MEASURING YEARS OF INACTIVITY, YEARS IN RETIREMENT, TIME TO RETIREMENT, AND AGE AT RETIREMENT WITHIN THE MARKOV MODEL*

GARY R. SKOOG AND JAMES E. CIECKA

Retirement-related concepts are treated as random variables within Markov process models that capture multiple labor force entries and exits. The expected number of years spent outside of the labor force, expected years in retirement, and expected age at retirement are computed—all of which are of immense policy interest but have been heretofore reported with less precisely measured proxies. Expected age at retirement varies directly with a person's age; but even younger people can expect to retire at ages substantially older than those commonly associated with retirement, such as age 60, 62, or 65. Between 1970 and 2003, men allocated most of their increase in life expectancy to increased time in retirement, but women allocated most of their increased life expectancy to labor force activity. Although people can exit and reenter the labor force at older ages, most 65-year-old men who are active in the labor force will not reenter after they eventually exit. At age 65, the probability that those who are inactive will reenter the labor force at some future time is .38 for men and .27 for women. Life expectancy at exact ages is decomposed into the sum of the expected time spent active and inactive in the labor force, and also as the sum of the expected time to labor force separation and time in retirement.

This article focuses on time spent in retirement, years to retirement, years of labor force inactivity, and functions of these variables, such as average age at retirement and average time in retirement for the U.S. population. We aim to provide a Markov process model that captures exits from, and reentries into, the labor market and to compute retirement-related characteristics of the U.S. population. We utilize new recursive formulae that treat retirement-related concepts as random variables with probability distributions implied by the Markov model. This approach to retirement enables us to account for the huge number of labor force activity/inactivity paths that can occur because transitions between activity and inactivity are possible throughout a person's life. Because entire probability distributions are captured, we are able to compute means and other distributional characteristics of retirement random variables, such as the probabilities that a person has actually retired or will never retire.

The article is organized as follows. We begin with a discussion of various notions of the term "retirement" and its measurement. Next, we explain our notation and counting conventions. Recursive probability formulae used to determine probability distributions are presented in the next section of the article, followed by empirical results. The final section contains a summary of results and discussion of the direction of future research.

MEANING AND MEASUREMENT OF RETIREMENT

Commonly Used Notions of Retirement

There is little agreement on the meaning of the term "retirement," how to best measure it, and how to account for labor force reentry. For some, to be retired is to have left a career

*Gary R. Skoog, Department of Economics, DePaul University; and Legal Econometrics, Glenview, IL. James E. Ciecka, Department of Economics, DePaul University, 1 East Jackson Boulevard, Chicago, IL 60604; e-mail: jciecka@depaul.edu. We gratefully acknowledge comments and suggestions made by anonymous referees, which greatly improved the quality and readability of the article. M. Gendell made valuable comments as well. We also thank Genevieve Thompson and Lauri Dietz for editorial assistance.

or occupation that may be followed by part-time or bridge jobs, ultimately ending in final separation from the labor force. Ruhm (1990), using 1971 data, estimated that 25% of those who viewed themselves as retired reentered the labor force after initially retiring. Reasons for reentry commonly include the need for health insurance before Medicare, loss of value of savings, improvements in health, changes in family status, and the erosion of a pension if not indexed (Blau 1994; Reimers and Honig 1993; Ruhm 1990; Rust and Phelan 1997).

Retirement may be self-identified, and different people may self-identify as retired or not retired even though they maintain substantially the same level of labor force attachment. Self-identification of retirement status may be important as a sociological variable that may influence behaviors for people who consider themselves either retired or not.

According to the Retirement Confidence Survey (RCS), people “who are age 65 or older and not employed full time” are retired (Helman, VanDerhei, and Copeland 2008:22). This definition is clear-cut once the term “not employed full time” is defined. As defined by the U.S. Bureau of Labor Statistics (BLS), “employed” means working at least one hour for pay per week or working at least 15 hours per week in an unpaid job, and “full-time employment” means usually working at least 35 hours per week (Bureau of Labor Statistics and U.S. Census Bureau 2006:5-2 and 5-3). Combining the RCS retirement definition and the BLS full-time employment definition, a person age 64 who works 20 hours per week is not retired, but a 65-year-old who works 25 hours per week is. The Current Population Survey (CPS), sponsored jointly by the U.S. Census Bureau and the BLS, identifies employed and unemployed people as composing the labor force; those not in the labor force are classified as retired, disabled, or others not in the labor force.[1] A 75-year-old who works five hours per week in gainful employment is “employed” and therefore cannot be retired according to the BLS but would be considered retired under the RCS standard. The BLS associates retirement with being out of the labor force and not wanting to work; being employed or unemployed (i.e., on layoff or not employed but actively seeking employment) automatically excludes the BLS’s retirement designation.

Retirement might be viewed as a state (e.g., no connection to the labor force or to a life-long occupation), or it could be viewed as a process of gradually reducing labor force activity. Hayward and Grady (1990:352) estimated multivariate increment-decrement worklife tables for older men for 1966–1983 and showed that “…retirement is not a single event in many men’s lives and that postretirement labor force participation represents a significant portion of older men’s working life expectancy.” They associated retirement with “a series of acts involving movements both out of and back into the labor force” and, finally, “the ending of the labor force career.” They viewed retirement as a process that ultimately ends with being out of the labor force.

Receipt of pension benefits may be a retirement marker to some people. The average age at which people elect to begin receiving Social Security payments may be the most common pension-related notion of retirement in the United States. From 1940 to 1985, the average age at which Social Security benefits were awarded for men fell from 68.1 to 63.7 (Social Security Administration 2007: Table 6.B5). This decline reflects many changes, including eligibility at age 62 beginning in 1956 for women and 1961 for men. In 2005, the

1. The first question in the CPS labor force module is “Did you work last week?” If the respondent is under age 50 and says “No, I’m retired,” that person is counted as out of the labor force because of retirement. If the respondent simply says “No,” questions continue until a reason other than retirement is established. If no clear reason is established, the respondent picks from a list of reasons (which includes retirement) for not having worked last week. If a person age 50 or older volunteers that he or she is retired, the person is immediately asked whether he or she wants to work. The person is marked as retired if the answer is “no,” and that part of the interview stops. If the person answers “yes,” that part of the interview continues in the usual manner to determine the reason for not working (Bureau of Labor Statistics and U.S. Census Bureau 2006:6-7). If a person is 50 or older, retirement is a voluntary declaration with the follow-up question about wanting to work. Retirement in the CPS corresponds to (1) not working and (2) not wanting to work.

average age at which benefits were awarded for men was 63.8. There was a similar decline for women (age 67.4 in 1940, 63.4 in 1985, and 63.7 in 2005).

In the literature, beginning with Reimers (1976), statisticians and demographers have attempted to discover a formula that would readily map observed labor force participation rates into the "average age of retirement," which is much more difficult to measure. Reimers formula essentially is

$$\overline{X}_R = \frac{\sum_{x=35}^{\infty}(x+0.5)(PR_x - PR_{x+1})P_{x+1}}{\sum_{x=35}^{\infty}(PR_x - PR_{x+1})P_{x+1}}\ , \tag{1}$$

where $\overline{X}_R$ is the average retirement age, x is age in years, PR_x is the labor force participation rate at age x, P_{x+1} is the number of people age $x + 1$, and $(PR_x - PR_{x+1})$ is the decline in labor force participation between ages x and $x + 1$. The formula maps participation rates into a number, dubbed "average retirement age."

This use of decreases in the labor force participation rates is reminiscent of the old conventional model of labor force participation and worklife expectancy, which was abandoned more than 25 years ago by the BLS (Bureau of Labor Statistics 1982). In Reimer's formulation, declining participation rates from age 35 are due to retirements. In addition, labor force movements go in only one direction in this formulation; reentry into the labor force is not allowed after an initial exit. The analytical structure mimics a life table in which movements go only from the living to the death state. Although one-way movements make perfect sense in a life-table context, labor markets are characterized by people entering, leaving, and then reentering—perhaps several times throughout their lives.

Gendell and Siegel (1992:23) noted the fundamental problem inherent in formulae like (1). They referenced the Markov model as a solution because "…exits and reentries to the labor force are not uncommon in the age range we examine (45 years to 75 years and over), making the criterion of withdrawal ambiguous. For this reason, we prefer the stricter criterion of permanent withdrawal, despite the challenge it poses in getting requisite data." Their skepticism of formulae like (1) notwithstanding, Gendell and Siegel (1992, 1993, 1996) and Gendell (1998, 2001) did not break the research link between measures of retirement and participation rates. In light of Ruhm's (1990) finding that 25% of those who leave the labor force later reenter, such models cannot measure retirement accurately.

A Markov Process Approach to Retirement

In this article, we eschew retirement formulae based on participation rates and all other formulae that allow only one-way labor force movements. We proffer new Markov process recursive formulae for retirement-related concepts that allow people to exit and reenter the labor force according to age-specific transition probabilities. Years spent in retirement, years inactive, years before retirement occurs, and age at retirement are the primary variables of interest. Our approach captures probability mass functions for these labor force random variables. Upon taking expectations, we compute age-specific average years in retirement (including people who never retire), expected years in retirement for those who actually retire, expected years of inactivity, the expected number of times people leave the labor force, and expected age at retirement. We decompose life expectancy in two ways: (1) as the sum of expected time spent in and out of the labor force, and (2) as the sum of expected time to final labor force separation and expected time in retirement. Aggregate measures of average retirement age and average time in retirement for the stationary and actual U.S. populations are also computed.

We use an unambiguous definition of "years in retirement" as inactive time after final separation from the labor force has occurred. We describe how gradual or abrupt the retirement process may be by comparing future years in retirement with all future inactive years at specific ages. We also measure the probability that an exit from the labor force is final (i.e., the probability of having arrived at the end of the reentry process), the time that passes before reentry if an exit from the labor force is not final, and the probability of spending no time in retirement.

NOTATION AND COUNTING CONVENTIONS

For the vast majority of individuals, especially those at younger ages, time will be spent both in the labor force (years of activity, YA) and out of the labor force (years of inactivity, YI). Years to final separation (YFS) measures time until the last exit from the labor force, which can occur because of death while active or through inactivity and then eventual death; and YIR denotes future years spent in retirement (i.e., future time after final labor force separation has occurred). YI, which counts all inactive time, is greater than or equal to YIR. Similarly, YFS exceeds or equals YA because periods of inactivity before final separation are included in YFS but not in YA. Letting YL_x denote the additional-years-of-life random variable at age x, the complete life expectancy $\mathring{e}_x \equiv E(YL_x)$ at any exact age x for either gender can be written as $\mathring{e}_x = E(YA_{x,m}) + E(YI_{x,m}) = E(YFS_{x,m}) + E(YIR_{x,m})$, where m indicates labor force status at age x, with $m = a$ for active and $m = i$ for inactive.[2] That is, life expectancy at any age can be thought of as the sum of expected additional years of labor force activity and inactivity; and life expectancy also is the sum of expected years to final separation from the labor force and years in retirement.

Not everyone retires; people who die while active never retire. In order to have a retirement, an individual must become inactive, stay inactive for some amount of time, and then die. Because the probability of dying while active is not counted in the probability mass function for the years-to-retirement random variable ($YTR_{x,m}$), the variable is defective in the sense that its probability mass sums to less than 1. The sum of the probability mass of $YTR_{x,m}$ over all times to retirement, when subtracted from 1, gives the probability that a person does not retire because death occurs while active. Years to retirement is a stopping time, and if the probability that the stopping time does not occur is positive, the associated random variable is defective. We may relate the size of this defect to another, closely related random variable, $YIR_{x,m}$, by defining the years in retirement as zero if one does not retire. With this identification, $YIR_{x,m}$ is not defective, and the probability that $YIR_{x,m} = 0$ is the defect in $YTR_{x,m}$ and the probability mass function. Conditioning on retirement actually occurring produces the random variable $CYTR_{x,m}$, which is not defective. The expected value of this random variable plus current age x, $E(CYTR_{x,m}) + x$, defines the average age at retirement for those in state m at age x, given that retirement actually occurs. In addition, by conditioning on $YIR_{x,i} > 0$, we create a new random variable, $CYIR_{x,m} > 0$, whose expected value is the average time spent in retirement by those who actually retire. See Table 1 for a summary of our random-variable notation.

As examples of our terminology, inactive individuals who experience the sequences below tally the associated values for various random variables commencing at age x:

Age x x+1 x+2 x+3 x+4 x+5	
(a) $i \to i \to i \to d$	$YIR_{x,i} = 2.5$, $YI_{x,i} = 2.5$, $YA_{x,i} = 0$, $YFS_{x,i} = 0$, $YTR_{x,i} = 0$
(b) $i \to i \to i \to a \to d$	$YIR_{x,i} = 0$, $YI_{x,i} = 2.5$, $YA_{x,i} = 1$, $YFS_{x,i} = 3.5$, $YTR_{x,i}$ undefined
(c) $i \to i \to i \to a \to i \to d$	$YIR_{x,i} = 1$, $YI_{x,i} = 3.5$, $YA_{x,i} = 1$, $YFS_{x,i} = 3.5$, $YTR_{x,i} = 3.5$

2. We use italicized capital letters with appropriate subscripts to denote random variables, but nonitalicized counterparts without subscripts are simply shorthand notation. For example, $YIR_{x,m}$ denotes the future-years-in-retirement random variable for a person in state m at age x, while the nonitalicized YIR with no subscripts simply stands for the term "years in retirement."

Table 1. Random-Variable Notation

Abbreviation	Meaning
YA_x	Years of future activity (i.e., being in the labor force, although not necessarily continuously) for a person age *x*.
YI_x	Years of future inactivity (i.e., being out of the labor force, although not necessarily continuously) for a person age *x*.
YFS_x	Years to final separation from the labor force for a person age *x*. YFS includes all time, whether active or inactive, before final labor force separation occurs.
YIR_x	Years in retirement for a person age *x*. YIR begins after a person's final labor force separation; if death causes labor force separation, we count zero years in retirement because no years are spent in the retirement.
YL_x	Years of future life for a person age *x*.
$CYIR_x$	Conditional years in retirement for a person age *x*, assuming some amount of time in retirement (i.e., final labor force separation has occurred but not because of death).
$CYTR_x$	Conditional years to retirement for a person age *x*, assuming some amount of time in retirement (i.e., final labor force separation will occur but not because of death).

An individual is observed at exact age x, and transitions take place at the midpoints between exact ages. Thus, at age x in example (a), the transition to the death state occurred at the midpoint between exact ages $x + 2$ and $x + 3$; and 2.5 years were spent inactive and in future retirement.[3] In example (b), despite 2.5 years of inactivity from age x, this person reentered the labor force at age $x + 2.5$ and died while active. Because the person did not die in retirement (i.e., while inactive), $YTR_{x,i}$ is undefined; but $YIR_{x,i} = 0$, and $YI_{x,i} = 2.5$. We draw a distinction between (a) and (b) in regard to YTR: in example (a), retirement has already occurred, implying zero years to retirement; but in example (b), retirement never occurred, leading to an undefined value for $YTR_{x,i}$. In addition, YFS and YTR are equal when YTR is defined. In example (c), retirement occurred 3.5 years after age x at final separation from the labor force, and only one year was spent in retirement even though $YI_{x,i} = 3.5$.

The probability that a person in state m at age x will be in state n at age $x + 1$ is denoted by ${}^{m}p_x^{n}$, with $m \in \{a,i\}$ and $n \in \{a,i,d\}$, where d denotes the death state. We assume that transitions between state m at age x and state n at age $x + 1$ occur at age $x + 0.5$ (i.e., mid-period transitions). We define BA (beginning age) to be the earliest exact age for labor force activity, and TA (truncation age) is the youngest exact age at which everyone is dead. On the assumption that labor force activity always is possible if a person is alive at age BA or beyond, TA is the youngest exact age at which no labor force activity can occur. Everyone alive at age $TA - 1$ dies at age $TA - 0.5$; thus, ${}^{a}p_{TA-1}^{d} = {}^{i}p_{TA-1}^{d} = 1$ because the only transition at age $TA - 1$ is to the death state. That is, ${}^{a}p_x^{a} = {}^{a}p_x^{i} = {}^{i}p_x^{a} = {}^{i}p_x^{i} = 0$ for $x \geq TA - 1$. For all ages between BA and TA, ${}^{a}p_x^{a} + {}^{a}p_x^{i} + {}^{a}p_x^{d} = 1$, ${}^{i}p_x^{a} + {}^{i}p_x^{i} + {}^{i}p_x^{d} = 1$, and the transition to the death state does not depend on labor force status, ${}^{a}p_x^{d} = {}^{i}p_x^{d} = p_x^{d}$.

If $RV_{x,m}$ denotes any of the random variables $YA_{x,m}$, $YI_{x,m}$, $YFS_{x,m}$, $YIR_{x,m}$, $CYIR_{x,m}$, $YTR_{x,m}$, or $CYTR_{x,m}$, then $p_{RV}(x,m,y)$ is the probability that $RV_{x,m} = y$ for a person who is in state m at exact age x. An individual's retirement date cannot be determined with certainty until that person dies, but we view that as a desirable feature of this formulation because some older people reenter the labor force at advanced ages for various reasons. Our framework does permit the computation of $p_{YA}(x,i,0)$, the probability of zero additional years of

3. That retirement, by our midpoint convention, took place at age $x - 0.5$, or earlier, and so it is left-censored. Thus, throughout this article, we mean years of retirement occurring in the future, beyond age x.

activity given inactivity at age x, which may be interpreted as the model's determination of the probability that a person will remain retired.

RECURSIONS FOR YA, YI, YFS, YIR, CYIR, YTR, AND CYTR

This section contains sets of recursions that enable us to compute probability mass functions for YA, YI, YFS, YIR, CYIR, YTR, and CYTR.[4] These recursions solve a difficult computational problem arising from an unmanageably large number of labor force–related paths that a person's life may traverse. For example, consider a 20-year-old and two labor force states (active or inactive). By age 80, there are $2^{60} = 1.15 \times 10^{18}$ possible paths to consider besides death, which might occur anywhere along these paths and would thereby increase the number of possible states that may occur. The huge number of possible labor force–related paths occurs because of possible exits and reentries into the labor force. This two-way movement is very different from the one-directional movement from living to dead in survivor and life expectancy studies. The following recursions also enable us to calculate entire probability distributions, not only averages such as life expectancy.

We provide a unique set of recursions for each random variable we consider; but all sets of recursions have the same general structure consisting of global conditions, boundary conditions, and main recursions. Global conditions apply to all random variables and are stated only once in (2a)–(2d).

$$p_{RV}(x,a,y) = p_{RV}(x,i,y) = 0 \quad \text{if } y < 0 \text{ or } y > TA - x - 0.5 \tag{2a}$$

$$p_{RV}(TA,a,0) = p_{RV}(TA,i,0) = 1 \tag{2b}$$

$${}^{a}p_x^{d} = {}^{i}p_x^{d} = 1 \quad \text{for } x = TA - 1 \tag{2c}$$

$${}^{a}p_x^{a} = {}^{a}p_x^{i} = {}^{i}p_x^{a} = {}^{i}p_x^{i} = 0 \quad \text{for } x = TA - 1 \tag{2d}$$

Boundary conditions pertain to initial values of random variables—values that occur with zero probability or probabilities that can be computed in a recursive manner without feedback effects. The main recursions compute all remaining probabilities for those initially active and inactive in the labor force.

The boundary conditions and main recursions for YA [(3a)–(3e)] and YI [(4a)–(4e)] for ages $x = BA,\ldots,TA - 1$ are mirror images of each other; one need only substitute a for i and i for a in the YA probability-mass-function recursions to generate the YI recursions. The essential observation is that ${}^{a}p_x^{a}$ transitions result in a one-year increase in YA but a zero-year increase in YI, ${}^{i}p_x^{i}$ transitions increase YI by one year but cause no change in YA, and ${}^{a}p_x^{i}$ and ${}^{i}p_x^{a}$ result in a half-year increase in both YA and YI. Similarly, ${}^{a}p_x^{d}$ produces a half-year gain in YA but no increase in YI; for ${}^{i}p_x^{d}$, the reverse is true.[5]

$$p_{YA}(x,a,0) = 0 \tag{3a}$$

$$p_{YA}(x,a,0.5) = {}^{a}p_x^{d} + {}^{a}p_x^{i}\,p_{YA}(x+1,i,0) \tag{3b}$$

$$p_{YA}(x,i,0) = {}^{i}p_x^{d} + {}^{i}p_x^{i}\,p_{YA}(x+1,i,0) \tag{3c}$$

4. YA and YFS recursions (Skoog 2002; Skoog and Ciecka 2002, 2003) are repeated here to facilitate comparisons drawn throughout the article, but the focus of this article is on the new recursions for YI, YIR, CYIR, YTR, and CYTR.

5. See Demography's Web page for illustrations of probability mass functions for YA, YI, YFS, CYIR, YTR, and CYTR. The Web site also has figures showing the probability of inactive men and women remaining inactive and average years in retirement, by exact age.

$$p_{YA}(x,a,y) = {}^a p_x^a p_{YA}(x+1,a,y-1) + {}^a p_x^i p_{YA}(x+1,i,y-0.5),$$
$$y = 1.5, 2.5, \ldots, TA - x + 0.5 \tag{3d}$$

$$p_{YA}(x,i,y) = {}^i p_x^a p_{YA}(x+1,a,y-0.5) + {}^i p_x^i p_{YA}(x+1,i,y), \quad y = 1,2,\ldots,TA-x \tag{3e}$$

$$p_{YI}(x,i,0) = 0 \tag{4a}$$

$$p_{YI}(x,i,0.5) = {}^i p_x^d + {}^i p_x^a p_{YI}(x+1,a,0) \tag{4b}$$

$$p_{YI}(x,a,0) = {}^a p_x^d + {}^a p_x^a p_{YI}(x+1,a,0) \tag{4c}$$

$$p_{YI}(x,i,y) = {}^i p_x^i p_{YI}(x+1,i,y-1) + {}^i p_x^a p_{YI}(x+1,a,y-0.5),$$
$$y = 1.5, 2.5, \ldots, TA - x + 0.5 \tag{4d}$$

$$p_{YI}(x,a,y) = {}^a p_x^i p_{YI}(x+1,i,y-0.5) + {}^a p_x^a p_{YI}(x+1,a,y), \quad y = 1,2,\ldots,TA-x \tag{4e}$$

The relationship between YFS and YIR is more intricate than that between YA and YI. First, the main recursions for YIR [(6f)–(6g)] hold for integer values, whereas the main recursions for YFS [(5e)–(5f)] are defined on the half integers starting at 1.5 years. For YIR, an active person must turn inactive in order for YIR to exceed zero; and because such a transition occurs at a half-integer age and death also occurs at a half-integer age, the resulting YIR will have an integer value. The same reasoning applies to an initially inactive person who turns active before once again turning inactive in retirement. The YFS argument on the right sides of the main YFS recursions is always $y - 1$ years at age $x + 1$ because all of the transitions ${}^a p_x^a$, ${}^a p_x^i$, ${}^i p_x^a$, and ${}^i p_x^i$ add one more year to YFS relative to age x. In a complementary manner, all of the transitions add nothing to YIR relative to age x; thus, the YIR argument on the right sides of the main YIR recursions is y, except for the term ${}^a p_x^i p_{YIR}(x+1,i,y-0.5)$. That term enters into the calculation of $p_{YIR}(x,a,y)$ in (6f) because the ${}^a p_x^i$ transition may be the final transition that leads to retirement, thereby causing $y - 0.5$ years in retirement at age $x + 1$ to be y years in retirement at age x. The probability $p_{YIR}(x+1,i,y-0.5)$ itself comes from the YIR boundary condition (6e), $p_{YIR}(x,i,y) = {}^i p_x^i p_{YIR}(x+1,i,y-1)$, defined on the half integers beginning with $y - 1 = 0.5$ years because an inactive person (who stays inactive and then eventually dies) will die at a half-integer age. The remaining YIR boundary conditions apply to YIR of zero, 0.5, or half integers for those initially active.

YFS probability mass functions for ages $x = BA,\ldots.,TA - 1$ are defined by (5a)–(5f):

$$p_{YFS}(x,a,y) = 0, \quad y = 0,1,2,3,\ldots,TA-1 \tag{5a}$$

$$p_{YFS}(x,i,y) = 0, \quad y = 0.5,1,2,3,\ldots,TA-1 \tag{5b}$$

$$p_{YFS}(x,a,0.5) = {}^a p_x^d + {}^a p_x^i p_{YFS}(x+1,i,0) \tag{5c}$$

$$p_{YFS}(x,i,0) = {}^i p_x^d + {}^i p_x^i p_{YFS}(x+1,i,0) \tag{5d}$$

$$p_{YFS}(x,a,y) = {}^a p_x^a p_{YFS}(x+1,a,y-1) + {}^a p_x^i p_{YFS}(x+1,i,y-1),$$
$$y = 1.5, 2.5, \ldots, TA - x + 0.5 \tag{5e}$$

$$p_{YFS}(x,i,y) = {}^i p_x^a p_{YFS}(x+1,a,y-1) + {}^i p_x^i p_{YFS}(x+1,i,y-1),$$
$$y = 1.5, 2.5, \ldots, TA - x + 0.5 \tag{5f}$$

YIR probability mass functions for ages $x = BA, \ldots, TA - 1$ are defined by (6a)–(6g):

$$p_{YIR}(x,i,0.5) = {}^{i}p_x^{d} \tag{6a}$$

$$p_{YIR}(x,a,y) = 0, \qquad y = 0.5, 1.5, 2.5, \ldots, TA - x - 0.5 \tag{6b}$$

$$p_{YIR}(x,i,0) = {}^{i}p_x^{a} p_{YIR}(x+1,a,0) + {}^{i}p_x^{i} p_{YIR}(x+1,i,0) \tag{6c}$$

$$p_{YIR}(x,a,0) = {}^{a}p_x^{a} p_{YIR}(x+1,a,0) + {}^{a}p_x^{i} p_{YIR}(x+1,i,0) + {}^{a}p_x^{d} \tag{6d}$$

$$p_{YIR}(x,i,y) = {}^{i}p_x^{i} p_{YIR}(x+1,i,y-1), \qquad y = 1.5, 2.5, \ldots, TA - x + 0.5 \tag{6e}$$

$$p_{YIR}(x,a,y) = {}^{a}p_x^{a} p_{YIR}(x+1,a,y) + {}^{a}p_x^{i} p_{YIR}(x+1,i,y-0.5) + {}^{a}p_x^{i} p_{YIR}(x+1,i,y), \qquad y = 1, 2, \ldots, TA - x \tag{6f}$$

$$p_{YIR}(x,i,y) = {}^{i}p_x^{a} p_{YIR}(x+1,a,y) + {}^{i}p_x^{i} p_{YIR}(x+1,i,y), \qquad y = 1, 2, \ldots, TA - x \tag{6g}$$

Recursion (6a) means that an inactive person age x gets credit for half a year of retirement if that person dies at age $x + 0.5$. Recursion (6b) recognizes that an active person must turn inactive at a half-integer age to accumulate any time in retirement; the person then dies at a half-integer age, implying an integer value of YIR and zero probability of a half-integer YIR. Recursion (6c) gives the probability that an inactive person age x spends zero years in retirement by turning active or remaining inactive but, in either case, not ever retiring from age $x + 1$. Recursion (6d) works the same way as the previous boundary condition but with the additional probability that an active person age x could die and thereby not ever retire.

CYIR (i.e., YIR conditioned on $YIR_{x,a} > 0$) probability mass functions are defined in (7a) and (7b) for ages $x = BA, \ldots, TA - 1$. Here we condition on some amount of time in retirement greater than zero.

$$p_{CYIR}(x,a,y) = \frac{p_{YIR}(x,a,y)}{\sum_{y=0.5}^{TA-x-0.5} p_{YIR}(x,a,y)}, \qquad y = 0.5, 1.0, 1.5, 2.0, \ldots, TA - x - 0.5 \tag{7a}$$

$$p_{CYIR}(x,i,y) = \frac{p_{YIR}(x,i,y)}{\sum_{y=0.5}^{TA-x-0.5} p_{YIR}(x,i,y)}, \qquad y = 0.5, 1.0, 1.5, 2.0, \ldots, TA - x - 0.5. \tag{7b}$$

The YTR recursions (8a)–(8f) are the same as the YFS recursions with one exception: the YFS boundary condition (5c), $p_{YFS}(x,a,0.5) = {}^{a}p_x^{d} + {}^{a}p_x^{i} p_{YFS}(x+1,i,0)$, reverts to the simpler YTR condition (8c), $p_{YTR}(x,a,0.5) = {}^{a}p_x^{i} p_{YTR}(x+1,i,0)$. That is, the death probability ${}^{a}p_x^{d}$ contributes half a year to YFS because death is one way to separate from the labor force; but ${}^{a}p_x^{d}$ does not contribute to YTR because retirement requires turning inactive and dying inactive. It is this missing probability piece that causes YTR to be a defective random variable and the sum of its probability mass to be less than 1. The CYTR "cures" this defect by dividing each YTR probability by the sum of all YTR probabilities [see (9a) and (9b)]. The CYTR random variable's expected value is the average years to retirement for people who actually retire. The expected age at retirement is the expected value of CYTR plus current age.

YTR defective probability mass functions for ages $x = BA, \ldots, TA - 1$ are defined by (8a)–(8f).

$$p_{YTR}(x,a,y) = 0, \qquad y = 0, 1, 2, 3, \ldots, TA - 1 \tag{8a}$$

$$p_{YTR}(x,i,y) = 0, \quad y = 0.5,1,2,3,\ldots,TA - 1 \tag{8b}$$

$$p_{YTR}(x,a,0.5) = {}^{a}p_x^{i} p_{YTR}(x + 1,i,0) \tag{8c}$$

$$p_{YTR}(x,i,0) = {}^{i}p_x^{d} + {}^{i}p_x^{i} p_{YTR}(x + 1,i,0) \tag{8d}$$

$$p_{YTR}(x,a,y) = {}^{a}p_x^{a} p_{YTR}(x + 1,a,y - 1) + {}^{a}p_x^{i} p_{YTR}(x + 1,i,y - 1), \quad y = 1.5,2.5,\ldots,TA - x + 0.5 \tag{8e}$$

$$p_{YTR}(x,i,y) = {}^{i}p_x^{a} p_{YTR}(x + 1,a,y - 1) + {}^{i}p_x^{i} p_{YTR}(x + 1,i,y - 1), \quad y = 1.5,2.5,\ldots,TA - x + 0.5 \tag{8f}$$

CYTR probability mass functions for ages $x = BA,\ldots,TA - 1$ are given in (9a) and (9b).

$$p_{CYTR}(x,a,y) = \frac{p_{YTR}(x,a,y)}{\sum_{y=0}^{TA-x-0.5} p_{YTR}(x,a,y)}, \quad y = 0,0.5,1.5,2.5,\ldots,TA - x - 0.5 \tag{9a}$$

$$p_{CYTR}(x,i,y) = \frac{p_{YTR}(x,i,y)}{\sum_{y=0}^{TA-x-0.5} p_{YTR}(x,i,y)}, \quad y = 0,1.5,2.5,\ldots,TA - x - 0.5 \tag{9b}$$

The foregoing recursions accomplish three important tasks. First, they allow for any number of labor force exits and entries for an individual according to age-specific transition probabilities. Second, they yield the probabilities $p_{RV}(x,m,y)$ for all of the random variables we consider, but they do so in an extremely efficient manner. Finally, they give us entire probability distributions whose characteristics can be tabulated.

DATA AND EMPIRICAL RESULTS

Data

The Current Population Survey (CPS), an ongoing survey of labor force characteristics carried out since 1940, is the primary source of U.S. labor force statistics.[6] The CPS provides data for official BLS estimates of labor force participation, employment, and unemployment for the noninstitutionalized U.S. population. One member of a household (the "household respondent") typically provides information on all members of a household unit, with each household member being designated as a "respondent." Interviewers gather information on labor force status (i.e., either active or inactive) for the week prior to interviews. The CPS has a 4-8-4 monthly structure. That is, interviews are conducted in four consecutive months, followed by an eight-month hiatus in which no interviews are conducted, and then followed by interviews in four consecutive months. Thus, the entire process takes 16 months for one group of respondents to complete. We designate the eight months of sample data by $S_1,S_2,\ldots,S_8$, keeping in mind that eight calendar months separate S_4 and S_5. The CPS thus affords four opportunities to make one-year-apart comparisons of labor force characteristics of specific individuals by using any of the pairs S_1S_5, S_2S_6, S_3S_7, or S_4S_8. Estimates of transition probabilities ${}^{a}p_x^{a}$, ${}^{a}p_x^{i}$, ${}^{i}p_x^{a}$, ${}^{i}p_x^{i}$, can be garnered from any of these pairs.[7] We discuss that procedure next and then consider the advantages of using the S_4S_8 pairing.

---

6. See Bureau of Labor Statistics and U.S. Census Bureau (2006) for more background and technical details about the CPS.

7. The CPS includes about 60,000 households, and it is the only survey that can produce the very large number of observations needed to estimate transition probabilities. Even with the large number of observations,

Let ${}^aN_x$ and ${}^iN_x$ denote the number of people active and inactive in the labor force at reported age x from CPS data. Then ${}^aN_x = {}^aN_x^a + {}^aN_x^i$, where ${}^aN_x^a$ denotes the number of those active at reported age x who remain active at reported age $x + 1$, and ${}^aN_x^i$ denotes the number who become inactive at reported age $x + 1$. Similarly, ${}^iN_x = {}^iN_x^a + {}^iN_x^i$, where ${}^iN_x^a$ denotes the number of those inactive at reported age x who become active at reported age $x + 1$, and ${}^iN_x^i$ is the number of those inactive at age reported x who remain inactive. Our estimated values of ${}^ap_x^a$, ${}^ap_x^i$, ${}^ip_x^a$, ${}^ip_x^i$, denoted by ${}^a\hat{p}_x^a, {}^a\hat{p}_x^i, {}^i\hat{p}_x^a, {}^i\hat{p}_x^i$, are computed with

$$ {}^a\hat{p}_x^a = \left[\frac{{}^aN_{x-1}^a + {}^aN_x^a}{{}^aN_{x-1} + {}^aN_x}\right](1 - \hat{p}_x^d) \tag{10a} $$

$$ {}^a\hat{p}_x^i = \left[\frac{{}^aN_{x-1}^i + {}^aN_x^i}{{}^aN_{x-1} + {}^aN_x}\right](1 - \hat{p}_x^d) \tag{10b} $$

$$ {}^i\hat{p}_x^i = \left[\frac{{}^iN_{x-1}^i + {}^iN_x^i}{{}^iN_{x-1} + {}^iN_x}\right](1 - \hat{p}_x^d) \tag{10c} $$

$$ {}^i\hat{p}_x^a = \left[\frac{{}^iN_{x-1}^a + {}^iN_x^a}{{}^iN_{x-1} + {}^iN_x}\right](1 - \hat{p}_x^d). \tag{10d} $$

The CPS records ages as integers; but people are, on average, one-half year older than their reported integer age. For example, people with recorded age 30 in the CPS may have just had their 30th birthday, or they be only a few days away from their 31st birthday. On average, data recorded for 30-year-old respondents in the CPS refer to people age 30.5. By averaging these respondents with those who declare themselves age 29 (who are actually 29.5, on average), we create transition probability estimates for exact age 30. Formulae (10a)–(10d) show this procedure. Formulae (10a)–(10d) also contain a mortality adjustment using age-specific mortality probabilities from the 2003 U.S. life tables (Arias 2007).

In order to estimate transition probabilities that span one year, we must match respondents using one of the months-in-sample pairs S_1S_5, S_2S_6, S_3S_7, or S_4S_8. However, we want this matched sample to be representative of the U.S. labor force. Since 1994, the CPS has computed what it calls "outgoing rotation weights," which force the S_4S_8 subsample of the CPS "to sum to the composited estimates of employment, unemployment, and not in the labor force each month" (Bureau of Labor Statistics and U.S. Census Bureau 2000:10–12). In other words, use of the outgoing rotation weights with S_4S_8 eliminates concern that these specific parts of the CPS may not agree with overall CPS measures of labor force characteristics. However, one additional problem remains in ensuring that we have a nationally representative data set to estimate transition probabilities: specific people must be matched between S_4 and S_8. That is, we want to know a specific person's labor force status (active or inactive) in the fourth month in sample and the same person's status one year later in the eighth month in sample. For whatever reason (e.g., people move, die, cannot be contacted), the matched sample of people is about 75% of the S_4 and S_8 samples. Krueger (2004) adjusted the outgoing rotation weights on the matched S_4S_8 sample to equal the sum of the outgoing rotation weights of the entire S_4S_8 sample. This final process eliminates any differences in employment, unemployment, and labor force participation between the matched S_4S_8 sample and the entire CPS sample.

it is desirable to use data for several consecutive years. Suppose that we want to estimate two independent transition probabilities for each exact age between 21 and 80 for both men and women for four levels of educational attainment with a set of 60,000 households, which generates 100,000 usable observations. The average number of observations per estimated transition probability would be only $104 \approx 100{,}000 / (2 \times 60 \times 2 \times 4)$.

The estimated transition probabilities used in this article are based on approximately 674,000 matched records for men and women combined from the S_4S_8 samples over the period 1999–2004 (Krueger 2004). Because this period includes the end of an economic expansion, a short recession, and the onset of a new expansion, it should not be subject to a particular part of the business cycle.

Our empirical work proceeded in the following manner. Estimated transition probabilities, using (10a)–(10d), were based on CPS data. These transition probabilities are the primitive terms in all recursions. Probability mass functions were then computed for our random variables using all recursions from (3a) through (9b). With probability mass functions for all ages in hand, we computed expected values and other distributional characteristics at individual ages. Tables 2–7 contain those characteristics for selected ages.[8]

Results for Specific Ages

Tables 2–7 contain empirical results for YA, YI, YFS, YIR, CYIR, and CYTR characteristics for initially active men at ages 20, 30, 40, 50, 55, 60, 65, 70, and 75. Each row contains the mean, median, mode, standard deviation, skewness, and kurtosis, as well as the interquartile range and the 10% and 90% probability points.

We draw several observations from these tables. First, complete life expectancy in 2003 (Arias 2007: Table 2) is decomposed as $\dot{e}_x = E(YA_{x,a}) + E(YI_{x,a}) = E(YFS_{x,a}) + E(YIR_{x,a})$ at each age x. This result holds because YA and YI are disjoint and inclusive of all time alive, as is the pair YFS and YIR. At any age, the sum of the entries in the "Mean" columns of Tables 2 and 3 and the sum of the entries in Tables 4 and 5 give life expectancy values. For example, at age 65, $E(YA_{65,a}) = 4.47$, $E(YI_{65,a}) = 12.28$, $E(YFS_{65,a}) = 5.85$, $E(YIR_{65,a}) = 10.90$, and $\dot{e}_x = 16.75 = 4.47 + 12.28 = 5.85 + 10.90$.

Second, the expected value of YIR, $E(YIR_{x,a})$, varies little (between 11.33 and 11.99 years) within the interval bounded by ages 20 and 60 (see Table 5). A similar pattern holds for $E(CYIR_{x,a})$: approximately 13.89 to 14.68 years between ages 20 and 60 (see Table 6). That is, active men ages 20 to 60 (including those who never retire) can expect to spend approximately 11 to 12 years in retirement and 14 or more years in retirement if they actually retire. Men who are still active at age 65 can expect 10.90 years in retirement after final labor force separation, and those who actually retire can expect 12.69 years in retirement after final labor force separation.

Third, YI and YIR differ less with increasing age, although $E(YI)$ always exceeds $E(YIR)$. For example, YI and YIR are 17.69 and 11.33 at age 20, but they are 12.28 and 10.90, respectively, at age 65 (see Tables 3 and 5). This result reflects the idea that inactivity is more closely associated with retirement at older ages than at younger ages. The exit and reentry process may take some time to work itself out; but at age 65, expected inactive and retirement time differ by only 1.38 years (= 12.28 – 10.90). That is, expected retirement time is approximately 89% of all expected future inactive time. However, at ages 40 and 50, expected years in retirement are only 0.73 (= 11.67 / 15.99) and 0.78 (= 11.93 / 15.37) of expected inactive time, respectively.[9]

8. All the numbers (transition probabilities) on the right sides of (10a)–(10d) can be computed from Krueger (2004) for all males, females, four levels of educational attainment, and regardless of education, as well as from mortality probabilities in Arias (2007). Tables 2–7 are based on these transition probabilities. As a separate exercise, we estimated transition probabilities for 1997–1998 and found very similar transition probabilities (Skoog and Ciecka 2002).

9. The expectancies in Tables 3 and 5 are related to those in Tables 2 and 4. For example, we noted that expected inactivity exceeds expected years in retirement by 1.38 years for active 65-year-old males. When looked at from an activity point of view, expected years of activity (4.47 in Table 2) are smaller than expected years to final separation (5.85 in Table 4) by the same 1.38 years. This is expected time spent outside of the labor force before expected retirement.

Table 2. YA Characteristics for Initially Active Men, Regardless of Education

Age	Mean	Median	Mode	SD	Skewness	Kurtosis	Interquartile Probability Interval		10%–90% Probability Interval	
							25%	75%	10th %	90th %
20	38.11	39.50	41.50	9.14	–1.06	5.10	33.50	44.50	26.50	48.50
30	29.98	30.50	32.50	8.02	–0.68	3.92	25.50	35.50	19.50	39.50
40	21.25	21.50	22.50	7.06	–0.34	3.20	17.50	25.50	11.50	29.50
50	13.10	13.50	12.50	5.74	0.15	2.94	9.50	16.50	5.50	20.50
55	9.47	9.50	7.50	4.95	0.49	3.14	6.50	12.50	3.50	16.50
60	6.31	5.50	2.50	4.18	0.89	3.63	3.50	8.50	1.50	12.50
65	4.47	3.50	0.50	3.51	1.12	4.13	1.50	6.50	0.50	9.50
70	3.49	2.50	0.50	2.90	1.22	4.20	1.50	4.50	0.50	7.50
75	2.91	2.50	0.50	2.28	0.93	3.30	0.50	4.50	0.50	6.50

Source: Calculated using YA recursions (3a)–(3e).

Table 3. YI Characteristics for Initially Active Men, Regardless of Education

Age	Mean	Median	Mode	SD	Skewness	Kurtosis	Pr (*YI* = 0)	Interquartile Probability Interval		10%–90% Probability Interval	
								25%	75%	10th %	90th %
20	17.69	17.00	0.00	11.87	0.34	2.41	.06	8.00	26.00	2.00	34.00
30	16.53	16.00	0.00	11.60	0.33	2.38	.10	7.00	25.00	0.00	32.00
40	15.99	16.00	0.00	11.30	0.32	2.34	.12	6.00	24.00	0.00	31.00
50	15.37	15.00	0.00	10.65	0.27	2.27	.11	6.00	23.00	0.00	30.00
55	14.86	15.00	0.00	10.12	0.25	2.25	.10	6.00	22.00	0.00	29.00
60	14.08	14.00	0.00	9.46	0.26	2.26	.10	6.00	21.00	1.00	27.00
65	12.28	12.00	0.00	8.55	0.33	2.31	.10	5.00	19.00	0.00	24.00
70	9.93	9.00	0.00	7.48	0.46	2.44	.13	4.00	15.00	0.00	20.00
75	7.54	7.00	0.00	6.31	0.65	2.70	.18	2.00	12.00	0.00	17.00

Source: Calculated using YI recursions (4a)–(4e).

Table 3 contains the probabilities that active men will never be inactive, Pr(YI = 0), and Table 5 shows the probability active men will not retire, Pr(YIR = 0). These probabilities are .10 and .14, respectively, at age 65. The former is smaller than the latter because a currently active person can withdraw from the labor force (making $YI > 0$), reenter the labor force, and never retire afterward (leaving $YIR = 0$).

Separate recursions give the expected number of labor force exits and accessions (see the appendix). At age 65, active males can expect 1.25 labor force exits (including an exit due to death) and 0.25 reentries.[10] We might say that approximately three-fourths of men

10. At ages 20, 30, 40, 50, 55, 60, 65, 70, and 75, expected exits are 2.60, 1.93, 1.74, 1.58, 1.49, 1.39, 1.25, 1.14, 1.08, respectively. Expected reentries are equal to expected exits less 1.00.

Table 4. YFS Characteristics for Initially Active Men, Regardless of Education

Age	Mean	Median	Mode	SD	Skewness	Kurtosis	Interquartile Probability Interval		10%–90% Probability Interval	
							25%	75%	10th %	90th %
20	44.47	45.50	42.50	10.83	–0.89	5.10	39.50	51.50	31.50	57.50
30	35.03	35.50	32.50	9.82	–0.44	3.85	30.50	41.50	22.50	47.50
40	25.57	25.50	22.50	9.02	–0.08	3.20	20.50	31.50	14.50	37.50
50	16.54	15.50	12.50	7.95	0.37	2.94	11.50	21.50	6.50	27.50
55	12.34	11.50	7.50	7.24	0.65	3.08	7.50	16.50	3.50	22.50
60	8.55	7.50	2.50	6.40	0.96	3.52	3.50	12.50	1.50	18.50
65	5.85	4.50	0.50	5.20	1.29	4.51	1.50	8.50	0.50	13.50
70	4.19	2.50	0.50	3.94	1.57	5.84	1.50	5.50	0.50	10.50
75	3.22	2.50	0.50	2.83	1.61	7.19	0.50	4.50	0.50	6.50

Source: Calculated using YFS recursions (5a)–(5f).

Table 5. YIR Characteristics for Initially Active Men, Regardless of Education

Age	Mean	Median	Mode	SD	Skewness	Kurtosis	Pr (YIR = 0)	Interquartile Probability Interval		10%–90% Probability Interval	
								25%	75%	10th %	90th %
20	11.33	9.00	0.00	10.47	0.74	2.81	.23	1.00	19.00	0.00	26.00
30	11.48	10.00	0.00	10.46	0.73	2.80	.22	1.00	19.00	0.00	27.00
40	11.67	10.00	0.00	10.43	0.70	2.76	.21	2.00	19.00	0.00	27.00
50	11.93	10.00	0.00	10.22	0.63	2.59	.18	3.00	19.00	0.00	27.00
55	11.99	11.00	0.00	9.92	0.57	2.49	.17	3.00	19.00	0.00	26.00
60	11.84	11.00	0.00	9.42	0.51	2.41	.15	3.00	19.00	0.00	25.00
65	10.90	10.00	0.00	8.54	0.50	2.43	.14	3.00	17.00	0.00	23.00
70	9.23	8.00	0.00	7.43	0.57	2.54	.15	3.00	14.00	0.00	20.00
75	7.23	6.00	0.00	6.25	0.71	2.79	.19	2.00	11.00	0.00	16.00

Source: Calculated using YIR recursions (6a)–(6g).

exit only once, and one-fourth exit twice [i.e., 1.25 exits = (1 exit)(3/4) + (2 exits)(1/4)]. Of course, other fractions are possible, and some active 65-year-olds will exit three or more times. However, a large number of exits for many in this group is not possible because the expected value is 1.25. At least three-fourths of the group has only one exit, and therefore most labor force activity and subsequent inactivity is uninterrupted after age 65.[11]

Fourth, expected age at retirement $x + E(CYTR_{x,a})$ for active men, the sum of the first two columns of Table 7, is approximately age 66.5 until $x = 40$. After age 40, it gradually increases to age 78.2 for men who were active at age 75.

11. The probability mass values for 1, 2, 3, and 4 labor force exits (including death) are .771, .206, .022, and .001, respectively, at age 65. Expected exits are .771(1) + .206(2) + .022(3) + .001(4) + … = 1.253.

Table 6. CYIR Characteristics for Initially Active Men, Regardless of Education

Age	Mean	Median	Mode	SD	Skewness	Kurtosis	Interquartile Probability Interval 25%	Interquartile Probability Interval 75%	10%–90% Probability Interval 10th %	10%–90% Probability Interval 90th %
20	14.68	13.00	1.00	9.63	0.63	2.90	7.00	21.00	3.00	28.00
30	14.71	13.00	1.00	9.63	0.63	2.90	7.00	21.00	3.00	28.00
40	14.72	13.00	1.00	9.60	0.62	2.86	7.00	21.00	3.00	28.00
50	14.62	13.00	1.00	9.41	0.57	2.69	7.00	21.00	3.00	28.00
55	14.39	13.00	1.00	9.15	0.52	2.58	7.00	21.00	3.00	27.00
60	13.89	13.00	5.00	8.71	0.48	2.51	7.00	20.00	3.00	26.00
65	12.69	12.00	8.00	7.89	0.49	2.53	6.00	18.00	3.00	24.00
70	10.90	10.00	5.00	6.86	0.55	2.65	5.00	16.00	2.00	21.00
75	8.89	8.00	2.00	5.77	0.67	2.89	4.00	13.00	2.00	17.00

Source: Calculated using CYIR recursions (7a)–(7b).

Table 7. CYTR Characteristics for Initially Active Men, Regardless of Education

Age	Mean	Median	Mode	SD	Skewness	Kurtosis	Interquartile Probability Interval 25%	Interquartile Probability Interval 75%	10%–90% Probability Interval 10th %	10%–90% Probability Interval 90th %
20	46.44	46.50	42.50	8.55	–0.32	4.91	41.50	51.50	36.50	57.50
30	36.50	36.50	32.50	8.28	–0.04	3.73	31.50	41.50	26.50	47.50
40	26.64	26.50	22.50	8.11	0.10	3.37	21.50	31.50	16.50	37.50
50	17.00	16.50	12.50	7.61	0.40	3.07	11.50	21.50	7.50	27.50
55	12.51	11.50	7.50	7.08	0.66	3.16	7.50	16.50	4.50	22.50
60	8.48	7.50	2.50	6.30	0.98	3.58	3.50	12.50	1.50	17.50
65	5.73	4.50	0.50	5.13	1.32	4.60	1.50	8.50	0.50	13.50
70	4.11	2.50	0.50	3.90	1.59	5.87	1.50	5.50	0.50	9.50
75	3.23	2.50	0.50	2.84	1.56	6.80	0.50	4.50	0.50	6.50

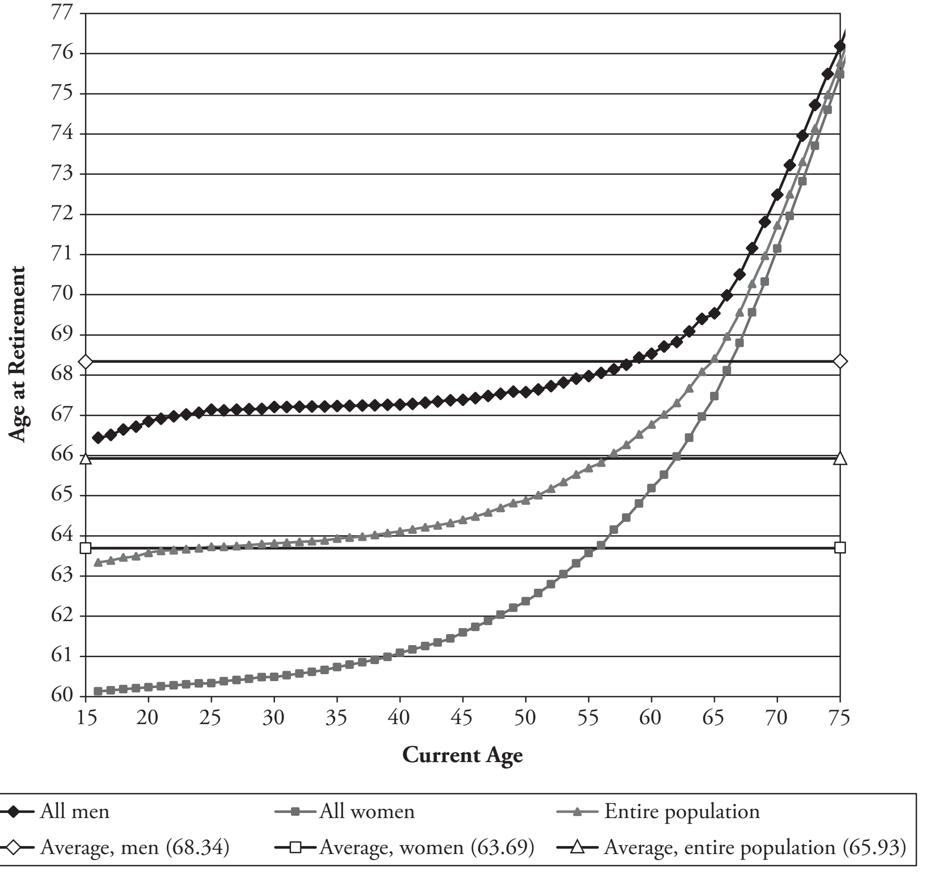
Source: Calculated using CYTR recursions (9a)–(9b).

Fifth, we know from life tables that the probability mass function of the additional-years-of-life random variable is skewed to the left at young ages and skewed to the right at older ages. The same relationship holds for YA and YFS (see Tables 2 and 4). However, YI and YIR are skewed to the right for all ages, implying that the mean usually exceeds the median (see Tables 3 and 5). For example, expected years in retirement at age 65 for active men is 10.90 years, but 50% of all such men will have 10.00 or fewer years in retirement.

Aggregation Results

Our methodology allows us to make broad statements about changes that occurred in the average age at retirement between 1970 and 2003. Based on stationary populations,

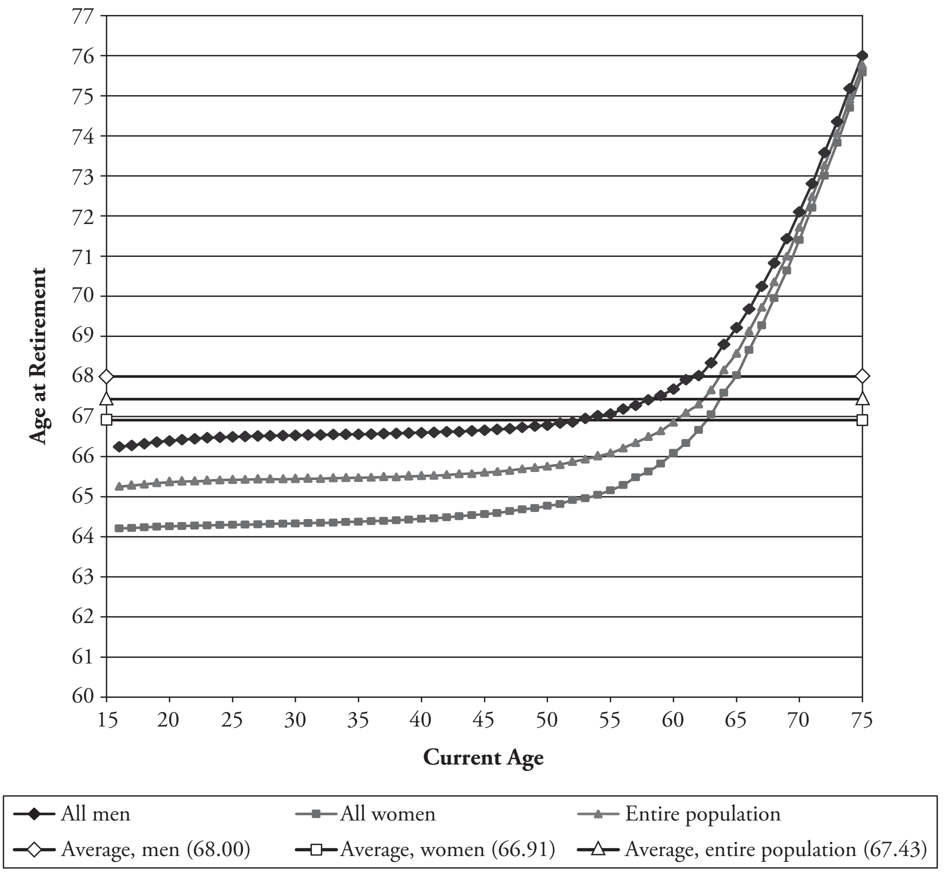
Figure 1. Average Age at Retirement for All Men, All Women, and the Entire Population in 1970



average age at retirement for males changed very little (from age 68.82 to 69.08) between 1970 and 2003. The increase in average retirement age of women over the same period was more dramatic (from age 65.16 to 68.22). For the actual U.S. populations in 1970 and 2003, there was a slight decline in men's average age at retirement, from age 68.34 to 68.00 (see the highest horizontal lines in Figures 1 and 2) and an increase of 3.22 years for women from age 63.69 to 66.91 (the lowest horizontal lines in Figures 1 and 2).[12] When aggregated over both genders, average age at retirement increased 1.50 from age 65.93 in 1970 to 67.43 in 2003 (the middle horizontal lines in Figures 1 and 2). Figures 1 and 2 also show the average age at retirement for any current age in 1970 and 2003. For

12. Figure 1 is based on 1970 transition probabilities from Bureau of Labor Statistics (1982); mortality probabilities are as of 2003 (Arias 2007). Figure 2 is based on the same transition probabilities as Tables 2–7. In both figures, we use participation rates and their complements as weights for active and inactive groups, and we use population weights in the aggregation process.

Figure 2. Average Age at Retirement for All Men, All Women, and the Entire Population in 2003



example, in 1970 a 40-year-old woman's expected retirement age was approximately age 61, but in 2003 a 40-year-old woman's expected age at retirement exceeded 64.

Average ages at retirement for the actual population (68.00 for men and 66.91 for women in 2003) are strikingly larger than the age at onset of Social Security payments, which occur at approximately age 63.8. This is not surprising, since people may collect Social Security and continue to work, either before normal retirement age (incurring a penalty if they earn too much) or after normal retirement age. Furthermore, for those who retire, reenter, and retire again, it is the final retirement that is incorporated into our calculations.

SUMMARY AND DIRECTION OF FUTURE RESEARCH

We treated YA, YI, YFS, YIR, CYIR, and CYTR as random variables within a Markov model and allowed for exit and entry into the labor force according to age-specific transition probabilities. Characteristics of these random variables are summarized in Tables 2–7 for active males for selected ages, but similar tables could have been produced for all exact ages between 18 and 75, for initially inactive men, for women, and for various educational levels. The recursive formulae we used enabled us to deal efficiently with what would otherwise be an unmanageably large number of paths of labor market activity and inactivity and also to compute entire probability distributions.

Many of our main conclusions can be stated as expectations. To this end, we think of a typical life table that has a final column showing complete life expectancy, $\mathring{e}_x$, at all ages x in the table. The life table could be extended to include several additional columns providing information about labor market activity and inactivity. Table 8, for currently active men in 2003, illustrates this idea. Age and complete life expectancy are in columns (1) and (2), respectively. Expected years of activity in column (3) and expected inactive time in column (4) sum to life expectancy given in column (2) at each age x, as do expected years to final separation from the labor force in column (5) and expected time in retirement in column (6). If a person actually retires, column (7) gives average years in retirement conditional on retirement actually occurring, which must exceed column (6) because $E(YIR)$ includes the probability of never retiring. Column (8) gives average waiting time to retirement under the assumption that some time will actually be spent in retirement. Expected age at retirement, the sum of age in column (1) and the conditional years to retirement expectancy in column (8), is shown in column (9). Note that expected age at retirement increases monotonically with age: it hovers about age 66.5 for active men who are currently between the ages 20 and 40, but is age 74.11 for active 70-year-old men. Even for young men, expected retirement age exceeds ages like 62 or 65 because we have allowed for the possibility of labor force reentry. Column (10) shows the expected number of exits from the labor force at each age—a number that must be at least 1.00 because the table refers to active members of the labor force who must exit the labor force through either retirement or death. Column (11) shows the probability of never being inactive beyond age x, and column (12) contains the probability of never retiring. The probability of retiring is given by [1 – column (12)], which, when multiplied by $E(CYIR)$ in column (7) and added to $E(YFS)$ in column (5), also yields life expectancy in column (2). Finally, similar extended life tables by gender, education, initial labor force status, and a fuller set of exact ages would improve understanding of retirement-related behaviors across society.

In addition to the findings summarized in the extended life table, we found the average age at retirement in the actual U.S. population in 2003 to be age 68.00 for all men and age 66.91 for women. Both ages are substantially older than the average age at the onset of Social Security benefits.

The increase in life expectancy from 1970 to 2003 has been used differently by men and women. At age 45, 90% of the increase in life expectancy for men had been devoted to additional years in retirement, but women had allocated only 20% of their increase in life expectancy to years in retirement.

We found the probability that an inactive person is actually retired given age and gender. If an inactive person were to reenter the labor force, we also found the probability for the length of time that passes before reentry into the labor force occurs. Inactive men and women age 65 have probabilities of .62 and .73 of remaining inactive for the remainder of their lives (i.e., staying retired). If reentry were to occur, however, the reentry probability declines with the length of time since becoming inactive; but the probability is approximately .40 that reentry will not occur until age 70 or beyond for both men and women. (See the supplemental figures on Demography's Web site: http://www.populationassociation.org/publications/demography.)

The Markov model is a singular improvement over measures of retirement that make once-out, always-out assumptions about being in the labor force, but it remains an autoregressive process of order 1 (AR1). Being active in the labor force at age x implies certain transition probabilities of remaining active or becoming inactive at age $x + 1$, regardless of a person's active and inactive states at ages $x - 1$, $x - 2$,…. Similarly, inactivity at age x affects activity status at age $x + 1$ in a manner that does not depend on status at ages

Table 8. Extended Life Table for Men Active in the Labor Force at Various Ages

Age (1)	Life Expectancy $\dot{e}_x$ (2)	Worklife Expectancy *E*(*YA*) (3)	Inactive Expectancy *E*(*YI*) (4)	Final Labor Force Separation Expectancy *E*(*YFS*) (5)	Years in Retirement Expectancy *E*(*YIR*) (6)	Conditional Years in Retirement Expectancy *E*(*CYIR*) (7)	Conditional Years to Retirement Expectancy *E*(*CYTR*) (8)	Retirement Age Expectancy Age + *E*(*CYTR*) (9)	Expected Number of Labor Force Exits *E*(*NE*) (10)	Pr(*YI* = 0) (11)	Pr(*YIR* = 0) (12)
20	55.80	38.11	17.69	44.47	11.33	14.68	46.44	66.44	2.55	.06	.23
30	46.51	29.98	16.53	35.03	11.48	14.71	36.50	66.50	2.00	.10	.22
40	37.24	21.25	15.99	25.57	11.67	14.72	26.64	66.64	1.77	.12	.21
50	28.47	13.10	15.37	16.54	11.93	14.62	17.00	67.00	1.59	.11	.18
55	24.33	9.47	14.86	12.34	11.99	14.39	12.51	67.51	1.49	.10	.17
60	20.39	6.31	14.08	8.55	11.84	13.89	8.48	68.48	1.38	.10	.15
65	16.75	4.47	12.28	5.85	10.90	12.69	5.73	70.73	1.25	.10	.14
70	13.42	3.49	9.93	4.19	9.23	10.90	4.11	74.11	1.13	.13	.15
75	10.45	2.91	7.54	3.22	7.23	8.89	3.23	78.23	1.07	.18	.19

Source: Column (2) is from Arias (2007: table 2); column (3) is from Table 2; column (4) is from Table 3; column (5) is from Table 4; column (6) is from Table 5; column (7) is from Table 6; column (8) is from Table 7; column (9) is the sum of column (1) and column (8); column (10) is computed with recursions given in the appendix; column (11) is from Table 3; column (12) is from Table 5.

younger than age x. Autoregressive models of order 2 or higher would be more realistic. Certainly, a 40-year-old person who has been in the labor force for the last 10 consecutive years has a stronger attachment to the labor force than a 40-year-old who was inactive between ages 31–39 and entered the labor force at age 40. However, both have the same Markov transition probability of remaining active or turning inactive at age 41. Two problems would occur, one analytical and the other empirical, if we were to attempt to consider higher-order autoregressive processes. New recursive formulae would have to be specified, and longer-term longitudinal data would be required to estimate transitional probabilities. The first problem is solvable for AR2 or AR3 processes, but the latter problem is more formidable. A sufficiently large data set (to estimate transition probabilities) does not exist that traces labor force activity for even two calendar years, the minimum amount of time in an AR2 model. Data sufficient for an AR3 process would be more unlikely. A possible approach to this problem might entail a parametric formulation generating AR2 or higher-order transition probabilities that would feed into appropriate higher-order recursions.

APPENDIX

Number of labor force exits (NE) probability mass functions for ages $x = BA,\ldots,TA-1$ are as follows:

$$p_{NE}(x,i,0) = {}^ip_x^d + {}^ip_x^i p_{NE}(x+1,i,0)$$

$$p_{NE}(x,a,1) = {}^ap_x^a p_{NE}(x+1,a,1) + {}^ap_x^d + {}^ap_x^i p_{NE}(x+1,i,0)$$

$$p_{NE}(x,a,y) = {}^ap_x^a p_{NE}(x+1,a,y) + {}^ap_x^i p_{NE}(x+1,i,y-1)$$

$$y = 2,\ldots,(TA+1-x)/2 \text{ for } x \text{ even}$$

$$y = 2,\ldots,(TA-x)/2 \text{ for } x \text{ odd}$$

$$p_{NE}(x,i,y) = {}^ip_x^a p_{NE}(x+1,a,y) + {}^ip_x^i p_{NE}(x+1,i,y)$$

$$y = 1,2,\ldots,(TA-1-x)/2 \text{ for } x \text{ even}$$

$$y = 1,2,\ldots,(TA-x)/2 \text{ for } x \text{ odd}$$

Years until a state switch (YSS) (to a living state) defective probability mass functions for $x = BA,\ldots,TA-1$ are shown below. The defect in this random variable occurs because ${}^ap_x^d$ and ${}^ip_x^d$ do not contribute to a state switch to a living state.

$$p_{YSS}(x,a,0) = 0 \text{ and } p_{YSS}(x,i,0) = 0$$

$$p_{YSS}(x,a,0.5) = {}^ap_x^i \text{ and } p_{YSS}(x,i,0.5) = {}^ip_x^a \quad \text{for } x = BA,\ldots,TA-1$$

$$p_{YSS}(x,a,y) = {}^ap_x^a p_{YSS}(x+1,a,y-1) \quad y = 1.5,2.5,\ldots,TA-x+0.5$$

$$p_{YSS}(x,i,y) = {}^ip_x^i p_{YSS}(x+1,i,y-1) \quad y = 1.5,2.5,\ldots,TA-x+0.5$$

Conditional years until a state switch (CYSS) probability mass functions for ages $x = BA,\ldots,TA-1$ are as follows:

$$p_{CYSS}(x,a,y) = \frac{p_{YSS}(x,a,y)}{\sum_{y=0.5}^{TA-x-0.5} p_{YSS}(x,a,y)} \qquad y = 0.5,1.5,2.5,\ldots,TA-x+0.5$$

$$p_{CYSS}(x,i,y) = \frac{p_{YSS}(x,i,y)}{\sum_{y=0.5}^{TA-x-0.5} p_{YSS}(x,i,y)} \qquad y = 0.5,1.5,2.5,\ldots,TA-x+0.5$$

REFERENCES

Arias, E. 2007. "United States Life Tables, 2003." *National Vital Statistics Reports* 54(14). Hyattsville, MD: National Center for Health Statistics.

Blau, D. M. 1994. "Labor Force Dynamics of Older Men." *Econometrica* 62:117–56.

Bureau of Labor Statistics. 1982. "Tables of Working Life: The Increment-Decrement Model." Bulletin 2135. U.S. Department of Labor, Washington, DC.

Bureau of Labor Statistics and U.S. Census Bureau. 2000. "Design and Methodology, Current Population Survey." Technical Paper 63. BLS and U.S. Census Bureau, Washington, DC.

———. 2006. "Design and Methodology, Current Population Survey." Technical Paper 66. BLS and U.S. Census Bureau, Washington, DC.

Gendell, M. 1998. "Trends in Retirement Age in Four Countries, 1965–95." *Monthly Labor Review* 121(2):20–30.

———. 2001. "Retirement Age Declines Again in 1990s." *Monthly Labor Review* 124(10):12–21.

Gendell, M. and J.S. Siegel. 1992. "Trends in Retirement Age by Sex, 1950–2005." *Monthly Labor Review* 115(7):22–29.

———. 1993. "Retirement Quandary: More Retirees at Younger Ages, Living Longer." *Population Today* 21(3):6–7,9.

———. 1996. "Trends in Retirement Age in the United States, 1955–1993, by Sex and Race." *Journal of Gerontology, Series B: Psychological and Social Sciences* 51:S132–S139.

Hayward, M.D. and W.R. Grady. 1990. "Work and Retirement Among a Cohort of Older Men in the Unites States, 1966–1983." *Demography* 27:337–56.

Helman, R., J. VanDerhei, and C. Copeland. 2008. "Retirement Confidence Survey, 2008: Americans Much More Worried About Retirement, Health Costs a Big Concern." EBRI Issue Brief No. 316. Employee Benefit Research Institute, Washington, DC.

Krueger, K.V. 2004. "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* 17:313–81.

Reimers, C. 1976. "Is the Average Age at Retirement Changing?" *Journal of the American Statistical Association* 71:552–58.

Reimers, C. and M. Honig. 1993. "The Perceived Budget Constraint Under Social Security: Evidence for Reentry Behavior." *Journal of Labor Economics* 11:184–204.

Ruhm, C.J. 1990. "Bridge Jobs and Partial Retirement." *Journal of Labor Economics* 8:482–501.

Rust, J. and C. Phelan. 1997. "How Social Security and Medicare Affect Retirement Behavior in a World of Incomplete Markets." *Econometrica* 65:781–831.

Skoog, G.R. 2002. "Worklife Expectancy: Theoretical Results." Paper presented at Allied Social Science Association Meeting, Atlanta.

Skoog, G.R. and J.E. Ciecka. 2002. "Probability Mass Functions for Labor Market Activity Induced by the Markov (Increment-Decrement) Model of Labor Force Activity." *Economics Letters* 77:425–31.

———. 2003. "Probability Mass Functions for Years to Final Separation From the Labor Force Induced by the Markov (Increment-Decrement) Model." *Journal of Forensic Economics* 16:51–86.

Social Security Administration (SSA). 2007. "Annual Statistical Supplement to the Social Security Bulletin, 2006." SSA Publication No. 13-11700. SSA, Washington, DC.