inactive due to one of the four causes above. Likewise, when a person who was disabled, retired or withdrawn in the previous period becomes active, we observe a record, whereas when such a person remains inactive for a second period, we observe that event by the non-existence of a record. In this way, an increment-decrement model may be constructed from event-history data that an actuary would maintain. However, in practice, we often do not have access to such detailed event-history data; rather

<sup>\*</sup>Respectively, DePaul University and Legal Econometrics, Inc., Glenview, IL; Economics Department, DePaul University, Chicago, IL.

we might have probabilities or rates of transition recorded by actuaries or deemed reasonable by actuaries. Our goal is to combine such actuarial inputs or data with proper statistical theory to estimate worklife expectancy of workers in an occupation.

Such worklife expectancy will differ conceptually from the measure we and others have traditionally used, since we are calculating years in both the occupation and the region. We therefore caution against indiscriminant use of worklives for younger non-railroad craft workers (since railroad workers are tracked nationally) who might move from the region but remain within the occupation. For older workers, who have established roots (human capital) in the region and/or occupation, including job contacts, reputation, possibly seniority and social connections, such withdrawal into another region but the same occupation will be relatively rare, and these tables will be appropriate. Since some multiplayer pension funds have reciprocity rights, one can sometimes gauge the extent of regional transfers. We provide here multiple decrement estimates for railroad workers because we can compare the results with earlier models, the actuarial assumptions are publicly available, they are reliably based on experience from a large number of employee records, and workers are tracked across regions.

### II. Multiple Decrement/Competing Risk Theory and the Markov Model

We have differentiated between event data and aggregated data. In the former, we have observations on individuals' year-to-year transitions, which we term micro-data or event-history data. Alternatively, we might have macrodata, estimates provided to us by actuaries derived from micro-data providing us with either probabilities of transitions or rates of transition, closely related to these underlying probabilities.

We need extensive notation. For the Markov model, such notation is in place. The innovation of Skoog-Ciecka (1998, 2002, and 2006) and this paper is to use railroad actuarial data to refine and restrict estimates of worklife. Fortunately the actuarial science literature provides notation and results facilitating this exercise; we follow Bowers, et al., (1987) and Jordan (1991).

As usual, we compute transition probabilities between the active and inactive states. Let  ${}^{a}p_{x}^{a}$  be the probability that a worker who is active (logs more than a non-trivial number of hours, typically 250) in the first period (year) remains active in the next year. Then  ${}^{a}p_{x}^{i}$  is the probability such an initially active worker goes inactive; the inactive state could represent a disability status, a retirement status, or a withdrawal status. That withdrawal status may in turn represent movement into another region or a movement out of the labor market altogether, for a reason other than those already mentioned. Death is another withdrawal from activity which is treated separately. Thus, many sources of decrement (transition out of the measured state) are now explicitly counted, and our task will be to build up an estimate of  ${}^{a}p_{x}^{i}$  from such data, a somewhat different situation from that encountered in usual worklife construction from Current Population Survey data. For each age x we have, of course, that  ${}^{a}p_{x}^{a} + {}^{a}p_{x}^{i} + p_{x}^{d} = 1$  and that  ${}^{i}p_{x}^{a} + {}^{i}p_{x}^{i} + p_{x}^{d} = 1$ . At some advanced age x,

244

death occurs with probability 1:  ${}^{a}p_{x}^{a} = {}^{a}p_{x}^{i} = {}^{i}p_{x}^{a} = {}^{i}p_{x}^{i} = 0$ . Similarly, to estimate a Markov model (or to estimate a special case thereof, such as what Skoog-Ciecka, 2004, termed the *LPd* model) and unlike the multiple decrement model, we need estimates of transition probabilities from the inactive state. These might be estimated given micro-data; alternatively, they may be provided by actuaries either as estimates or, in the multiple decrement setting, as assumptions that certain probabilities or rates are zero. Finally, regarding mortality, in the Markov setting, it is commonly assumed that  ${}^{a}p_{x}^{d} = {}^{i}p_{x}^{d} \equiv p_{x}^{d}$ . This is implied by the assertion that transition to death is independent of the initial living state, active or inactive. This assumption of independence is maintained more generally in the multiple decrement literature.

In the multiple decrement setting, let  $q_x^{(j)}$  be the probability of decrement (movement out of the counted state, *a*, years of activity in the usual Markov model above, or years in the occupation specific railroad activity, below) due to cause *j* for a person age *x* between ages *x* and *x*+1. The causes *j* are:

(1) 
$$q_x^{(1)}$$
 probability of death  
 $q_x^{(2)}$  probability of disability  
 $q_x^{(3)}$  probability of retirement, and  
 $q_x^{(4)}$  probability of withdrawal.

Since these are mutually exclusive probabilities, the probabilities are additive. We let  $\tau$  indicate "all causes" for such transitions, so that

(2a) 
$$q_x^{(\tau)} = q_x^{(1)} + q_x^{(2)} + q_x^{(3)} + q_x^{(4)}$$

and

(2b) 
$$p_x^{(\tau)} = 1 - q_x^{(\tau)}$$

where  $p_x^{(\tau)}$  gives the probability of remaining in the initial state throughout the interval.

Once the  $q_x^{(j)}$  are in hand, one may define decrements into state j as  $d_x^{(j)} = l_x^{(\tau)} q_x^{(j)}$  from an initial number of persons in the initially active state,  $l_x^{(\tau)}$ , and proceed to construct a multiple decrement table in the same way that a single-decrement life table is constructed.

There is a fair amount of theory beneath these symbols. Let T be a continuous random variable, denoting additional time of exit from the active state or occupation. The equations above implicitly set T's realization t to 1, since this is the first time period after initial age x which we observe. Consequently  $_{I}q_{x}^{(j)} = _{t}q_{x}^{(j)}$ , for all j. T, a survival time, possessing a continuous instantaneous force  $\mu_{x}^{(j)}$  of transition, and the state J, j=1,...,m are two random variables described by a joint pmf function  $f_{T,J}(t,j)$ . This function is continuous in the first argument and discrete in the second argument, since the number of states is discrete. Usual definitions and relations hold:

(3) 
$$f_{T,J}(t,j) = \Pr\{(t < T \le t + dt) \cap (J = j)\},\$$

the definition of a continuous-discrete density;

(4) 
$$\sum_{j=1}^{J=m} \int_{a}^{b} f_{T,J}(t,j) dt = Pr\{a < T \le b\},$$

the probability of decrement due to all causes between times *a* and *b*;

(5) 
$$Pr\{(0 < T \le t) \cap (J = j)\} = \int_{0}^{t} f_{T,J}(s,j) ds$$

the probability of transition before *t* into (due to cause) state *j*;

(6) 
$${}_{t}q_{x}^{(j)} = \int_{0}^{t} f_{T,J}(s,j)ds, t \ge 0, j = 1,...,m$$

generalizing (1) and re-expressing (5);

(7) 
$$f_T(j) = \int_0^\infty f_{T,J}(s,j) ds = {}_\infty q_x^{(j)}, j = 1,...,m$$

the marginal distribution for J, the probability of decrement due to cause j at any time in the future.

Now listed among the actuarial assumptions in any pension valuation will be specifications of mortality, disability, retirement, and withdrawal decrements. One must determine whether these are *net probabilities*, sometimes called *independent rates of decrements*, or equivalently *absolute rates of decrements*  $q'_{x}^{(1)}$ ,  $q'_{x}^{(2)}$ ,  $q'_{x}^{(3)}$ ,  $q'_{x}^{(4)}$ , or whether they are instead the  $q^{(j)}_{x}$  probabilities above. These net probabilities, also known as *complementary associated single decrement probabilities* are referred to with primes, refer to remaining in the state, and are given by

(8) 
$${}_{t}p'_{x}^{(j)} = exp\{-\int_{0}^{t}\mu_{x}^{(j)}(s)ds\} = 1 - {}_{t}q'_{x}^{(j)}.$$

The  $q_x^{(j)}$  necessarily exceed the corresponding  $q_x^{(j)} = 1 - p_x^{(j)}$ , as may be seen in the next equation, (9).

We have said that one problem involves determining when actuarial assumptions or determinations appearing in pension plan valuations refer to probabilities as opposed to net probabilities. Another possible problem is that the plan assumptions, which are often called *rates*, refer to central rates of decrement, denoted by an actuarial symbol such as  $m_x^{(j)}$ , an estimator for  $\mu_x^{(j)}$ , above. The relation between a central rate of decrement and the probability generally is  $m_x = q_x/_{1-.5q_x}$  when there is only one source of decrement. We need to be able to convert these single decrement  $q'_x^{(j)}$  net probabilities or  $m_x^{(j)}$ rates appearing in actuarial reports into the corresponding probabilities of multiple decrement  $q_x^{(j)}$ . This need arises when we have only actuarial macrodata. If we had access to the underlying actuarial event data, we could directly estimate the decrements. Alternatively, we could group the retirement, disability and withdrawal states into an inactive state, and proceed as with the ordinary analysis of CPS data, and use the mortality experience of the group (if its size is large enough) or impose the mortality experience from an external source. (U.S. Life Tables or an annuitants' mortality table)

246

As a special case, but one we have noted in practice, suppose we have  $q_x^{(2)}$  the probability of disability,  $q_x^{(3)}$  the probability of retirement, and  $q_x^{(4)}$  the probability of withdrawal, while  $q_x'^{(1)}$  comes from an associated mortality table such as the Group Annuitants' Mortality 1983, which provides an absolute rate of decrement. We need to be able to convert this latter net rate (the rate at which the initial population would diminish, if mortality were the only factor causing decrements) to the corresponding  $q_x^{(1)}$  probability. Starting with  $l_x$  persons alive at x, we know that  $d_x^{(2)} = q_x^{(2)}l_x$  will be disabled at x+1,  $d_x^{(3)} = q_x^{(3)}l_x$  will retire and  $d_x^{(4)} = q_x^{(4)}l_x$  will withdraw at x+1, by definition of the  $q_x^{(j)}$  as probabilities. Now assuming that these decrements proceed linearly between x and x+1, at age  $x + \frac{1}{2}$  there will be  $l_x - .5(d_x^{(2)} + d_x^{(3)} + d_x^{(4)})$  remaining, and it is this number, and not  $l_x$ , which is on average exposed to death. Thus  $q_x'^{(1)}$  times this number or  $q_x'^{(1)}$  { $1-.5(q_x^{(2)} + q_x^{(3)} + q_x^{(4)})$ }  $l_x = q_x^{(1)}l_x$  will die. Consequently we have, equating coefficients, and dividing out  $l_x$ ,

(9) 
$$q_x^{(1)} = q_x^{\prime(1)} \{ 1 - .5(q_x^{(2)} + q_x^{(3)} + q_x^{(4)}) \}$$

as a simple approximation (to second order). Other approximations are possible, as are solutions based on the more complicated and exact expressions. (Bowers, 1987, p. 321, 323)

(9a) 
$$1 - q_x^{\tau} = \prod_{i=1}^{i=4} (1 - q_x'^{(i)})$$

and

(9b) 
$$1 - q'_{x}^{(j)} = (1 - q_{x}^{(\tau)})^{q_{x}^{(j)}/q_{x}^{(\tau)}}$$

The left hand side of equation (9) shows that the probability of decrement due to death when all four sources of decrement are present (i.e., are "competing" with each other) depends on the strength of the competition-higher  $q_x^{(2)}$ ,  $q_x^{(3)}$  or  $q_x^{(4)}$  probabilities make it less likely that source 1 is the reason for decrement.  $q_x^{(1)}$  is only equal to the commonly reported  $q_x'^{(1)}$  when it is the only cause of decrement, the implicit assumption in ordinary mortality tables. Indeed, a literature parallel to the multiple decrement model, as this subject is called in actuarial science, was developed in biostatistics and statistics generally, where it is called *competing risk theory*. Seal (1977) offers historical context, while modern books on failure time data (Kalbfleisch and Prentice, 2002, Chapter 8) provide textbook treatments.

For ensic economists have some experience with the need to derive the probabilities of interest in the usual construction of the Markov model, where we obtain death probabilities from one source (U.S. Life Tables) and (conditional on survival) transition probabilities from another source, the CPS. A life table gives us the ordinary mortality probability, referred to above as a net probability,  $q'_x{}^{(1)} = q_x$ , where the left-hand side is in the new notation and the right-hand side reflects the usual single decrement life table notation. This last expression is equivalent to  $p'_x{}^{(1)} = 1 - q_x$ . Consequently, it is standard in the Markov model construction to form the relevant transition probability  ${}^a p_x^a = (1 - q_x)^A p_x^A = p'_x{}^{(1)A} p_x^A$  and  ${}^a p_x^i = (1 - q_x)^A p_x^I = p'_x{}^{(1)A} p_x^{(2)}$  out of a net probability and a conditional probability, which here is also another net probability.

Returning to (9), the objective was to calculate  $q_x^{(1)}$  given  $q_x^{\prime(1)}$ ,  $q_x^{(2)}$ ,  $q_x^{(3)}$ , and  $q_x^{(4)}$ . More generally, given only the rates  $q_x^{\prime(1)}$ ,  $q_x^{\prime(2)}$ ,  $q_x^{\prime(3)}$ , and  $q_x^{\prime(4)}$ , then by symmetry

(10) 
$$\begin{array}{l} q_{x}^{(1)} = q_{x}^{\prime (1)} [1 - .5(q_{x}^{(2)} + q_{x}^{(3)} + q_{x}^{(4)})] \\ q_{x}^{(2)} = q_{x}^{\prime (2)} [1 - .5(q_{x}^{(1)} + q_{x}^{(3)} + q_{x}^{(4)})] \\ q_{x}^{(3)} = q_{x}^{\prime (3)} [1 - .5(q_{x}^{(1)} + q_{x}^{(2)} + q_{x}^{(4)})] \\ q_{x}^{(4)} = q_{x}^{\prime (4)} [1 - .5(q_{x}^{(1)} + q_{x}^{(2)} + q_{x}^{(3)})] . \end{array}$$

Re-arranging (10) by moving the second, third and fourth terms from the righthand side to the left-hand side and solving the matrix equation  $A_{q'_x}q_x = q'_x$  for  $q_x = A_{q'_x}^{-1}q'_x$  yields the corresponding probabilities contained in the vector  $q_x$ , where

$$(11) A_{q'_{x}} = \begin{bmatrix} 1 & .5q'_{x}^{(1)} & .5q'_{x}^{(1)} & .5q'_{x}^{(1)} \\ .5q'_{x}^{(2)} & 1 & .5q'_{x}^{(2)} & .5q'_{x}^{(2)} \\ .5q'_{x}^{(3)} & .5q'_{x}^{(3)} & 1 & .5q'_{x}^{(3)} \\ .5q'_{x}^{(4)} & .5q'_{x}^{(4)} & .5q'_{x}^{(4)} & 1 \end{bmatrix}, \ q_{x} = \begin{bmatrix} q_{x}^{(1)} \\ q_{x}^{(2)} \\ q_{x}^{(3)} \\ q_{x}^{(4)} \end{bmatrix}, \text{ and } q'_{x} = \begin{bmatrix} q'_{x}^{(1)} \\ q'_{x}^{(2)} \\ q'_{x}^{(2)} \\ q'_{x}^{(3)} \\ q'_{x}^{(4)} \end{bmatrix}.$$

The Appendix contains formulae for the calculation of probabilities given that data come in the form of three probability decrements and one net rate (as in (9)), two probability decrements and two net rates, one probability decrement and three net rates, and four net rates (i.e., the solution for the vector  $q_x$  in (11)). Expression (A7a) in the appendix effectively inverts the matrix  $A_{q'_x}$ .

### **III. Application to Railroad Workers**

As an example of the competing risks/multiple decrement theory, we utilize data contained in *The Twenty-Third Actuarial Valuation* of the US Railroad Retirement Board. We expand on our previous notation to incorporate the importance of service years on disability, retirement, and withdrawal probabilities in the following manner:

x	denotes exact age, here $x = 17, \dots, 75$ ;
ω	denotes the youngest age for which the probability of being active
	in the railroad industry is zero, here $\omega = 95$ ;
s	denotes years of railroad service, here $s = 0,, x - 17$ ;
$q_{x}^{\prime(1)}$	denotes the mortality rate between age $x$ and $x+1$ ;
$q_{x,s}^{(2)}$	denotes the probability of a railroad disability retirement (the
	RRB Actuary's term) between <i>x</i> and <i>x</i> +1 given <i>s</i> years of service;
$q_{x,s}^{(3)}$	denotes the probability of a railroad age retirement between $x$
	and <i>x</i> +1 given <i>s</i> years of service;
$q_{x,s}^{(4)}$	denotes the probability of withdrawal from railroad work be-
	tween <i>x</i> and <i>x</i> +1 given <i>s</i> years of service;

- $\begin{array}{l} q_x^{(1)} = q_x'^{(1)} [1 .5(q_{x,s}^{(2)} + q_{x,s}^{(3)} + q_{x,s}^{(4)})], \quad q_x^{(1)} \quad \text{denotes mortality probability,} \\ q_x'^{(1)} \text{ measures the net rate of mortality, and we use (9) above to} \\ & \text{transform } q_x'^{(1)} \text{ into } q_x^{(1)}; \\ WLE_{x,s}^{CR} \quad \text{denotes competing risks railroad worklife expectancy for an indi-} \end{array}$
- $WLE_{x,s}^{CR}$  denotes competing risks railroad worklife expectancy for an individual at age x with s years of railroad service under the assumption of mortality, disability, age retirement, and withdrawal as the competing risks.

As the notation suggests, in *The Twenty-Third Actuarial Valuation* the mortality probability  $q_x^{\prime(1)}$  (and therefore the transformed  $q_x^{(1)}$ ) is not a function of years of service; disability retirement probability  $q_{x,s}^{(2)}$  depends on both age and years of service; age retirement probability  $q_{x,s}^{(3)}$  is zero prior to age 60 and it depends upon both x and s at age 60 and beyond; and the withdrawal probability  $q_{x,s}^{(4)}$  is a function of both age and years of service.<sup>1</sup> Disability, retirement, and withdrawals are reported as probabilities but mortality is given as a net rate, requiring a conversion to a probability as given in (9).

Consider a person who is in the railroad industry at age x and has s years of service. Then the probability of that individual remaining in the railroad industry at age x+1 is

(12) 
$${}_{1}p_{x,s} = 1 - (q_x^{(1)} + q_{x,s}^{(2)} + q_{x,s}^{(3)} + q_{x,s}^{(4)})$$

which is the same as (2b) with the time of exit from the active state being one year.

The probability of continuing as a railroad worker is defined recursively by

(13) 
$$_{i+1}p_{x,s} = {}_{i}p_{x,s} [1 - (q_{x+i}^{(1)} + q_{x+i,s+i}^{(2)} + q_{x+i,s+i}^{(3)} + q_{x+i,s+i}^{(4)})]$$

where  $i = 1, ..., \omega - x - 1$  and  $_{\omega-x} p_{x,s} = 0$ . We note that (13) is the discrete counterpart of one minus the probability in (4) where, using the notation in (4), a = 0, b = i + 1, and m = 4 sources of decrement.

The worklife expectancy in railroad service at exact age x for an individual with s years of railroad service is

(14) 
$$WLE_{x,s}^{CR} = .5(1 + {}_{1}p_{x,s}) + .5({}_{1}p_{x,s} + {}_{2}p_{x,s}) + ... + .5({}_{\omega-x-1}p_{x,s} + {}_{\omega-x}p_{x,s})$$
  
=  $.5 + \sum_{i=1}^{\omega-x-1} {}_{i}p_{x,s}$ 

where each term on the right-hand side of (14) reflects the usual averaging of beginning and ending period probabilities (i.e., assuming mid-period transitions) typically used in expectancy calculations, such as average years of life calculations (Jordan, 1991, p. 173). The term New CR Expectancy in Table 1 refers to railroad worklife expectancies computed with formulae (12)-(14).<sup>2</sup>

<sup>&</sup>lt;sup>1</sup>The mortality-rate table in the *Twenty-Third Actuarial Valuation* is Table S-4, disability retirements are from Table S-11, age retirements are found in Table S-10, and withdrawals are Table S-12.

<sup>&</sup>lt;sup>2</sup>In Table 1 we illustrate worklife expectancies commencing at age 20 but in five-year age steps with service years also in five-year steps. We have produced more extensive calculations. One set is for the Association of American Railroads. Another set of additional tables and some extended calculations in a different format may be found at authors' web sites: www.legaleconometrics.com, http://fac.comtech.depaul.edu/gskoog/.and http://fac.comtech.depaul.edu/jciecka/

	Service	New CR	Standard	Old AAR-Type	ID	Standard
Age	Years	Expectancy	Deviation	Expectancy	Expectancy	Deviation
20	0	15.08	15.50	35.59	35.02	6.60
25	0	15.40	13.80	31.33	31.08	6.02
25	<b>5</b>	20.44	13.36	30.91	31.08	6.02
30	0	14.21	12.03	27.05	28.81	6.58
30	5	18.51	11.34	26.62	26.71	5.52
30	10	20.20	10.74	26.20	26.71	5.52
35	0	13.68	11.46	24.92	24.49	6.53
35	<b>5</b>	16.55	9.26	22.31	24.04	5.92
35	10	17.72	8.67	21.87	22.24	5.04
35	15	18.43	8.22	21.45	22.24	5.04
40	0	12.18	9.81	20.84	20.35	6.10
40	<b>5</b>	15.52	8.82	20.31	20.06	6.10
40	10	14.83	6.89	17.66	19.61	5.45
40	15	15.13	6.60	17.21	17.80	4.53
40	20	15.21	6.57	16.84	17.80	4.53
45	0	10.10	7.99	16.70	16.07	5.56
45	<b>5</b>	13.38	7.06	16.44	16.07	5.56
45	10	13.75	6.56	15.90	16.00	5.45
45	15	11.99	4.87	13.15	14.92	4.60
45	20	11.77	5.06	12.75	13.49	3.86
45	25	11.88	4.95	12.94	13.49	3.86
50	0	7.44	5.88	12.55	11.88	4.97
50	<b>5</b>	10.39	5.60	12.54	11.88	4.97
50	10	10.80	5.26	12.26	11.88	4.97
50	15	10.61	5.03	11.69	11.80	4.82
50	20	8.34	3.50	8.83	10.11	3.59
50	25	8.41	3.43	8.97	9.25	3.11
50	30	8.36	3.46	9.15	9.25	3.11
55	0	4.89	4.05	8.61	8.09	4.20
55	<b>5</b>	6.64	4.28	8.61	8.09	4.20
55	10	7.65	4.06	8.60	8.09	4.20
55	15	7.62	3.86	8.28	8.09	4.20
55	20	7.20	3.81	7.71	7.96	3.99
55	25	4.81	2.01	5.01	5.27	2.15
55	30	4.79	2.04	5.15	5.27	2.15
55	35	4.79	2.04	5.15	5.27	2.15
60	0	3.67	2.91	5.97	4.75	3.19
60	5	3.85	2.87	4 75	4 83	3 36

Table 1 Worklife Expectancies of Railroad Workers Utilizing Competing Risk Theory and the Increment-Decrement Model

(continued)

	Service	New CR	Standard	Old AAR-Type	ID	Standard
Age	Years	Expectancy	Deviation	Expectancy	Expectancy	Deviation
60	10	4.34	2.97	4.75	4.83	3.36
60	15	4.44	2.97	4.74	4.83	3.36
60	20	4.18	2.92	4.41	4.83	3.36
60	25	4.09	2.65	4.28	4.67	3.09
60	30	1.35	1.53	1.44	1.27	1.51
60	35	1.35	1.53	1.44	1.27	1.51
60	40	1.35	1.53	1.44	1.27	1.51
65	0	4.39	3.31	7.30	4.03	3.38
65	5	2.09	2.44	2.63	2.41	2.82
65	10	2.39	2.68	2.63	2.41	2.82
65	15	2.46	2.74	2.63	2.41	2.82
65	20	2.47	2.75	2.62	2.41	2.82
65	25	2.49	2.73	2.62	2.41	2.82
65	30	2.09	2.34	2.27	2.09	2.46
65	35	2.09	2.34	2.27	2.09	2.46
65	40	2.09	2.34	2.27	2.09	2.46
65	45	2.09	2.34	2.27	2.09	2.46
70	0	4.18	3.12	6.84	3.38	2.98
70	5	2.52	2.52	3.12	3.16	2.97
70	10	2.85	2.72	3.12	3.16	2.97
70	15	2.93	2.76	3.12	3.16	2.97
70	20	2.96	2.78	3.12	3.16	2.97
70	25	2.98	2.76	3.12	3.16	2.97
70	30	2.87	2.69	3.12	3.16	2.97
70	35	2.87	2.69	3.12	3.16	2.97
70	40	2.87	2.69	3.12	3.16	2.97
70	45	2.87	2.69	3.12	3.16	2.97
70	50	2.87	2.69	3.12	3.16	2.97
75	0	3.75	2.87	6.01	2.39	2.04
75	<b>5</b>	2.31	2.27	2.83	2.90	2.73
75	10	2.60	2.44	2.83	2.90	2.73
75	15	2.66	2.48	2.83	2.90	2.73
75	20	2.69	2.50	2.83	2.90	2.73
75	25	2.71	2.49	2.83	2.90	2.73
75	30	2.61	2.43	2.83	2.90	2.73
75	35	2.61	2.43	2.83	2.90	2.73
75	40	2.61	2.43	2.83	2.90	2.73
75	45	2.61	2.43	2.83	2.90	2.73
75	50	2.61	2.43	2.83	2.90	2.73
75	55	2.61	2.43	2.83	2.90	2.73

# Table 1 (continued) Worklife Expectancies of Railroad Workers Utilizing Competing Risk Theory and the Increment-Decrement Model

Formulae (12) and (13) induce probability mass functions (pmf's) on years of railroad activity. Here we think of additional years of railroad work  $YA_{x,s}^{CR}$  as a random variable with pmf  $p_{YA}^{CR}(x,s,y)$ , which denotes the probability that a railroad worker age x with s service years will accumulate y additional years of railroad service. This pmf at age x and s service years consists of the boundary condition and a main recursion in (15).

(15) Boundary Condition:  

$$p_{YA}^{CR}(x,s,.5) = 1 - {}_{1}p_{x,s}$$
  
Main Recursion:  
 $p_{YA}^{CR}(x,s,y) = {}_{y-.5}p_{x,s} - {}_{y+.5}p_{x,s}$   $y = 1.5, 2.5, ..., \omega - x - .5$ 

Of course,  $E(YA_{x,s}^{CR}) = WLE_{x,s}^{CR}$  as previously calculated in (14).<sup>3</sup> The pmf defined in (15) captures the entire probability distribution of time spent in railroad activity and therefore one can compute any measure of central tendency (e.g., the mean, median, and mode), measures of dispersion and shape (e.g., standard deviation, skewness, and kurtosis), and any probability interval of interest. Table 1 shows two such characteristics—the mean and standard deviation of railroad activity based on the age and years of service and subject to decrements for mortality, disability, retirement, and other withdrawals.

decrements for mortality, disability, retirement, and other withdrawals. When worklife expectancies depend only on  $q_x^{(1)}$ ,  $q_{x,s}^{(2)}$ , and  $q_{x,s}^{(3)}$  [i.e., with withdrawals  $q_{x,s}^{(4)}$  excluded from formulae (12)–(14)], we refer to such expectancies as the Old AAR-Type Expectancy.<sup>4</sup> These expectancies resemble those previously published by the Association of American Railroads which depended on mortality, disability and age retirement decrements, but not withdrawals; these expectancies also appear in Table 1.

Our previous Increment-Decrement (ID) model, expanded slightly to allow for other than 30-year/age 60 retirements and updated with *Twenty Third Valuation* data produces the last two columns in Table 1, showing the ID Expectancies (i.e., the means) and standard deviations. In (16)  $YA_{x,s}^{ID}$  denotes the random variable measuring additional time spent in railroad activity. Transitions from *a* (active) and *i* (inactive) occur at the midpoint of any year using economy-wide transition probabilities until a railroad worker qualifies for a railroad retirement and railroad transition probabilities thereafter (Krueger, 2004; Skoog and Ciecka, 1998, 2001a, 2001b, 2002, 2004, and 2006).<sup>5</sup> The boundary conditions and recursions in (16) define the pmf's.

<sup>3</sup>Since  $E(YA_{x,s}^{CR}) = .5(1 - _1p_{x,s}) + 1.5(_1p_{x,s} - _2p_{x,s}) + 2.5(_2p_{x,s} - _3p_{x,s}) + ... + (\omega - x - .5)(_{\omega - x - 1}p_{x,s} - _{\omega - x}p_{x,s})$ , we have after collecting terms,  $E(YA_{x,s}^{CR}) = .5 + _1p_{x,s} + _2p_{x,s} + _3p_{x,s} + ... + _{\omega - x - 1}p_{x,s} = WLE_{x,s}^{CR}$ <sup>4</sup>See Skoog and Ciecka (1998 and 2006) for Old AAR-Type expectancies and Markov process

<sup>&</sup>lt;sup>4</sup>See Skoog and Ciecka (1998 and 2006) for Old AAR-Type expectancies and Markov process expectancies for railroad workers based on the *Twentieth Actuarial Valuation* and *Twenty-Second Actuarial Valuation*, respectively.

<sup>&</sup>lt;sup>5</sup>The Markov calculations in the last two columns of Table 1 utilize data from the *Twenty-Third Actuarial Valuation* for mortality probabilities and retirement probabilities. These results also are based on transition probabilities for all males in the U.S. population for ages less than 60 (Krueger, 2004). Transition probabilities are scaled to the mortality experience of railroad workers, and we set active-to-inactive transition probabilities equal to retirement probabilities (adjusted for mortality) of railroad workers for ages 60 and above if they have sufficient service credits to retire

Boundary Conditions:  

$$p_{YA}^{ID}(x, a, 0) = 0$$
  
 $p_{YA}^{ID}(x, a, .5) = {}^{a}p_{x}^{d} + {}^{a}p_{x}^{i}p_{YA}^{ID}(x + 1, i, 0)$   
 $p_{YA}^{ID}(x, i, 0) = {}^{i}p_{x}^{d} + {}^{i}p_{x}^{i}p_{YA}^{ID}(x + 1, i, 0)$ , for  $x = BA, ..., TA - 1$ 

(16)

Main Recursions

$$p_{YA}^{ID}(x, a, y) = {}^{a} p_{x}^{a} p_{YA}^{ID}(x+1, a, y-1) + {}^{a} p_{x}^{i} p_{YA}^{ID}(x+1, i, y-.5) ,$$
  
y = 1.5,2.5,3.5,...,TA - x + .5 for x = BA,...,TA - 1

$$p_{YA}^{ID}(x,i,y) = {}^{i}p_{x}^{a}p_{YA}^{ID}(x+1,a,y-.5) + {}^{i}p_{x}^{i}p_{YA}^{ID}(x+1,i,y),$$
  

$$y = 1,2,3,...,TA - x \text{ for } x = BA,...,TA - 1$$

where *BA* and *TA* denote beginning age and truncation age, respectively. Table 1 illustrates the following theoretical and empirical results:

- 1. Old AAR-Type expectancies exceed the New CR expectancies. This is due to the inclusion of other withdrawals as well as decrements related to death, disability, and retirement in the New CR expectancies.
- 2. The gap between the Old AAR-Type expectancies and New CR expectancies is quite large at young ages, but it narrows with age and years of service.
- 3. The ID expectancies hover around the values of the Old AAR-Type expectancies except at older ages with no railroad service. Within the context of the Old AAR-Type model, retirement is not possible for the latter group within the railroad (until the accumulation of five years of service) and decrements only occur through death or disability. This leads to relatively large Old AAR-Type worklife expectancies at older ages and no service.
- 4. In the competing risks model, standard deviations tend to be large relative to means at young ages; and, given age, standard deviations decline as service years increase. Standard deviations also tend to become large relative to their respective means (i.e., large coefficients of variation) at ages 60 and above.

within railroad pension rules. Active-to-active transition probabilities are thereby also determined. For example, consider a 45-year-old with 16 years of service. In 15 years, at age 60, such a person will have accumulated approximately an additional 13 years of service, on average; reaching a total service accumulation of 29 years but insufficient for railroad retirement at age 60. This person would be eligible for railroad retirement in approximately one year; and, at that point, Table S-10, "30 & Over" probabilities are used. However, a 45-year-old with 10 years of service will have accumulated service of approximately 23 years at age 60; and, at age 62, would still have less than 30 years of service. This person would qualify for railroad retirement at age 62, but we then use the "5-29" years of service column probabilities in Table S-10, noting that this person will never accumulate 30 years of service since worklife expectancy for such a person is only 16 years. We assume that inactive to active transition probabilities are zero once a person qualifies for retirement; this can occur at age 60 at the earliest. However, it may occur later, as indicated in the above examples. In the case of a 45-year-old with 10 service years, early retirement can occur at age 62 and active-to-active and inactive-to active transition probabilities are the economy-wide probabilities for ages 60 and 61; inactive-to active probabilities are zero for age 62 and beyond and active-to-inactive probabilities are taken from Table S-10.

254

5. Standard deviations from the competing risks model exceed Markov standard deviations at younger ages; the former are more than double the latter until approximately age 30. Markov standard deviations are slightly bigger than competing risks standard deviations after age 50.

Figures 1-3 and Table 2 reveal important information underlying the New CR Expectancies in Table 1. For example, Column (2) of Table 2 contains probabilities that a 20-year-old railroad worker who has no years of service will remain in that occupation for the number of years in Column (1). The probabilities are .431, .333, .250, and .035 for 10.5, 20.5, 30.5, and 40.5 years of additional railroad activity. The corresponding pmf in Figure 1 shows a mode of .5 years and a secondary peak at 40.5 years, and the standard deviation is large relative to the mean (the coefficient of variation is 1.03). Column (3) of Table 2 and Figure 2 tell a similar story for a 40-year-old, also with no railroad service. Figure 3, the pmf for a 40-year-old railroad worker with 15 years of service, exhibits much less variation and more symmetry about the mean. The probability is .574 of attaining another 15.5 years on the railroad as indicated in the last column in Table 2. The likelihood of such a person attaining the 30/60 retirement requirement is much greater than that of his younger or contemporary counterparts who have no previous railroad service.



Figure 1. PMF for Railroad Workers Age 20 with Zero Service



Figure 2. PMF for Railroad Workers Age 40 with Zero Service



Figure 3. PMF for Railroad Workers Age 40 with 15 Years of Service

Probability in Railroad Work for at Least the Number of Years Indicated in Column (1)					
Years	Age 20, Zoro Sorvico	Age 40, Zoro Sorvico	Age 40, 15 Sorvice Veers		
(1)	(9)	(2)	(4)		
(1)	(2)	0.849	(4)		
0.5	0.815	0.849	0.975		
1.5	0.704	0.705	0.932		
2.5	0.644	0.707	0.930		
5.5	0.000	0.605	0.805		
4.5	0.504	0.628	0.850		
5.5 6.5	0.534	0.576	0.000		
0.5	0.508	0.576	0.802		
7.5 8.5	0.466	0.534	0.300		
9.5	0.400	0.515	0.7752		
10.5	0.431	0.496	0.725		
11.5	0.451	0.478	0.697		
12.5	0.410	0.470	0.669		
13.5	0.402	0.445	0.603		
14.5	0.380	0.493	0.609		
15.5	0.371	0.423	0.574		
16.5	0.363	0.384	0.538		
17.5	0.356	0.365	0.505		
18.5	0.349	0.345	0.473		
19.5	0.342	0.321	0.300		
20.5	0.333	0.295	0.106		
20.5 21.5	0.325	0.235	0.052		
21.0 22.5	0.317	0.188	0.022		
23.5	0.310	0.150	0.016		
20.0 24 5	0.302	0.102	0.009		
25.5	0.294	0.060	0.005		
26.5	0.286	0.040	0.003		
20.0 27.5	0.278	0.028	0.002		
28.5	0.269	0.020	0.001		
29.5	0.260	0.015	0.001		
30.5	0.250	0.010	0.001		
31.5	0.240	0.007	0.000		
32.5	0.230	0.005	0.000		
33.5	0.220	0.004	0.000		
34.5	0.209	0.003	0.000		
35.5	0.197	0.002	0.000		
36.5	0.184	0.001	0.000		
37.5	0.173	0.001	0.000		
38.5	0.162	0.001	0.000		
39.5	0.098	0.000	0.000		
40.5	0.035	0.000	0.000		
41.5	0.017	0.000	0.000		
42.5	0.009	0.000	0.000		
43.5	0.005	0.000	0.000		
44.5	0.003	0.000	0.000		

 Table 2

 Probability of Future Time in Railroad Activity by Age and Service Years

### **IV. Conclusion**

Using competing risks or multiple decrement theory, which will be shown in future work to be a special case of the Markov model, and proper actuarial pension plan data, we can calculate worklife expectancies for time spent in an occupation in a region. Some actuarial data (e.g., for railroad workers) enable us to calculate occupational worklives regardless of region. These occupational worklife expectancies and their underlying pmf's do not tell us everything about a person's labor force activity, especially at young ages; but they do shed light on the reasonableness of assuming that future labor-force time will be in only one occupation. The 20-year-old railroad worker considered above has a significant expectation of future railroad time-15.08 years; but that is quite different from simply assuming that (say) the next 40 years will be in railroad activity. Skoog and Ciecka (2001b) show a worklife of 37.28 for initially active males (Krueger, 2004, reports 38.00 years), regardless of occupation and education; and, as a very rough approximation, we might estimate approximately 22 or 23 additional years in other occupations.<sup>6</sup> Among other variables, personal injury or wrongful death-related losses depend on duration of loss and a level of earnings and benefits. The earnings level and benefits of a 20-year-old railroad worker who has no years of service is linked to a worklife of 15.08 years; and the remainder of labor-force time is tied to the earnings and benefits level of non-railroad workers. This may be important if earnings in the railroad sector differ significantly from other occupations. In addition to railroad workers, it may be possible to calculate worklife expectancies for other occupations from actuarial and pension data.

### References

- Bowers, N. L., Hans U. Gerber, James C. Hickman, Donald A. Jones and Cecil J. Nesbitt, Actuarial Mathematics, Second Edition, Schaumburg, IL: Society of Actuaries, 1987.
- Bureau of the Actuary, *Twenty-Third Actuarial Valuation*, Chicago, IL: US Railroad Retirement Board, 2006.
- Jordan, C. W., Life Contingencies, Chicago, IL: The Society of Actuaries, 1991.
- Kalbfleisch, John D., and Ross L. Prentice, *The Statistical Analysis of Failure Time Data, Second Edition*, Hoboken, NJ: John Wiley and Sons, 2002.
- Krueger, Kurt V., "Tables of Inter-Year Labor Force Status of the U.S. Population (1998-2004) to Operate the Markov Model of Worklife Expectancy," *Journal of Forensic Economics*, 2004, 17(3), 313-381.

Seal, Hilary L., "Studies in the History of Probability and Statistics. XXXV, Multiple decrements or competing risks," *Biometrika*, 1977, 64(3), 429-439.

Skoog, Gary, and James Ciecka, "Worklife Expectancies of Railroad Workers," Journal of Forensic Economics, 1998, 11(3), 237-252.

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

<sup>&</sup>lt;sup>6</sup>This is only a very rough approximation of time spent in non-railroad activity. The exact calculation requires that one properly account for different transition probabilities in railroad and nonrailroad work and whether an exiting railroad worker remains active or transitions into inactivity.

\_\_\_\_\_, 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 \_\_\_\_\_, "Probability Mass Functions for Labor Market Activity Induced by the Markov (Increment-Decrement) Model of Labor Force Activity," *Economics Letters*, 2002, 77(3), 425-431.

\_\_\_\_\_, 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 \_\_\_\_\_, "Markov Model Worklife Expectancies and Association of American Railroads Type Worklife Expectancies of Railroad Workers Based on the *Twenty-Second Actuarial Valuation* of the US Railroad Retirement Board," *The Earnings Analyst*, 2006, 8, 13-25.

### Appendix

In this paper, we consider four sources of decrements and recognize that data may come in the form of net rates of decrements, probability decrements, or a mixture of both. This appendix contains formulae for the calculation of probabilities given that data come in the form of three probability decrements and one net rate, two probability decrements and two net rates, one probability decrement and three net rates, and four net rates. From symmetry in the superscripts, the problems below each solve a group of problems, i.e., there are  $_4C_1 = \binom{4}{1} = 4$  problems of type 1,  $_4C_2 = \binom{4}{2} = 6$  problems of type 2 below,  $_4C_3 = \binom{4}{3} = 4$  problems of type 3, and  $_4C_4 = \binom{4}{4} = 1$  problem of type 4.

Problem 1 (treated in the text): Given  $q_x^{(2)}$ ,  $q_x^{(3)}$ ,  $q_x^{(4)}$ , and  $q_x'^{(1)}$ , determine  $q_x^{(1)}$ .

Solution to Problem 1:

(A1) 
$$q_x^{(1)} = q_x^{(1)} [1 - .5(q_x^{(2)} + q_x^{(3)} + q_x^{(4)})].$$

Problem 2: Given  $q_x^{(3)}$ ,  $q_x^{(4)}$ , and  $q_x^{\prime(1)}$ ,  $q_x^{\prime(2)}$ , determine  $q_x^{(1)}$ ,  $q_x^{(2)}$ .

Solution to Problem 2: We use the equations

(A2) 
$$q_x^{(1)} = q_x^{(1)} [1 - .5(q_x^{(2)} + q_x^{(3)} + q_x^{(4)})]$$
$$q_x^{(2)} = q_x^{(2)} [1 - .5(q_x^{(1)} + q_x^{(3)} + q_x^{(4)})].$$

Solving (A2) for  $q_r^{(1)}$  yields

(A3) 
$$q_x^{(1)} = \frac{q_x^{\prime(1)} [1 - .5(q_x^{(3)} + q_x^{(4)})](1 - .5q_x^{\prime(2)})}{1 - (.5)^2 q_x^{\prime(1)} q_x^{\prime(2)}}.$$

To solve for  $q_x^{(2)}$ , replace superscript (1) with (2) and replace superscript (2) with (1) in (A3).

Problem 3: Given  $q_x^{(4)}$ , and  $q_x^{\prime(1)}$ ,  $q_x^{\prime(2)}$ ,  $q_x^{\prime(3)}$ , determine  $q_x^{(1)}$ ,  $q_x^{(2)}$ ,  $q_x^{(3)}$ .

Solution to Problem 3: We use the equations

(A4)  

$$q_{x}^{(1)} = q_{x}^{(1)} [1 - .5(q_{x}^{(2)} + q_{x}^{(3)} + q_{x}^{(4)})]$$

$$q_{x}^{(2)} = q_{x}^{(2)} [1 - .5(q_{x}^{(1)} + q_{x}^{(3)} + q_{x}^{(4)})]$$

$$q_{x}^{(3)} = q_{x}^{(3)} [1 - .5(q_{x}^{(1)} + q_{x}^{(2)} + q_{x}^{(4)})]$$

Solving (A4) for  $q_x^{(1)}$  yields

(A5) 
$$q_x^{(1)} = \frac{q_x^{\prime(1)}(1 - .5q_x^{\prime(4)})[1 - .5(q_x^{\prime(2)} + q_x^{\prime(3)}) + .5^2 q_x^{\prime(2)} q_x^{\prime(3)}]}{1 - .5^2 (q_x^{\prime(1)} q_x^{\prime(2)} + q_x^{\prime(1)} q_x^{\prime(2)} + q_x^{\prime(2)} q_x^{\prime(3)}) + 2(.5^3) q_x^{\prime(1)} q_x^{\prime(2)} q_x^{\prime(3)}]}$$

To solve for  $q_x^{(2)}$ , replace superscript (1) with (2) and replace superscript (2) with (1) in (A5); and to solve for  $q_x^{(3)}$ , replace superscript (1) with (3) and replace superscript (3) with (1) in (A5).

Problem 4: Given  $q_x^{\prime(1)}, q_x^{\prime(2)}, q_x^{\prime(3)}, q_x^{\prime(4)}$  determine  $q_x^{(1)}, q_x^{(2)}, q_x^{(3)}, q_x^{(4)}$ .

Solution to Problem 4: We use the equations

(A6)  

$$q_{x}^{(1)} = q_{x}^{\prime(1)} [1 - .5(q_{x}^{(2)} + q_{x}^{(3)} + q_{x}^{(4)})]$$

$$q_{x}^{(2)} = q_{x}^{\prime(2)} [1 - .5(q_{x}^{(1)} + q_{x}^{(3)} + q_{x}^{(4)})]$$

$$q_{x}^{(3)} = q_{x}^{\prime(3)} [1 - .5(q_{x}^{(1)} + q_{x}^{(2)} + q_{x}^{(4)})]$$

$$q_{x}^{(4)} = q_{x}^{\prime(4)} [1 - .5(q_{x}^{(1)} + q_{x}^{(2)} + q_{x}^{(3)})].$$

Solving (A6) for  $q_x^{(1)}$  yields

(A7a) 
$$q_x^{(1)} = \frac{q_x^{\prime(1)} \left[ 1 - .5 \sum_{i=2}^{4} q_x^{\prime(i)} + .5^2 \sum_{i=2}^{3} \sum_{j>i}^{4} q_x^{\prime(i)} q_x^{\prime(j)} - .5^3 q_x^{\prime(2)} q_x^{\prime(3)} q_x^{\prime(4)} \right]}{1 - .5^2 \sum_{i=1}^{3} \sum_{j>i}^{4} q_x^{\prime(i)} q_x^{\prime(j)} + 2(.5^3) \sum_{i=1}^{2} \sum_{j>i}^{3} \sum_{k>j}^{4} q_x^{\prime(i)} q_x^{\prime(j)} q_x^{\prime(j)} - 3(.5)^4 q_x^{\prime(1)} q_x^{\prime(2)} q_x^{\prime(3)} q_x^{\prime(4)} }$$

To solve for  $q_x^{(2)}$ , first write all summation terms in (A7a) term by term. Then replace superscript (1) with (2) and replace superscript (2) with (1); to solve for  $q_x^{(3)}$ , replace superscript (1) with (3) and replace superscript (3) with (1); and to solve for  $q_x^{(4)}$ , replace superscript (1) with (4) and replace superscript (4) with (1).We observe that (A7a) to second order is  $q_x'^{(1)} \{1-.5 \sum q_x'^{(i)}\}$  and so to first order is  $q_x'^{(1)}$ . Using (9a) and (9b) an exact solution for (A7a) requiring no linearity is given by:

$$q_{x}^{(1)} = \frac{\ln[1 - q_{x}^{\prime(1)}] \left[ \left\{ q_{x}^{\prime(1)} + q_{x}^{\prime(2)} + q_{x}^{\prime(3)} + q_{x}^{\prime(4)} \right\} - \sum_{i < j} q_{x}^{\prime(i)} q_{x}^{\prime(j)} + \sum_{i < j < k} q_{x}^{\prime(i)} q_{x}^{\prime(j)} q_{x}^{\prime(k)} + q_{x}^{\prime(1)} q_{x}^{\prime(2)} q_{x}^{\prime(3)} q_{x}^{\prime(4)} \right]}{\ln[1 - q_{x}^{\prime(1)}] + \ln[1 - q_{x}^{\prime(2)}] + \ln[1 - q_{x}^{\prime(3)}] + \ln[1 - q_{x}^{\prime(4)}]}$$

$$ln[1-q'^{(j)}_{x}] = -q'^{(1)}_{x} - \frac{(q'^{(1)}_{x})^{2}}{2} - \frac{(q'^{(1)}_{x})^{3}}{3} - \cdots$$

in (A7b) and grouping terms of the same order would permit a comparison of the approximation. The right hand sides of (A7a) and (A7b) are both  $q'^{(1)}_x$  to first order, using the fact that the leading term in the expanded denominator is  $\{q'^{(1)}_x + q'^{(2)}_x + q'^{(4)}_x\}$  and so cancels with that term in the numerator. It is the second order term that would provide a comparison with (A7a).

An alternative and more rigorous derivation of the key result (A1) above follows from use of the expansion  $(1+x)^a = 1 + ax + \frac{a(a-1)}{2}x^2 + o(x^2), -1 < x < 1$ . From (9b), which is exact, with  $x = -q_x^{(\tau)}$  and  $a = \frac{q_x^{(\tau)}}{q_x^{(\tau)}}$ , there results the equation, when

j is set to 1,

(A0) 
$$q'^{(1)}_{x} = q^{(1)}_{x} [1 + .5(q^{(2)}_{x} + q^{(3)}_{x} + q^{(4)}_{x})]$$

Since we can invert by  $\frac{1}{1 + .5(q_x^{(2)} + q_x^{(3)} + q_x^{(4)})} = 1 - .5(q_x^{(2)} + q_x^{(3)} + q_x^{(4)}) + o(1)$ , substitu-

tion into (A0) produces (9) in the text, which was derived differently, with the aid of a "linearity" assumption. This argument proves that (9) is not only correct without reliance on "linearity" but that it is correct to second order.

In fact, since (A1), (A3), (A5) and (A7a) were based on equalities to second order from (9), they can be no more accurate than to second order, despite their containing terms to third and higher orders (spurious precision). Below we drop terms of higher order than the second:

(A1) 
$$q_x^{(1)} = q_x^{\prime(1)} [1 - .5(q_x^{(2)} + q_x^{(3)} + q_x^{(4)})] + o(2)$$

(A3) 
$$q_r^{(1)} = q_r^{(1)} [1 - .5(q_r^{(3)} + q_r^{(4)} + q_r^{(2)})] + o(2)$$

(A5) 
$$q_{r}^{(1)} = q_{r}^{\prime(1)} \left[ 1 - .5(q_{r}^{(4)} + q_{r}^{\prime(2)} + q_{r}^{\prime(3)}) \right] + o(2)$$

(A7a) 
$$q_x^{(1)} = q_x^{\prime(1)} [1 - .5(q_x^{\prime(2)} + q_x^{\prime(3)} + q_x^{\prime(4)})] + o(2).$$

Inspection of the right hand sides of these 4 equations above is striking: we can replace any primed symbol with any un-primed symbol in the bracketed right hand side terms and still maintain equality, to within second order. What is true for the superscript (1) on the left hand side is true for all superscripts (2), (3) and (4). Given any set of primed and unprimed quantities, the right hand side of one of these equations will yield any unprimed quantity and inversion above permits one to go in the opposite direction.

260