Expected Labor Force Activity and Retirement Behavior by Age, Gender, and Labor Force History

James E. Ciecka and Gary R. Skoog

ABSTRACT
We find and estimate probability mass functions for labor force related random variables. Complete life expectancy (by age, gender, and two years of labor force history) is decomposed into expected years of future labor force activity and inactivity as well as into expected years until final separation from the labor force and expected years in retirement. We also calculate expected age at retirement and expected years in retirement for people who actually retire. Two consecutive years of inactivity, especially in middle age, is a key indicator for both men and women when accounting for future labor force participation and retirement. For example, women (men) who are out of the labor force at age 49 and again out of the labor force at age 50, can expect to be in the labor force seven (eight) fewer years in the future than their counterparts who were in the labor force at ages 49 and 50. In addition, they have expected retirement ages 4.5–5.5 years younger than their active counterparts.

1. Introduction
There has been a long-term secular decline in the U.S. male labor force participation rate from approximately 86% in 1950 to 69% in 2016. Female labor force participation grew in the second half of the last century but peaked at slightly over 60% in 2000 and is approximately 57% today. These trends reflect the composite influence of many factors such as private and public retirement policies, macroeconomic conditions, the age distribution and technological skills of the U.S. population, and immigration. The participation rate, and its complement, tells us the fraction of a population that is active in the labor force at a point in time and the fraction of a population inactive in the labor force. However, participation rates tell us little about what we might expect regarding future years of labor force activity, inactivity, time people spend in retirement, and age at retirement. In this article, we provide estimates of expected future years of activity, expected years of inactivity, expected years in retirement, and expected age at retirement, given a person's age, gender, and labor force history, while recognizing that people can enter and leave the labor force multiple times throughout life. We accomplish our objective by defining various labor force related behaviors as random variables, finding their probability mass functions, and computing their expectations. These expectations enable us to make two observations for public policy: (1) when inactive, it is important to move people back to activity before inactivity extends for two years and (2) the largest gains in future activity accrue to people in middle age moving back into the labor force.

We define a person's future years of activity as time in the labor force and years of inactivity as years out of the labor force. Their sum is additional years of life. We use an unambiguous definition of years in retirement as inactive time after final separation from the labor force has occurred, and age at retirement is the age at which final labor market separation occurs due to inactivity but not death. The complement of years in retirement is years to final separation from the labor force, and, when added to years in retirement, also gives a person's additional years of life. Therefore, complete life expectancy can be thought of as the sum of expected years of labor force activity and expected years of inactivity as well as the sum of expected years until final separation from the labor force and expected years in retirement.

It may be the case that declining participation rates portend problems about the long-term ability of the U.S. economy to provide private and public goods. If so, then there are important public policy questions regarding how best to extend expected years in the labor force (reduce expected years of labor force inactivity), extend expected years until final separation from the labor force (reduce expected years in retirement), and extend expected age at retirement. We show that the largest expected gains accrue to moving people in middle age from inactivity to activity and lesser, but still substantial, gains for younger and older groups. We also show that quick reversal for labor force inactivity to activity is important.

In part, a period life table affords an apt analogy to tables in this article. A life table computed with age-specific mortality probabilities shows life expectancies at various exact ages. There typically is no claim to predict future longevity but rather the more modest goal to measure the implications of current age-specific mortality probabilities on life expectancy and other
characteristics of future years of life. In this article, we have a similar objective but for labor-force related behavior, including time outside of the labor force and time in retirement. As with a life table, there is no claim that expectations predict the future but only that they provide estimates of the future if current age-specific transition probabilities remain unchanged. Beyond this point, the analogy to life tables breaks down because life tables contain only one-way transitions from living to dead, whereas transitions between labor market activity and inactivity may occur several times within a person’s life with death being the only absorbing state. Two Bureau of Labor Statistics (BLS) publications [Bulletin 2135 (Smith 1982); and Bulletin 2254 (Smith 1986)] recognized this very issue in their increment–decrement (Markov process) models. Previously, Schoen and Land (1979) gave a general algorithm for estimating increment–decrement tables and applied their methods to marital states which could change throughout a person’s life. They also noted that transitions between labor force states could be analyzed with the increment-decrement model. The BLS publications applied increment–decrement methods to calculate expected future time in the labor force (i.e., worklife expectancy) and conditioned transition probabilities on age, gender, race or ethnicity, and current labor force state. In this article, we recognize that more labor force history upon which current models rest. In 2002, Skoog and Ciecka (2002) treated time in the labor force as a random variable and specified its entire probability distribution within the increment–decrement model; and they showed that worklife expectancy was one of the characteristics of the time-in-the-labor-force random variable. Skoog, Ciecka, and Krueger (2011) produced the most current set of tables for worklife expectancy and other distributional characteristics such as median, standard deviation, skewness, kurtosis, and quantiles for time in the labor force. However, all of the foregoing work has been based on a person’s current labor force state. In this article, we recognize that more labor force history may provide valuable information. We capture more information by conditioning on a person’s current labor force status and the person’s labor force status one year earlier, thereby doubling the labor force history upon which current models rest. We also move well beyond worklife expectations and calculate expectations for future years in retirement, future years until final separation from the labor force, future years of inactivity, and the probability of retirement with our second-order models that allow any number of labor force accessions and exits. Others (e.g., Gendell and Siegel, 1992) have recognized the importance of this approach when studying retirement but were unable to deal with in-and-out nature of labor market activity. The article is organized as follows. Section 2 explains our concept of retirement, notation, and counting conventions. Section 3 describes data used to estimate probability mass functions. Empirical results are in Section 4. The final section is a conclusion.

2. Retirement, Notation, and Counting Conventions

2.1. Retirement

Before introducing our notation and counting conventions, we explain our concept of retirement. The drop in labor force participation rates that begins to occur in middle age is a possible gauge of inactivity and retirement. Reimers (1976) proposed a formula in which declining participation rates after age 35 measure retirements; exits from the labor force were assumed to be final and conceptually similar to deaths in a life table used to calculate life expectancy and expected age at death. Of course, one-way movements from living to dead make sense in a life table; but one-way labor market transitions are inappropriate because people can enter and exit the labor force several times throughout their lives. Measures like Reimer’s do not capture the dynamic nature of labor markets in which people may make multiple labor force entries and exits throughout their lives (Ruhm 1990; Maestas 2010).

Some measures of retirement link retirement to important life events such as the onset of a pension or departure from a career job (Gustman and Steinmeier 2000). We eschew these notions of retirement because two people receiving the same pension or having left a career job at the same age will both be “retired” even though one person may continue to be active in the labor force and the other out of the labor force. Another concept of retirement entails being out of the labor force after some arbitrary age (such as age 65 or age 67); but that concept of retirement leads to another concept of “re-retirement” if a person were to reenter the labor force and then subsequently leave the labor force and a possible “re-re-retirement” in the future if the labor force entry/exit sequence were to be repeated. In other words, labor force inactivity may be permanent or temporary. In our view, age at retirement is not fixed but rather a random variable. In addition, things change over time. For example, since 2000, labor force participation rates and the age at which people claim Social Security retirement benefits have increased. From 2000–2015, participation rates rose from 30% to 37% for men age 65–69; the increase was from 19% to 28% for women in the same age group. At the same time, the proportion of fully insured claiming Social Security retirement benefits at age 62 has declined substantially (Purcell 2016). Regardless of whether people have left a career job or receive a private pension (i.e., retirement markers to some), these facts indicate that the trend is to stay in the labor force at ages, and after life events, associated with what some call “retirement.”

In this article, retirement means labor force inactivity after final separation from the labor force has occurred (Skoog and Ciecka 2010). Our concept of retirement is retrospective. One cannot be sure a retirement has occurred until a person dies. While a person is alive, labor force reentry is possible even though the probability of reentry eventually approaches zero because of advanced age and morbidity. We view this definition of retirement as desirable because some people may be out of the labor force and think of themselves as “retired” but may reenter the labor force at some future time. Others, who do not intend to leave the labor force, may withdraw from the labor force in the future for any number of reasons and may return or perhaps never return to the labor force. In other words, intentions notwithstanding, future labor force status may change.

A person may leave the labor force, remain outside the labor force for several years, reenter for a short period of time, and then die while active. Under our construction, this person accumulates several years of inactivity but zero years in retirement. Consider the following example. A 60-year-old leaves the labor
force and may consider himself or herself "retired." Suppose this person remains inactive for the next 15 years; but returns to the labor force at age 75 and dies the day after returning to labor market activity. Although it might be intuitive to say that this person had been retired for 15 years and then died, we count zero years in retirement and 15 years of inactivity from ages 60 to 75. In this extreme case, the ratio of our retirement years to inactive time is 0 ≠ 0/15; and, for this reason, it might be argued that our measure of years in retirement is downward biased. On the other hand, we believe it is worthwhile to measure both years of inactivity and years in retirement as separate entities, with years in retirement being a subset of years of inactivity counted after final separation from the labor force has occurred. The type of extreme case in the foregoing example notwithstanding, we find that 60-year-old men and women will spend about 88%–97% of their future inactive years in an uninterrupted block of time that we call retirement. The remainder of their inactive time occurs intermittently over a period of years in which there are labor force exits and entries.

### 2.2. Notation

Let $x$ denote exact age, $a$ denotes active in the labor force, $i$ denotes inactivity, $d$ is the death state, $BA$ is the beginning age for labor market activity, and $TA$ (terminal age) is the youngest exact age at which everyone has died. Discrete random variables $RV$ measure years of activity ($YA$), years of inactivity ($YI$), years to final separation from the labor force ($YFS$), and years in retirement ($YIR$). Let $RV = YA, YI, YFS, YIR, m \in \{a, i\}$ and $n \in \{a, i\}$. The notation $RV_{x,m,n} = y$ indicates that the value of the random variable is $y$ years for a person exact age $x$ who had labor force status $m$ at age $x+1$ and has status $n$ at age $x$. This value of the random variable occurs with probability $p_{RV}(x, m, n, y)$. For example, $YI_{x,m,n} = y$, means a person exact age $x$ who had labor force status $m$ at age $x+1$ and status $n$ at age $x$ will accumulate $y$ future years of inactivity starting from age $x$ with probability $p_{YI}(x, m, n, y)$. Similar interpretations hold for $YA_{x,m,n} = y$ years of labor force activity, $YFS_{x,m,n} = y$ years to final separation from the labor force, and $YIR_{x,m,n} = y$ years in retirement.

The random variable $YAL_x$ measures additional years of life from age $x$, and $YAL_x = y$ with probability $p_{YAL}(x, y)$. We assume that $YAL_x$ does not depend on labor force state. However, $YAL_x = YA_{x,m,n} + YI_{x,m,n}$ because activity and inactivity cannot occur simultaneously but together they account for all future years of life. Similarly, years to final separation from the labor force plus years in retirement sum to additional years of life, that is, $YAL_x = YFS_{x,m,n} + YIR_{x,m,n}$.

The following table contains a summary of our notation. For each random variable, the second and third columns of the table define the probability mass function (pmf) for the random variable in the first column with support values in the last column. Implicit in this notation is the general Markov assumption that labor force states are conditionally independent going back more than two years. We require that $0 \leq p_{RV}(x, m, n, y) \leq 1$ and $\sum_y p_{RV}(x, m, n, y) = 1$ for all $RV = YA, YI, YFS, YIR$; and $0 \leq p_{YAL}(x, y) \leq 1$ and $\sum_y p_{YAL}(x, y) = 1$ for $YAL$.

| $RV$ | Value of $RV$ | Probability of $RV$ | Support for $RV$ |
|------|---------------|---------------------|-----------------|
| $YA_{x,m,a}$ | $YA_{x,m,a} = y$ | $p_{YA}(x, m, a, y)$ | $y = 0, 1, 2, 5, \ldots$ |
| $YA_{x,m,i}$ | $YA_{x,m,i} = y$ | $p_{YA}(x, m, i, y)$ | $y = 0, 1, 2, 5, \ldots$ |
| $YI_{x,m,a}$ | $YI_{x,m,a} = y$ | $p_{YI}(x, m, a, y)$ | $y = 0, 1, 2, 5, \ldots$ |
| $YI_{x,m,i}$ | $YI_{x,m,i} = y$ | $p_{YI}(x, m, i, y)$ | $y = 0, 1, 2, 5, \ldots$ |
| $YFS_{x,m,a}$ | $YFS_{x,m,a} = y$ | $p_{YFS}(x, m, a, y)$ | $y = 0, 1, 2, 5, \ldots$ |
| $YFS_{x,m,i}$ | $YFS_{x,m,i} = y$ | $p_{YFS}(x, m, i, y)$ | $y = 0, 1, 2, 5, \ldots$ |
| $YIR_{x,m,a}$ | $YIR_{x,m,a} = y$ | $p_{YIR}(x, m, a, y)$ | $y = 0, 1, 2, 5, \ldots$ |
| $YIR_{x,m,i}$ | $YIR_{x,m,i} = y$ | $p_{YIR}(x, m, i, y)$ | $y = 0, 1, 2, 5, \ldots$ |
| $YAL_x$ | $YAL_x = y$ | $p_{YAL}(x, y)$ | $y = 0, 1, 2, 5, \ldots$ |

#### 2.3. Counting Conventions

The following diagram illustrates counting conventions for $YA$, $YI$, $YFS$, and $YIR$.

![Labor force state notation diagram](image-url)

In this diagram, labor force state notation appears below exact ages. For example, $ia$ appears directly below exact age $x + 1$. The notation means that a person was active at age $x + 1$ but inactive one year earlier at age $x$. Beginning with age $x$, the diagram depicts a labor force path of a person who was inactive at age $x$ and also inactive at age $x − 1$, active at age $x + 1$, active at age $x + 2$, active at age $x + 3$, inactive at age $x + 4$, and dead by age $x + 5$. Assume that transitions occur at mid points between exact ages. A transition from inactivity to activity occurs at age $x + 5$, another transition occurs from activity to inactivity at $x + 3.5$, and a final transition to death occurs at age $x + 4.5$. Starting from state $ii$ at age $x$, we count 3 years of activity from $x + 5$ to $x + 3.5$, and 1.5 years of inactivity consisting of .5 years from $x$ to $x + .5$ plus 1 year of inactivity from $x + 3.5$ to $x + 4.5$. Thus, $YA_{x,i,i} = 3.0$ and $YI_{x,i,i} = 1.5$. We count 3.5 years to final separation from the labor force which occurs at $x + 3.5$ and 1 year in retirement from $x + 3.5$ to death at age $x + 4.5$; thus, $YFS_{x,i,i} = 3.5$ and $YIR_{x,i,i} = 1.0$. Finally, $YAL_x = 4.5$ is decomposed into 3 years of activity plus 1.5 years of inactivity; but $YAL_x$ also is decomposed into 3.5 years to final separation from the labor force plus 1.0 years in retirement.

Probability mass functions for random variables are defined recursively. Recursions are backwards in the sense that they begin at age $TA−1$ and move to successively younger ages. They transform probabilities at age $x + 1$ into probabilities at age $x$. In principle, it is possible to work in the other direction, i.e., start at age $x$ and trace all possible activity/inactivity paths a person could follow and eventually end with death. However, this is computationally impractical because there are $2^{TA−x} − 1$ such paths. For example, we assume that $TA = 111$ and, if age $x = 30$, there are $2^{111−30} − 1 = 2.41 \times 10^{50}$ paths for a person active at age 30 and an equal number for an inactive 30 year old. Our backward recursions accomplish the same task as tracing all possible labor force paths a person might follow but do so in an efficient manner in which probability mass functions calculate in a few seconds on a typical desktop or laptop computer. Recursive formulae defining probability mass functions are given in the Appendix (available in the online supplementary information). (For related work, but only for years of activity, see Cushing and Rosenbaum 2011 and Skoog and Cieck 2012.)
3. Data and Estimation of Transition Probabilities

Let \( m_{x} \) denote transition probabilities for a person who had labor force status \( m \) at age \( x-1 \) and has status \( n \) at age \( x \) and who will be \( a, i, \) or \( d \), respectively, at age \( x + 1 \). We estimate transition probabilities with data from The Survey of Income and Program Participation (SIPP) from Panels 2001, 2004, and 2008 (using SIPP Utilities 2013 and Data Ferret [2008] application). SIPP is a longitudinal survey administered by the U.S. Census Bureau (2008). The main objective of SIPP is to gather information about income and program participation of households and individuals in the United States. SIPP consists of panels created from a nationally representative sample of households in the civilian, non-institutionalized U.S. population. Panels are divided into four approximately equal subgroups, called rotation groups; and one rotation group is surveyed each month. A wave refers to the 4-month period required to interview every individual in a panel. Since a panel typically lasts 4 years, there are generally 12 waves per panel. This design results in an individual typically being interviewed once every 4 months during the life of a panel and a total of 12 interviews with each individual over the 4-year span of the panel.

SIPP is divided into core and topical modules. The core module includes questions asked during every interview and covers demographic characteristics, labor force participation, and program participation. Topical modules contain detailed information on specific subjects or measure characteristics that are not expected to change frequently, such as disability or marital status. An interviewee provides data on each of the preceding four months (the reference period for an interview) for most core items, but various topical module questions are asked less frequently. An interview begins with the same set of core questions and ends with a different set of topical interview questions that reappear periodically. (See the online supplementary material for more information about our SIPP sampling procedure.)

We utilize data from the 2001, 2004, and 2008 SIPP panels. The following table shows panel life (defined by dates of first and last interviews) and usable sample size.

| Panel       | Panel Duration           | Sample Size |
|-------------|--------------------------|-------------|
| 2001        | February 2001–January 2004 | 36,989      |
| 2004        | February 2004–January 2008 | 54,759      |
| 2008        | September 2008–December 2013 | 50,558      |

The National Bureau of Economic Research (NBER 2014) uses March 2001 as the date of a business cycle peak in the U.S., November 2001 as a trough, December 2007 as the next peak, June 2009 as the next trough, followed by a weak recovery that persisted into 2013. Using NBER business cycle dating, our SIPP data set includes two peak-to-trough periods and the slow recovery of the U.S. economy since mid-2009. There is some manifestation of the 2001 contraction in our transition data, but our second yearly observation and the third observation from the 2001 SIPP panel are in expansion months. Data from the 2004 SIPP panel include all expansion months. The 2008 SIPP panel beginning in late 2008 does include the Great Recession but our second yearly observation and the third observation in that panel are expansion months.

The following table contains participation rates computed with SIPP data and BLS participation rates averaged over 2001–2012. Although we do not use participation rates in this article, we report them in order to show that SIPP data yield participation rates that are quite similar to participation rates computed from BLS participation rates for the U.S. economy.

| Age     | SIPP | BLS† | SIPP | BLS‡ |
|---------|------|------|------|------|
| 25–29   | 0.91 | 0.77 | 0.77 | 0.74 |
| 30–34   | 0.94 | 0.92 | 0.76 | 0.74 |
| 35–39   | 0.93 | 0.93 | 0.77 | 0.75 |
| 40–44   | 0.92 | 0.91 | 0.79 | 0.77 |
| 45–49   | 0.90 | 0.89 | 0.79 | 0.77 |
| 50–54   | 0.87 | 0.86 | 0.76 | 0.74 |
| 55–59   | 0.78 | 0.78 | 0.67 | 0.66 |
| 60–64   | 0.57 | 0.59 | 0.46 | 0.48 |
| 65–69   | 0.34 | 0.35 | 0.24 | 0.25 |
| 70–74   | 0.22 | 0.21 | 0.12 | 0.13 |

†Computed with SIPP data for 2001–2012 using person weights.
‡Computed from BLS reported participation rates for 2001–2012 using U.S. non-institutional population.

We assume, as with life tables, that transitions occur at the midpoint between exact ages, which implies non-zero probabilities may only occur at integers and half-integers. To estimate transition probabilities, we let \( m_{x} \) denote the count of SIPP respondents who had labor force status \( m \) at age \( x-1 \) and status \( n \) at age \( x \). The number active at age \( x+1 \) is \( m_{x}+n_{x} \), and \( n_{x} \) denotes the number inactive. Then, \( m_{x} = m_{x} + n_{x} \). We estimate transition probabilities \( m_{x} \) using

\[
\begin{align*}
\hat{m}_{x} & = \left[ \frac{m_{x} + n_{x}}{m_{x} + n_{x}} \right] \left( 1 - \hat{p}_{x} \right) \\
\hat{n}_{x} & = \left[ \frac{m_{x} + n_{x}}{m_{x} + n_{x}} \right] \left( 1 - \hat{p}_{x} \right)
\end{align*}
\]

SIPP records ages as integers; but people are, on average, one-half year older than their reported integer age. That is, those reporting their age as \( x \) are on average age \( x + 0.5 \) and those reporting their age as \( x-1 \) are on average age \( x-0.5 \). The foregoing formulae center \( \hat{m}_{x} \) and \( \hat{n}_{x} \) on exact age \( x \). The survival probability adjustment factor \( (1 - \hat{p}_{x}) \) is from U.S. life tables (Arias 2010).

4. Empirical Results and Policy

Separate probability mass functions were estimated for all random variables at ages 30, 40, 50, and 60 for states \( aa, ia, ai, \) and \( ii \) for men and women. We illustrate a few of these pmfs in Figures 1 and 2 for \( Y_{1} \) and \( Y_{FS} \), respectively, for 40-year-old men who were in state \( aa \) and compare them to 40-year-old men in state \( ii \). The most striking feature in Figure 1 is its large spike of capturing the probability a 40-year-old man in state \( aa \) will never be inactive throughout the remainder of his life. That is, he will be in the labor force throughout the remainder of his life and will never retire.

On the other hand, the probability is zero that a 40-year-old man in state \( ii \) will have no future inactivity, and the
probability that he will be inactive for only 0.5 of a year is only $p_{YFFS}(40, i, 0, 0.5) = 0.015$. However, there are consistently higher probabilities (relative to the $aa$ state) that he will be inactive for 25 or more years throughout the remainder of his life.

In Figure 2, for $YFS$, $p_{YFYS}(40, i, i, 0) = 0.15$ measures the probability that men in state $ii$ will have made their final separation from the labor force, never to reenter after age 40. That is, the probability is 0.15 that men have retired if they are inactive at age 39 and again inactive at age 40. The probability that a 40-year-old man in state $aa$ will make a final separation from the labor force in 0.5 years, the earliest an active person can exit the labor force, only is $p_{YFFS}(40, a, a, 0.5) < 0.005$. Except for the probability spike at $YFS = 0$ for $ii$ men, $aa$ men have higher probabilities of labor force separation in the next 35 years; but the cumulative probability of labor force separation for $ii$ men remains higher than for $aa$ men, thereby showing greater activity for $aa$ men. Figures 1 and 2 show men in state $aa$ have less future labor force inactivity and separate later from the labor force than their $ii$ counterparts.

Tables 1 and 2 contain expectations for men and women by labor force state at ages 30, 40, 50, and 60. We use 50-year-old men as an example to read our tables. In Table 1, males active at age 49 and active at age 50 spend about one and a half more years of their future life outside of the labor force than in the labor force; they have expected years of activity and inactivity of $E[Y_{A50,a,a}] = 13.6$ and $E[Y_{I50,a,a}] = 15.2$. Their sum is life expectancy $E[Y_{AL50}] = 28.8$ years. The picture differs greatly for males inactive at age 49 and inactive at age 50. Their life expectancy (also 28.8 years) divides into expected years of labor force activity $E[Y_{A50,i,i}] = 5.7$ and expected years of inactivity $E[Y_{I50,i,i}] = 23.1$. That is, 50-year-old men in state $ii$ are inactive about eight additional years (and corresponding fewer years in labor force activity) than their $aa$ counterparts.

Continuing with our example for 50-year-old men, Table 1 contains the decomposition of life expectancy into years until final separation from the labor force occurs and years in retirement. Here, we have expected years to final separation from the labor force $E[Y_{FYS50,a,a}] = 15.6$ and expected years in retirement $E[Y_{IR50,a,a}] = 13.2$ for men active at ages 49 and 50, but $E[Y_{FYS50,i,i}] = 11.2$ and $E[Y_{IR50,i,i}] = 17.6$ for those inactive at ages 49 and 50. That is, 50-year-old men in state $ii$ are in retirement for 4.4 additional years than their $aa$ counterparts. Remembering that the $YIR$ random variable counts zero years in retirement if a person dies while active, expected years in retirement for those who actually retire (conditional years in retirement) is $E[Y_{IR50,a,a}] = 16.3$ for 50-year-old men in state $aa$ and $E[Y_{IR50,i,i}] = 19.7$ for those in state $ii$. Expected age at retirement (current age plus conditional years to retirement for those who actually retire) differs by 5.5 years for 50-year-old men in states $aa$ and $ii$, i.e., age 50 + $E[Y_{CYIR50,a,a}] = 65.8$ in comparison to age 50 + $E[Y_{CYIR50,i,i}] = 60.3$.

The probability that 50-year-old men in state $ii$ are retired or will retire (at some point in the future) is $P(Y_{IR50,i,i} > 0) = 0.89$ but only 0.81 for those in state $aa$ who are more likely to die while active and therefore never retire; 50-year-old men in state $ii$ will spend 76% of their inactive years in retirement, whereas retirement years comprise 87% of future inactive years for men in state $aa$ (reflecting the idea that once men in state $aa$ leave the labor force, they are more likely to stay out of the labor force than $ii$ men); and 80% of life expectancy for 50-year-old men in state $ii$ is spent in inactivity, whereas men in state $aa$ are inactive for only 53% of their life expectancy. The picture for 50-year-old $aa$ men relative to $ii$ men is clear. In the future, they are in the labor force longer (about 8 years), they separate from the labor force later (about 4.5 years), they retire later (about 5.5 years), and a much smaller fraction of their life will be out of the labor force.

We have two main policy insights. First, consider a person who is inactive at age $x = 1$. Excluding a transition to the death state, the labor force path for this person leads either to activity or inactivity at age $x$. This transition is important for expected future labor force activity and retirement behavior. For example, consider a woman who is inactive at age 49. If a transition to activity were to occur, Table 2 shows expected future years of activity $E[Y_{A50,i,i}] = 11.4$ years; but only $E[Y_{A50,i,i}] = 5.3$ years if the transition had been to inactivity. Similarly, expected age at retirement would be 50 + $E[Y_{CYTR50,i,i}] = 63.5$ if the transition were to activity but only age 50 + $E[Y_{CYTR50,i,i}] = 59.5$ if a transition to inactivity occurs. This suggests that once inactivity occurs, it is important to have policies and programs that prevent inactivity to extend for another year—the quicker a person can be brought back into the labor force, the better. Two consecutive years of inactivity for a 50-year-old woman translates into 6.1 fewer years of future labor force activity and earlier retirement age of 4.0 years. An analogy to a health issue may not be inappropriate in this regard in that the outcome is usually better when a health issue is addressed sooner rather than later.
Second, policies that can move people from inactivity to activity are more important when measured in terms of "years-of-activity gained" for the middle-aged (people 40 and 50) than younger people (age 30) or older people (age 60). As an example, consider 30-year-old men and 60-year-old men. Table 1 shows expected years of activity $E[Y_{A30,i,a}] = 30.2$ and $E[Y_{A30,i,i}] = 24.8$ – a difference of 5.4 years in future lifetime activity when comparing men in states $ia$ and $ii$ at age 30. The gain in years of activity is similar for 60-year-old men: $E[Y_{A60,i,a}] = 7.3$ and $E[Y_{A60,i,i}] = 2.0$; a gain of 5.3 years in future lifetime activity when comparing states $ia$ and $ii$. To be sure, these are large gains; and they are an especially large fraction of 60-year-old person's remaining life. However, time in the labor force increases even more for 40 and 50-year-old men when comparing $ia$ and $ii$ states. At age 40, $E[Y_{A40,i,a}] = 21.0$ in comparison to $E[Y_{A40,i,i}] = 13.2$—a gain of 7.8 years. Similarly, at age 50, there is a gain of 7.1 years when moving from $ia$ and $ii$ states with $E[Y_{A50,i,a}] = 12.8$ and $E[Y_{A50,i,i}] = 5.7$.

Similar gains in years-of-activity accrue to women when moving from $ia$ and $ii$ states as well, but gains are about a year less at all ages relative to men. This suggests that resources devoted to moving people from inactivity to activity may be more efficient when devoted to people in middle age rather than to younger people. (However, costs of moving people in middle age back into labor force activity may be higher than for younger people.) Although more current inactivity portends more future inactivity across all ages, younger people may be better able to sort out inactivity issues by themselves. Older people may not gain as much in future activity because mortality, morbidity, and other life decisions leading to retirement may affect them more than those in the middle age. On the other hand, those in the middle age have long life expectancies and are relatively free of

Table 1. Labor force characteristics for men

| Characteristics | $aa$ | $ia$ | $ai$ | $ii$ |
|-----------------|------|------|------|------|
| **Age 30**      |      |      |      |      |
| 1. $E[Y_{A}]$   | 46.9 | 46.9 | 46.9 | 46.9 |
| 2. $E[Y_{A}]$   | 30.7 | 30.2 | 27.4 | 24.8 |
| 3. $E[Y_{Y}]$   | 16.2 | 16.7 | 19.5 | 22.1 |
| 4. $E[Y_{FS}]$  | 33.8 | 33.7 | 33.4 | 33.0 |
| 5. $E[Y_{YIR}]$ | 13.1 | 13.2 | 13.5 | 13.9 |
| 6. $E[Y_{CTR}]$ | 16.9 | 16.9 | 17.2 | 17.5 |
| 7. $E[Y_{CYTR}]$| 34.8 | 34.7 | 34.1 | 33.4 |
| 8. $x + E[Y_{CYTR}]$ | 64.8 | 64.7 | 64.1 | 63.4 |
| 9. $Pr[Y_{IR} > 0]$ | 0.78 | 0.78 | 0.79 | 0.80 |
| 10. $E[Y_{YIR}/E[Y_{Y}]$ | 0.81 | 0.79 | 0.69 | 0.63 |
| 11. $E[Y_{Y}]$ | 0.35 | 0.36 | 0.42 | 0.47 |
| **Age 40**      |      |      |      |      |
| 1. $E[Y_{A}]$   | 37.6 | 37.6 | 37.6 | 37.6 |
| 2. $E[Y_{A}]$   | 21.7 | 21.9 | 16.8 | 13.2 |
| 3. $E[Y_{Y}]$   | 15.9 | 16.6 | 20.8 | 24.4 |
| 4. $E[Y_{FS}]$  | 24.4 | 24.1 | 22.8 | 21.5 |
| 5. $E[Y_{YIR}]$ | 13.2 | 13.5 | 14.8 | 16.1 |
| 6. $E[Y_{CTR}]$ | 16.8 | 17.0 | 18.1 | 19.2 |
| 7. $E[Y_{CYTR}]$| 25.0 | 24.7 | 22.7 | 20.9 |
| 8. $x + E[Y_{CYTR}]$ | 65.0 | 64.7 | 62.7 | 60.9 |
| 9. $Pr[Y_{IR} > 0]$ | 0.79 | 0.79 | 0.82 | 0.84 |
| 10. $E[Y_{YIR}/E[Y_{Y}]$ | 0.83 | 0.81 | 0.71 | 0.72 |
| 11. $E[Y_{Y}]$ | 0.42 | 0.44 | 0.55 | 0.60 |
| **Age 50**      |      |      |      |      |
| 1. $E[Y_{A}]$   | 28.8 | 28.8 | 28.8 | 28.8 |
| 2. $E[Y_{A}]$   | 13.6 | 12.8 | 8.2  | 5.7  |
| 3. $E[Y_{Y}]$   | 15.2 | 16.0 | 20.6 | 23.1 |
| 4. $E[Y_{FS}]$  | 15.6 | 15.2 | 12.7 | 11.2 |
| 5. $E[Y_{YIR}]$ | 13.2 | 13.6 | 16.1 | 17.6 |
| 6. $E[Y_{CTR}]$ | 16.8 | 16.6 | 18.6 | 19.7 |
| 7. $E[Y_{CYTR}]$| 15.8 | 15.2 | 12.0 | 10.3 |
| 8. $x + E[Y_{CYTR}]$ | 65.8 | 65.2 | 62.0 | 60.3 |
| 9. $Pr[Y_{IR} > 0]$ | 0.81 | 0.82 | 0.87 | 0.89 |
| 10. $E[Y_{YIR}/E[Y_{Y}]$ | 0.87 | 0.85 | 0.78 | 0.76 |
| 11. $E[Y_{Y}]$ | 0.53 | 0.56 | 0.72 | 0.80 |
| **Age 60**      |      |      |      |      |
| 1. $E[Y_{A}]$   | 20.7 | 20.7 | 20.7 | 20.7 |
| 2. $E[Y_{A}]$   | 8.0  | 7.3  | 3.2  | 2.0  |
| 3. $E[Y_{Y}]$   | 12.7 | 13.4 | 17.5 | 18.7 |
| 4. $E[Y_{FS}]$  | 9.0  | 8.4  | 5.3  | 4.2  |
| 5. $E[Y_{YIR}]$ | 11.7 | 12.3 | 15.4 | 16.5 |
| 6. $E[Y_{CTR}]$ | 14.3 | 14.7 | 16.8 | 17.5 |
| 7. $E[Y_{CYTR}]$| 8.9  | 8.3  | 4.7  | 3.6  |
| 8. $x + E[Y_{CYTR}]$ | 68.9 | 68.3 | 64.7 | 63.6 |
| 9. $Pr[Y_{IR} > 0]$ | 0.82 | 0.83 | 0.92 | 0.94 |
| 10. $E[Y_{YIR}/E[Y_{Y}]$ | 0.92 | 0.92 | 0.88 | 0.88 |
| 11. $E[Y_{Y}]$ | 0.61 | 0.65 | 0.85 | 0.90 |
morbidty problems for several years; they thereby may gain the most from policies that can move them from inactive to active in the labor force.

The foregoing estimates depend on a person’s age, gender, and recent labor force history. Many other factors exert strong influences on inactivity and retirement such as macro-economic conditions and private and public retirement policies (e.g., the Social Security tax penalty which costs $1 in Social Security benefits for every $2 in earned income at ages 62, 63, and 64). On a more individual level, important factors influencing labor force inactivity and retirement decisions include a person’s wealth, education, possession of modern-day technological skills, health (including disability and health of family members), and age of spouse and ages of children. Although we do not deal with the underlying causes of inactivity and retirement decisions beyond age, gender, and recent labor force history, we recognize the importance of other factors. Our goal only was to measure the current relation between labor force states and future inactivity and retirement but doing so with the probability distributions of all of the random variables we study and within the context of an auto-regressive model of order two. In so doing, this article attempts to move the discussion beyond participation rates (which tell us little about the expected values of the random variables we study) to actually measuring the relation between current of labor force states and expectations for lost years of future labor force activity and additional years spent in retirement. Better understanding of the magnitude of foregone years of activity may help focus attention on finding appropriate polices to move people back into the labor force by addressing the underlying causes of inactivity.

Table 2. Labor force characteristics for women

| Characteristics | State | aa | ia | ai | ii |
|-----------------|-------|----|----|----|----|
| Age 30          | 1. E[YAL] | 51.3 | 51.3 | 51.3 | 51.3 |
|                 | 2. E[YA] | 26.4 | 25.6 | 22.6 | 21.4 |
|                 | 3. E[YI] | 24.9 | 25.7 | 28.7 | 29.9 |
|                 | 4. E[YFS] | 31.7 | 31.6 | 31.3 | 31.1 |
|                 | 5. E[YIR] | 19.6 | 19.7 | 20.0 | 20.2 |
|                 | 6. E[CYIR] | 22.0 | 22.1 | 22.3 | 22.5 |
|                 | 7. E[CYTR] | 32.0 | 31.8 | 31.4 | 31.2 |
|                 | 8. x + E[CYTR] | 62.6 | 61.8 | 61.4 | 61.2 |
|                 | 9. Pr[YIR > 0] | 0.89 | 0.89 | 0.89 | 0.90 |
|                 | 10. E[YIR]/E[YI] | 0.79 | 0.77 | 0.70 | 0.68 |
|                 | 11. E[YI]/E[YAL] | 0.49 | 0.50 | 0.56 | 0.58 |
| Age 40          | 1. E[YAL] | 41.7 | 41.7 | 41.7 | 41.7 |
|                 | 2. E[YA] | 19.1 | 18.1 | 13.9 | 12.0 |
|                 | 3. E[YI] | 22.6 | 23.6 | 27.8 | 29.7 |
|                 | 4. E[YFS] | 22.4 | 22.1 | 20.8 | 20.1 |
|                 | 5. E[YIR] | 19.3 | 19.6 | 20.9 | 21.6 |
|                 | 6. E[CYIR] | 21.6 | 21.9 | 23.0 | 23.5 |
|                 | 7. E[CYTR] | 22.6 | 22.2 | 20.6 | 19.8 |
|                 | 8. x + E[CYTR] | 62.6 | 62.2 | 60.6 | 59.8 |
|                 | 9. Pr[YIR > 0] | 0.89 | 0.90 | 0.91 | 0.92 |
|                 | 10. E[YIR]/E[YI] | 0.85 | 0.83 | 0.75 | 0.73 |
|                 | 11. E[YI]/E[YAL] | 0.54 | 0.57 | 0.67 | 0.71 |
| Age 50          | 1. E[YAL] | 32.5 | 32.5 | 32.5 | 32.5 |
|                 | 2. E[YA] | 12.4 | 11.4 | 7.0 | 5.3 |
|                 | 3. E[YI] | 20.1 | 21.1 | 25.5 | 27.2 |
|                 | 4. E[YFS] | 14.1 | 13.5 | 11.0 | 9.9 |
|                 | 5. E[YIR] | 18.4 | 19.0 | 21.5 | 22.6 |
|                 | 6. E[CYIR] | 20.4 | 20.9 | 23.0 | 23.9 |
|                 | 7. E[CYTR] | 14.1 | 13.5 | 10.6 | 9.5 |
|                 | 8. x + E[CYTR] | 64.1 | 63.5 | 60.6 | 59.5 |
|                 | 9. Pr[YIR > 0] | 0.90 | 0.90 | 0.93 | 0.94 |
|                 | 10. E[YIR]/E[YI] | 0.92 | 0.90 | 0.84 | 0.83 |
|                 | 11. E[YI]/E[YAL] | 0.62 | 0.65 | 0.78 | 0.84 |
| Age 60          | 1. E[YAL] | 23.8 | 23.8 | 23.8 | 23.8 |
|                 | 2. E[YA] | 7.3 | 6.4 | 2.5 | 1.5 |
|                 | 3. E[YI] | 16.5 | 17.4 | 21.3 | 22.3 |
|                 | 4. E[YFS] | 7.8 | 7.0 | 3.8 | 2.8 |
|                 | 5. E[YIR] | 16.0 | 16.8 | 20.0 | 21.0 |
|                 | 6. E[CYIR] | 17.7 | 18.3 | 20.4 | 21.5 |
|                 | 7. E[CYTR] | 7.8 | 6.9 | 3.5 | 2.6 |
|                 | 8. x + E[CYTR] | 67.8 | 66.9 | 63.5 | 62.6 |
|                 | 9. Pr[YIR > 0] | 0.90 | 0.91 | 0.96 | 0.98 |
|                 | 10. E[YIR]/E[YI] | 0.97 | 0.97 | 0.94 | 0.94 |
|                 | 11. E[YI]/E[YAL] | 0.69 | 0.73 | 0.89 | 0.94 |
5. Conclusion

We found probability mass functions (fully specified in the Appendix, available in the online supplemental information) for random variables measuring years of future activity, years of future inactivity, years to final separation from the labor force, years in retirement, conditional years in retirement, and years to retirement. These probability mass functions enabled us to decompose life expectancy into labor force related states at ages 30, 40, 50, and 60. We conditioned on two years of most recent labor force history. We showed that states \( aa, ia, ai, ii \) play important roles in influencing expectations of labor force random variables and that their roles grow in importance as people enter middle age. The effects of two successive years of inactivity are large. These conclusions argue for pursuing policies that would quickly move people out of inactivity and targeting middle-aged inactives if the goal is to increase labor force participation and extend time to retirement.

Acknowledgments

We wish to thank an anonymous referee and the editor for many helpful insights and comments that greatly improved the article. We also thank Sean O’Farrell for excellent work retrieving SIPP data.

Supplementary Materials

Supplemental data for this article can be accessed on the publisher’s website.

References

Arias, E. (2010), *National Vital Statistics Reports, United States Life Tables 2006*, Hyattsville: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention. 58, Table 2 and Table 3. [4]

Cushing, M. J., and Rosenbaum, D. I. (2011), “Higher Order Markov Estimates of Worklife: Comparing SIPP and CPS Results,” Paper presented at Western Economics Association, San Diego, CA. [3]

Gendell, M., and Siegel, J. S. (1992), “Trends in Retirement Age by Sex, 1950–2005,” *Monthly Labor Review*, 115, 22–29. [2]

Gustman, A. L., and Steinmeier, T. L., (2000), “Retirement Outcomes in the Health and Retirement Study,” *Social Security Bulletin*, 63, 57–71. [2]

Maestas, N. (2010), “Back to Work, Expectations and Realizations of Work after Retirement,” *Journal of Human Resources*, 45, 718–748. [2]

National Bureau of Economic Research web site http://www.nber.org/cycles.html, accessed January 2014. [4]

Purcell, P. (2016), “Employment at Older Ages and Social Security Benefit Claiming,” *Social Security Bulletin*, 76, 1–17. [2]

Reimers, C. (1976), “Is the Average Age at Retirement Changing?” *Journal of the American Statistical Association*, 71, 552–558. [2]

Ruhm, C. (1990), “Bridge Jobs and Partial Retirement,” *Journal of Labor Economics*, 8, 482–501. [2]

Schoen, R., and Land, K. (1979), “A General Algorithm for Estimating a Markov-Generated Increment–Decrement Life Table with Applications to Marital-Status Patterns,” *Journal of the American Statistical Association*, 74, 761–776. [2]

SIPP Utilities (2013), Unicon Research Corporation, 1640 Fifth Street, Suite 100 Santa Monica, CA 90401. [4]

Skoog, G., and Ciecka, J. (2002), “Probability Mass Functions for Additional Years of Labor Market Activity Induced by the Markov (Increment–Decrement) Model,” *Economics Letters*, 77, 425–431. [2]

——— (2010), “Measuring Years of Inactivity, Years in Retirement, Time to Retirement, and Age at Retirement within the Markov Model,” *Demography*, 47, 609–628. [2]

——— (2012), “An Autoregressive Model of Order Two for Worklife Expectancies and Other Labor Force Characteristics with an Application to Major League Baseball Hitters,” *Journal of Legal Economics*, 18, 47–78. [3]

Skoog, G., Ciecka, J., and Krueger, K. (2011), “The Markov Process Model of Labor Force Activity: Extended Tables of Central Tendency, Shape, Percentile Points, and Bootstrap Standard Errors,” *Journal of Forensic Economics*, 22, 165–229. [2]

Smith, S. (1982), *Tables of Working Life: The Increment-Decrement Model*, *Bulletin 2135*, Washington, DC: Bureau of Labor Statistics. [2]

——— (1986), *Worklife Estimates: Effects of Race and Education, and Bulletin 2254*, Washington, DC: Bureau of Labor Statistics. [2]

U.S. Census Bureau (2008) Data Ferrett, 2008 Survey of Income and Program Participation reference. [4]