University of Michigan Retirement Research Center

Working Paper
WP 2003-044

Key equations in the Tuljapurkar-Lee model of the Social Security system

Ryan D. Edwards, Ronald D. Lee, Michael W. Anderson,
Shripad Tuljapurkar, and Carl Boe

MRRC Project #: UM03-Q1

“Key equations in the Tuljapurkar-Lee model of the Social Security system”

Ryan D. Edwards,

Ronald D. Lee, PhD
University of California at Berkeley

Michael W. Anderson, PhD

Shripad Tuljapurkar, PhD
Stanford University

Carl Boe
Stanford University

March 2003

Michigan Retirement Research Center
University of Michigan
P.O. Box 1248
Ann Arbor, MI 48104

Acknowledgements

This work was supported by a grant from the Social Security Administration through the Michigan Retirement Research Center (Grant # 10-P-98358-5). The opinions and conclusions are solely those of the authors and should not be considered as representing the opinions or policy of the Social Security Administration or any agency of the Federal Government.

Key equations in the Tuljapurkar-Lee model of the Social Security system

Shripad Tuljapurkar, and Carl Boe

Abstract

We present stochastic forecasts of the Social Security trust fund by modeling key demographic and economic variables as historical time series, and using the fitted models to generate computer simulations of future fund performance. We evaluate several plans for achieving long-term solvency by raising the normal retirement age (NRA), increasing taxes, or investing some portion of the fund in the stock market. Stochastic population trajectories by age and sex are generated using the Lee-Carter and Lee- Tuljapurkar mortality and fertility models. Interest rates, wage growth and equities returns are modeled as vector autoregressive processes. With the exception of mortality, central tendencies are constrained to the Intermediate assumptions of the 2002 Trustees Report. Combining population forecasts with forecasted per-capita tax and benefit profiles by age and sex, we obtain inflows to and outflows from the fund over time, resulting in stochastic fund trajectories and distributions. Under current legislation, we estimate the chance of insolvency by 2038 to be 50%, although the expected fund balance stays positive until 2041. An immediate 2% increase in the payroll tax rate from 12.4% to 14.4% sustains a positive expected fund balance until 2078, with a 50% chance of solvency through 2064. Investing 60% of the fund in the S&P 500 by 2015 keeps the expected fund balance positive until 2060, with a 50% chance of solvency through 2042. An increase in the NRA to age 69 by 2024 keeps the expected fund balance positive until 2047, with a 50% chance of solvency through 2041. A combination of raising the payroll tax to 13.4%, increasing the NRA to 69 by 2024, and investing 25% of the fund in equities by 2015 keeps the expected fund balance positive past 2101 with a 50% chance of solvency through 2077.

Key equations in the Tuljapurkar-Lee model of the Social Security system

Ryan D. Edwards, Ronald D. Lee, Michael W. Anderson,
Shripad Tuljapurkar, and Carl Boe*

March 27, 2003

The Tuljapurkar-Lee (henceforth "TL") model generates a large set of Monte Carlo simulations of future outcomes in the U.S. Social Security system. Four types of key macrodemographic and macroeconomic variables are modeled as stochastic components using standard time series methods. These include age-specific mortality rates, age-specific fertility rates, the rate of growth in real covered wages per capita, and real rates of return on two classes of financial assets: the special-issue Treasury obligations in the Social Security Trust Fund, and the S&P 500 stock index.

This paper is a revision of the technical appendix in Lee and Edwards (2002), updated to reflect the modeling techniques employed during the latest round of revision to the TL model.

1 Population forecasts

1.1 Mortality

Let $m_{x,t}$ be a central death rate for age $[x, x+5)$, and time $[t, t+1)$. Suppose we have a matrix of X age specific death rates over T years. The Lee-Carter method estimates the model:

$$\log(m_{x,t}) = a_x + b_x k_t + \epsilon_{x,t} \tag{1}$$

---

*Stanford University and University of California at Berkeley. Corresponding author: Boe, `boe@demog.berkeley.edu`

using a Singular Value Decomposition (SVD) or some other appropriate method. This yields estimates of the vectors a_x, b_x, and k_t. A second stage procedure adjusts k_t so that life expectancy at birth is exactly matched by the model for each year t. Estimates for a_x and b_x are listed in Table 2.2.

We now have a time series of k_t over T years. This time series is modeled using standard Box-Jenkins methods. (Tests for covariance with the residuals from the fertility model described below showed no association, so they were modeled independently). In most applications, it is well-fitted by a random walk with drift. In the Tuljapurkar-Lee model, Lee-Carter is estimated on mortality rates by single years of age and time, with both sexes combined, from the Social Security actuarial tables from 1950 to 1997. Model estimates are

$$\begin{array}{ccccccc} k_t & = & k_{t-1} & - & 1.279864 & + & \eta_t, \\ & & & & (0.224871) & & (1.574093) \end{array} \tag{2}$$

where standard errors are in parentheses.

The fitted model for k_t can then be used to forecast k for each sex separately over the desired horizon, together with a probability distribution for each forecast year. A fixed additive offset factor is used for male and female k_t's throughout the entire forecast period. The level of the offset is set so as to match the male/female mortality differential in 2001; for males, $k_{2001} = 21.87738$, and for females, $k_{2001} = -28.359$.

Using these sex-specific forecasts of k and equation (1), probability distributions and mean or median values of $m_{x,t}$ and the implied life expectancies can be calculated, along with probability distributions. These probability distributions reflect the innovation error in k, η, along with the uncertainty of the estimate of the drift in the k process. They typically will not include the ϵ terms, nor the uncertainty in the estimates of the a_x and b_x vectors, which do not add much to the uncertainty after the first decade or two. On all of this, see Lee and Carter (1992) and Lee and Miller (2001).

1.2 Fertility

A similar approach is followed, but the fertility rates themselves, rather than their logs, are modeled. The model for age specific fertility $\mathcal{F}$ is:

$$\mathcal{F}_{x,t} = c_x + d_x f_t + \nu_{x,t}, \tag{3}$$

which is again estimated using a SVD. Estimates for c_x and d_x are listed in Table 2.2. Time series models applied to the history of fertility in the U.S.

do not provide a plausible model or forecast for fertility for various reasons, so the mean of the forecast is constrained to equal a level specified ex ante, and in practice taken to equal the ultimate level of fertility assumed by the Social Security Actuaries, currently 1.95 children per woman. The fitted time series model then provides crucial information about the variability and autocovariance of fertility. See Lee (1993) for a discussion of all these issues, and exploration of some alternative modeling strategies.

The model is fitted to fertility rates by single years of age from 1933. Several sources are used to construct these data, including Whelpton (1954), Heuser (1976, 2003), and for more recent years, vital statistics data from the National Center for Health Statistics. The NCHS data on fertility among 5-year age groups is converted to data by single years of age using interpolation techniques. The fitted fertility time series model is a constrained ARMA that takes the following form:

$$\begin{array}{ccccccccc} f_t & = & \underset{(0.0315)}{-0.0924(1-0.9600)} & + & \underset{(0.0315)}{0.9600\, f_{t-1}} & + & \underset{(0.1101)}{\nu_t} & + & \underset{(0.0959)}{0.5232\, \nu_{t-1}}, \end{array} \tag{4}$$

where the -0.0924 value is hard-wired in order to achieve a long-run TFR of 1.95. In order to start generating trajectories with these estimates, the last innovation value is required. It is given by

$$\frac{f_{2001} + 0.0924 - 0.9600(f_{2000} + 0.0924)}{0.5232} = 0.0137. \tag{5}$$

Combining the forecast of f_t with estimates of c_x and d_x using (3) yields a set of stochastic fertility projections.

1.3 Formulating population forecasts

Immigration was projected deterministically following the intermediate cost assumption of the Social Security Actuary, since it was thought better to treat it as a policy instrument than to attempt to forecast future policy. Population forecasts are constructed by setting initial conditions using the base period population age distribution from Social Security data. A single stochastic sample path is generated by drawing random numbers for the errors in the fertility and mortality equations, and thereby generating a trajectory of age specific fertility and mortality rates over the desired horizon, say 100 years. Sample paths containing a total fertility rate below 0 or greater than 4 are discarded. In remaining paths, any negative age specific birth rates are set to 0. These are combined with the deterministic immigration rates. Using well-known accounting identities, the population forecast by age group is then calculated for this single sample path. The procedure is then repeated many times, sometimes 1,000 times and sometimes 10,000 times. The frequency distributions of outcomes of interest then provide estimates of the probability distributions for these outcomes, and joint distributions can be provided in a similar way.

2 Economic projections

2.1 Productivity (growth in covered wages)

The relevant concept of productivity growth in the Social Security system is the real rate of growth in average covered wages. Although there are crucial differences between average covered wages and productivity, or total output per worker or per worker-hour, the TL model treats them as essentially interchangeable for several reasons. First, Social Security taxes and benefits both grow according to the same concept, rather than some mixture of the two. Second, it is difficult to obtain a time series of average covered wage growth, while productivity growth measures are quite abundant. Third, we believe that the variability in the two measures over time has been similar. Fourth, since we choose to assume a fixed long-run trend growth rate in average covered wages identical to that assumed by the Trustees, we believe there is little precision lost by using historical productivity series to estimate the variance structure of covered wage growth.

For modeling purposes, a demographically adjusted productivity growth series was constructed. First, an average wage profile by age and sex was calculated from the 1997 March CPS. Data on the age-sex composition of the labor force were also taken from CPS, from 1948 to the present. The effect of the changing age-sex composition of the labor force, based on these age-sex weights for wages, was then calculated for each year since 1948 and used to adjust the official measure of productivity growth in the private nonfarm business sector to remove the effect of changing demographic structure of the labor force. The adjustment made relatively little difference in general, and is discussed in greater detail in Lee and Tuljapurkar (1998).

Next, a constrained mean time series model was fit to the adjusted productivity growth series. As with fertility, the time series model provides information about the variance, autocovariance and cross covariance of the series, but not about the long run mean, which is imposed as a value of 1.1 percent. An autoregressive model of order one was found to fit the data best:

$$\begin{array}{ccccccc} g_t - 1.1 & = & 0.5327 & (g_{t-1} - 1.1) & + & \epsilon_{g,t}. \\ & & (0.1197) & & & (1.3962) \end{array} \tag{6}$$

Productivity growth g_t is expressed in percentage points.

2.2 Asset returns

The bonds held in the Social Security Trust Fund are a special Treasury Issue with a rate of return equal to an average of rates on longer term Treasury bonds. The Social Security Administration's website contains a time series of the effective interest rates on Trust Fund assets from 1940 to 2002.[1] We use this special issue rate, minus the rate of inflation as measured by the CPI-U, as our baseline real interest rate. Historical stock returns, defined as total returns on the S&P 500 Index adjusted for the reinvestment of dividends, are available over the same period from Ibbotson Associates (2002) as well as from other sources. The jump-off points for the two series are 3.8 and −15.84 percent respectively.

We fit a VAR of order three that recognizes the conjoined behavior of real bond returns, r_t, and real stock returns, s_t, subject to the assumption that they will tend to revert to their respective long-run means of 3 and 7 percent. The equations take the following form, where an asterisked variable denotes its level minus its long-run mean:

$$r_t^* = \begin{pmatrix} 1.1555 \\ -0.7993 \\ 0.4772 \end{pmatrix} \cdot \begin{pmatrix} r_{t-1}^* \\ r_{t-2}^* \\ r_{t-3}^* \end{pmatrix} + \begin{pmatrix} 0.0131 \\ -0.0165 \\ 0.0093 \end{pmatrix} \cdot \begin{pmatrix} s_{t-1}^* \\ s_{t-1}^* \\ s_{t-1}^* \end{pmatrix} + \epsilon_{r,t} \tag{7}$$

$$s_t^* = \begin{pmatrix} 1.4392 \\ -0.3591 \\ 0.0091 \end{pmatrix} \cdot \begin{pmatrix} r_{t-1}^* \\ r_{t-2}^* \\ r_{t-3}^* \end{pmatrix} + \begin{pmatrix} 0.0227 \\ -0.2088 \\ 0.0039 \end{pmatrix} \cdot \begin{pmatrix} s_{t-1}^* \\ s_{t-1}^* \\ s_{t-1}^* \end{pmatrix} + \epsilon_{r,t}. \tag{8}$$

The 144-element variance-covariance matrix is not presented here due to space considerations. Shocks for the probabilistic trajectories are generated by resampling from the residuals.

---

[1] http://www.ssa.gov/OACT/ProgData/effectiveRates.html

References

Heuser, Robert L. 1976. *Fertility tables for birth cohorts by color: United States, 1917–1973.* U.S. Department of Health, Education, and Welfare. DHEW Publication No. (HRA) 76-1152.

Heuser, Robert L. 2003. "U.S. Cohort and Period Fertility Tables, 1917–1980." Made available by the Office of Population Research, Princeton University: http://opr.princeton.edu/archive/cpft/.

Ibbotson Associates. 2002. *Stocks, Bonds, Bills, and Inflation Yearbook.* Chicago: Ibbotson Associates.

Lee, Ronald D. 1993. "Modeling and Forecasting the Time Series of U.S. Fertility: Age Patterns, Range, and Ultimate Level." *International Journal of Forecasting* 9:187–202.

Lee, Ronald D. and Lawrence R. Carter. 1992. "Modeling and Forecasting U.S. Mortality." *Journal of the American Statistical Association* 87(419):659–671.

Lee, Ronald D. and Ryan D. Edwards. 2002. The Fiscal Impact of Population Aging in the US: Assessing the Uncertainties. In *Tax Policy and the Economy*, ed. J. M. Poterba. Vol. 16 pp. 141–180.

Lee, Ronald D. and Shripad Tuljapurkar. 1998. Stochastic Forecasts for Social Security. In *Frontiers in the Economics of Aging*, ed. David Wise. Chicago: University of Chicago Press.

Lee, Ronald D. and Timothy Miller. 2001. "Evaluating the Performance of the Lee-Carter Approach to Modeling and Forecasting Mortality." *Demography* 38(4):537–549.

Whelpton, Pascal K. 1954. *Cohort Fertility: Native White Women in the United States.* Princeton: Princeton University Press.

Table 1: Estimates of mortality SVD: a_x and b_x

age	a_x	b_x	age	a_x	b_x	age	a_x	b_x
0	−4.9255	0.0254	41	−6.1166	0.0096	82	−2.5803	0.0086
1	−7.5241	0.0223	42	−6.0476	0.0101	83	−2.4786	0.0085
2	−7.9330	0.0201	43	−5.9764	0.0105	84	−2.3757	0.0083
3	−8.1690	0.0202	44	−5.9093	0.0107	85	−2.2731	0.0081
4	−8.4324	0.0207	45	−5.8332	0.0109	86	−2.1724	0.0079
5	−8.5048	0.0204	46	−5.7568	0.0110	87	−2.0736	0.0076
6	−8.5304	0.0193	47	−5.6760	0.0111	88	−1.9766	0.0075
7	−8.5827	0.0187	48	−5.5875	0.0113	89	−1.8813	0.0073
8	−8.6393	0.0189	49	−5.4976	0.0114	90	−1.7875	0.0071
9	−8.7618	0.0195	50	−5.4045	0.0115	91	−1.6948	0.0069
10	−8.8642	0.0206	51	−5.3116	0.0115	92	−1.6029	0.0067
11	−8.8641	0.0204	52	−5.2224	0.0114	93	−1.5121	0.0064
12	−8.6099	0.0168	53	−5.1391	0.0111	94	−1.4225	0.0062
13	−8.1828	0.0120	54	−5.0575	0.0107	95	−1.3370	0.0060
14	−7.8184	0.0085	55	−4.9758	0.0103	96	−1.2571	0.0058
15	−7.5044	0.0061	56	−4.8891	0.0100	97	−1.1815	0.0058
16	−7.2948	0.0048	57	−4.7973	0.0097	98	−1.1121	0.0058
17	−7.1536	0.0043	58	−4.6993	0.0096	99	−1.0467	0.0058
18	−7.0707	0.0045	59	−4.5992	0.0095	100	−0.9815	0.0058
19	−7.0336	0.0050	60	−4.5009	0.0095	101	−0.9156	0.0058
20	−7.0034	0.0056	61	−4.4047	0.0094	102	−0.8499	0.0058
21	−6.9632	0.0060	62	−4.3100	0.0091	103	−0.7796	0.0058
22	−6.9459	0.0062	63	−4.2186	0.0088	104	−0.7146	0.0058
23	−6.9397	0.0060	64	−4.1291	0.0083	105	−0.6418	0.0058
24	−6.9393	0.0055	65	−4.0380	0.0079	106	−0.5807	0.0058
25	−6.9391	0.0050	66	−3.9485	0.0076	107	−0.4853	0.0058
26	−6.9440	0.0046	67	−3.8670	0.0075	108	−0.4414	0.0058
27	−6.9277	0.0043	68	−3.7937	0.0074	109	−0.3560	0.0058
28	−6.8907	0.0041	69	−3.7238	0.0075	110	−0.2207	0.0058
29	−6.8451	0.0040	70	−3.6508	0.0076			
30	−6.7874	0.0040	71	−3.5733	0.0077			
31	−6.7328	0.0041	72	−3.4929	0.0078			
32	−6.6768	0.0043	73	−3.4099	0.0079			
33	−6.6197	0.0047	74	−3.3240	0.0080			
34	−6.5620	0.0051	75	−3.2341	0.0081			
35	−6.5004	0.0056	76	−3.1416	0.0081			
36	−6.4354	0.0061	77	−3.0495	0.0082			
37	−6.3741	0.0068	78	−2.9587	0.0084			
38	−6.3131	0.0075	79	−2.8674	0.0085			
39	−6.2500	0.0083	80	−2.7746	0.0086			
40	−6.1876	0.0090	81	−2.6790	0.0087			

Table 2: Estimates of fertility SVD: c_x and d_x

age	c_x	d_x
15	0.0079	−0.0025
16	0.0205	−0.0037
17	0.0381	−0.0010
18	0.0606	0.0082
19	0.0840	0.0233
20	0.1012	0.0420
21	0.1115	0.0591
22	0.1136	0.0722
23	0.1141	0.0793
24	0.1142	0.0814
25	0.1147	0.0768
26	0.1154	0.0679
27	0.1157	0.0567
28	0.1144	0.0487
29	0.1116	0.0404
30	0.1059	0.0358
31	0.1009	0.0260
32	0.0909	0.0246
33	0.0804	0.0225
34	0.0695	0.0248
35	0.0593	0.0251
36	0.0501	0.0256
37	0.0407	0.0257
38	0.0319	0.0264
39	0.0244	0.0251
40	0.0180	0.0234
41	0.0130	0.0180
42	0.0088	0.0146
43	0.0055	0.0109
44	0.0031	0.0086
45	0.0017	0.0058
46	0.0006	0.0038
47	0.0002	0.0024
48	0	0.0012
49	0	0.0006
50	0	0.0003