Journal of Official Statistics, Vol. 22, No. 3, 2006, pp. 453–485

Forecasting Labor Force Participation Rates

Edward W. Frees[1]

Motivated by the desire to project the financial solvency of Social Security, the article discusses forecasts of labor force participation rates. The data are highly multivariate in the sense that, at each time point, rates are disaggregated by gender, age, marital status as well as the presence of young children in the household; taken together, there are 101 demographic cells at each point in time. Thirty-one years of data from the U.S. Bureau of Labor Statistics are available. As input to Social Security projections, it is desirable to produce forecasts over a 75-year time horizon.

The purposes of this article are to give a structure to the problem of forecasting labor force participation rates, to summarize the data and to show how to implement different types of basic models.

For the basic structure, the article explores the use of statistical techniques commonly employed in demography for forecasting vital rates such as mortality and fertility. Specifically, we examine principal components techniques popularized by Lee and Carter (1992), as well longitudinal data techniques, for forecasting.

In summarizing the data, we find that differencing rates captures much of the dynamic movement of rates, as measured by short-range out-of-sample validation. Moreover, a logistic transformation stabilizes long-range forecasts.

The long-range forecasts varied substantially over the type of model selected, ranging from an average of 88.9% for females (78.4% for males) to 65.1% (61.3%) over a 75-year horizon. This large difference reflects the well-known fact that sex-specific models can summarize small historical differences yet project them into large differences in the distant future. To remedy this, we propose a simple autoregressive model that captures gender differences with common parameters. This simple model generates a 71.4% average participation rate for females (77.5% for males), a modest increase (decrease) compared to current rates.

Key words: Longitudinal data; principal components; demographic forecasting; JEL classification; Primary J110; Secondary J210.

1. Introduction

Labor force participation rate (LFPR) forecasts, coupled with forecasts of the population, provide us with a picture of a nation's future workforce. This picture provides insights into the future workings of the overall economy, and thus LFPR projections are of interest to a number of government agencies. In the United States, LFPRs are projected by the Social Security Administration (SSA), the U.S. Bureau of Labor Statistics, the Congressional Budget Office, and the Office of Management and Budget. In the context of Social

[1] University of Wisconsin, School of Business, 975 University Avenue, Madison, WI 53706, U.S.A. Email: jfrees@bus.wisc.edu
Acknowledgments: I thank anonymous reviewers for comments on the article. The Assurant Health Insurance Professorship and the National Science Foundation Grants SES-0095343 and SES-0436274 provided funding to support this research.

Security, policy-makers use labor force projections to evaluate proposals for reforming the Social Security system and to assess its future financial solvency.

Is the quantity of the future labor force important to Social Security financial projections? If Social Security were funded like private pension plans, the answer would be murky. Funding of private pension plans is based on the principle of individual equity so that each plan member supports his or her own benefits. As contributions increase through employment, benefits increase commensurably. In contrast, funding for Social Security lacks this direct link between contributions and benefits. Under Social Security, as the population works more, contributions increase. Further, benefits also increase because more people are becoming eligible for Social Security, but the increase in benefits is lower than would be under a private system. Thus, the number of workers can have a strong influence on the financial solvency of the Social Security system. Because essentially all workers are required to make Social Security contributions, this article focuses on the number of people participating in the labor force. In this sense, forecasts of labor force participation rates are critical to forecasts of financial status of the Social Security system.

This article is directed towards developing models that can be used for forecasting labor force participation rates for Social Security projections or other policy purposes. As such, we will develop LFPR projections by demographic cell, specifically using age, sex and marital status breakdowns. Further, in addition to short-range, we will also be interested in medium- and long-range projections; SSA currently publishes projections up to 75 years. Other government agencies focus their LFPR projections on a ten-year period that is critical for budgeting purposes. In lieu of projections by demographic cell, we could decompose the labor force by sector of the economy or geographic region. However, because of the interest in longer run projections, the SSA does not decompose LFPRs by means of economic or geographic breakdowns, focusing instead on demographic breakdowns. Because the purpose of this article is to propose models that can be used in SSA projections, we also focus on disaggregation of the labor force by demographic cell.

The literature on forecasting labor force participation is relatively small, given the extensive literature on the determinants of LFPRs in the labor economics, management and pensions literatures. Most of this modeling research has focused on the causal nature of these determinants, in contrast to the forecasting interest of this article. Still, these literatures provide insights into the problem of forecasting; Section 2 contains a brief survey. The modeling research on forecasting labor force participation seems to ignore the age and marital status shape patterns, focusing instead on aggregate rates; see, for example, Smith and Ward (1985). In contrast, there is substantial government published work that considers projections of disaggregated labor force participation rates. Section 3 describes the projections by the U.S. Bureau of Labor Statistics.

This article uses statistical forecasting techniques for disaggregated rates that were developed in demographic contexts. Bell (1997) describes two methods for fitting and forecasting age-specific fertility and mortality rates: the curve fitting approach and the principal components approach. That is, specialized time series techniques are required because at each time point there are many rates when disaggregated by age. The many rates can be expressed as a vector at a point in time; however, this vector has a high dimension. Standard multivariate time series techniques are inadequate because of dimensionality problems.

The first approach for overcoming the dimensionality problem consists of fitting parametric curves to rates and using multivariate time series techniques to forecast the curve parameters. To illustrate, Thompson et al. (1989) fit Gamma curves to fertility rates, treating each curve as a function of age of the mother. Four parameter Gamma curves, that were scale and shape shifted, were fit to each year of data, thus reducing 32 observations (one for each age of the mother, 14–45) to four statistics. The parameters were then forecast using multivariate time series techniques, and these forecast parameters produced a forecast fertility function. This approach not only reduces the dimensionality of the problem but also ensures that forecast rates have a smooth shape across ages similar to that of historical data.

The curve fitting approach is not considered in this article because we wish to disaggregate labor force participation not only by age but also by marital status and the presence of children. These discrete categorizations do not lend themselves readily to smooth curve fitting with just a few parameters. Moreover, even when indexed only by age, the economic labor force participation rates do not display the same smooth age patterns that are enjoyed by the biological fertility rates. For example, there are abrupt discontinuities in labor force participation at age 65 because of the "normal retirement age" that was available for many generations as part of their public and private pension plans.

The second approach summarized by Bell (1997) is the principal components technique that has been popularized by Lee and Carter (1992) in the demographic literature on forecasting mortality and fertility rates. Here, one decomposes the vector of rates and focuses on those components that account for most of the variability, the so-called "principal" components. The principal components are forecast using multivariate time series techniques; forecasts of the principal components are then used to provide forecasts of the vector of rates. For our application, an important advantage compared to the parametric curve fitting approach is that the vector need not be indexed by a smooth function, such as age. This allows us to readily incorporate discrete indices such as marital status and the presence of children. Like the parametric curve fitting approach, the principal components approach reduces the dimensionality of the problem and ensures that forecast rates have a smooth shape across ages similar to that of historical data. Lee and his co-workers have established that this approach has been very fruitful for forecasting mortality and fertility in a Social Security context; see Carter and Lee (1992), Lee (2000), Lee, Carter, and Tuljapurkar (1995), Lee and Miller (2001), and Lee and Tuljapurkar (1994, 1997).

We also use a third type that we refer to as the "longitudinal data technique." Here, the rates are organized by demographic cell and each rate is considered a response in a regression model. We use longitudinal data methods to investigate the sharing of parameters across responses from different demographic cells. This being a regression based methodology, we can easily incorporate the effects of explanatory (exogenous) variables. Thus, in contrast to the principal components approach, this approach will allow us to readily incorporate known future events that affect specific demographic cells, such as the increase in the normal retirement age for Social Security. We will also be able to assess the effect of different policy reform proposals that may affect future labor force participation rates. Moreover, following the "seemingly unrelated regressions" literature, we can incorporate contemporaneous correlations through the covariance structure. A drawback of this approach is that, without appropriate modifications, it does not guarantee that forecast rates have a smooth shape across age, marital status and the presence of children similar to that of historical data, as with the first two types of techniques considered.

Section 4 describes the longitudinal data techniques in greater detail. The details of forecasting with principal components are widely available; the Appendix provides an introduction in order to keep this article self-contained.

An advantage of the projection models that are described in this article is that they are stochastic. Stochastic projection models provide a mechanism for assessing the uncertainty of the projections. We anticipate that underlying economic and demographic conditions in which the system operates will change in the future. Models that incorporate uncertainty will allow us to assess the robustness and sustainability of the system in the presence of changing economic and demographic conditions. For example, some policy designs will be adaptable ("immune") to changing economic conditions whereas other designs are insensitive to current economic conditions. This adaptability can be measured using ideas of variance, a concept that is not available when using deterministic projection techniques. In the interest of brevity, we will not take advantage of stochastic features of the model in this article. However, the advantages of stochastic projection systems have been made forcefully elsewhere; see, for example, Lee and Tuljapurkar (1997).

We do not compare our forecasts with those produced by the Social Security Administration (SSA). The SSA forecasts are based on a deterministic projection system. Forecasts of LFPRs use a static structural model that relates LFPRs to various economic and demographic factors. As such, it is highly multivariate. The goal of this article is to present and interpret the LFPR data and show how some simple models can be used to provide desirable forecasts.

The outline for the rest of the article follows. Section 2 describes some background literature on labor force participation, focusing on aspects that may influence future labor force participation patterns. Section 3 examines labor force participation data that are in focus in this article. Section 4 describes the statistical models used for forecasting. Results for short-range forecasts are in Section 5 and implications for long-range forecasts are in Section 6. A simple model that addresses many of the concerns raised in this article is provided in Section 7. Section 8 provides a summary and concluding remarks.

2. Literature Review

The composition of the workforce has changed dramatically in the last century. These changes have been influenced heavily by our changing attitudes towards retirement and by the larger presence of women in the workforce. This section reviews the bases for each of these influences, with a focus on forecasting.

2.1. The Role of Retirement

The word *retirement* generally connotes a complete and permanent withdrawal from paid labor. Some definitions supplement this with the idea of the receipt of income from pensions, Social Security, and other retirement plans (Purcell 2000).

In contrast, long-range historical retirement patterns are typically assessed based on the concept of "gainful" employment (Costas 1998). This construct measures the proportion of individuals who claim to have had an occupation in the year before the census was taken. With this construct, it can be seen that labor force participation for the elderly has changed dramatically over the last century. In 1900, 65% of men age 65 and over were in the labor force, primarily in the agricultural sector. Many left the labor force because of poor health or diminished employment prospects and became dependent on their children for support. By 1990, fewer than 20% of men age 65 and over were in the labor force, primarily in white-collar jobs. A large fraction of the elderly retires to enjoy leisure time in retirement. Health, unemployment, and income also play important roles in the retirement decision.

The data that we examine beginning in Section 3 is based on the current definition of labor force used by the Current Population Survey produced by the U.S. Bureau of Labor Statistics. This definition, initiated in 1940, is based on whether an individual worked for pay or sought employment during the survey week (not on his or her having an occupation). In particular, the labor force includes both the employed and unemployed.

There are several potential determinants that may explain the much lower labor force participation of elderly men at the end of the twentieth century compared to the beginning. Perhaps the most important reason is the availability of retirement income. In 1900, very few private pension plans were in operation, a national social retirement scheme was not available and organized tax-shelters to encourage private savings, such as 401(K)s, were not available. By the end of the twentieth century, each of these three institutional mechanisms provides important sources of income for the nation's elderly population.

In addition to providing retirement income, both Social Security and private pension plans offer substantial financial incentives to retire early (more so for private plans). Estimates of the effect of private plans on labor supply are large, although there is considerable disagreement (Anderson, Gustman, and Steinmeier 1999).

The size of wealth and wages may also influence retirement rates through eventual retirement income. To illustrate, Costas (1998) argues that increased wealth for a group of Civil War Union Army pensioners had a substantial effect on retirement.

For assessing retirement patterns over the last century, it is also important to note the shift in the economy from its agricultural basis to the manufacturing and service sectors. Retirees age 65 and over in 1900 had begun their working lives when the U.S. was primarily agricultural. Those in farming generally had phased retirements rather than a complete and abrupt withdrawal from the labor force. In contrast, by the end of the century, relatively few switch from full to part time and about 75% switch from full to zero percent time (Rust 1990). Further, the employment structure may have shifted to encourage occupations with lower retirement ages (Perrachi and Welch 1994).

The changing role of disability benefits may have an important influence of the changing retirement patterns. When introduced, the Social Security disability program was small and inconsequential. The liberalization of qualifying rules for disability benefits and the increasing generosity of these benefits may be partially responsible for declines in LFPR for workers age 50–64.

Modeling changes in attitudes towards retirement may be the most daunting task. It seems clear that retirees today enjoy better health than in the past. Quinn and Burkhauser (1990) suggest that health is an important factor in the retirement decision. Retirees are unlikely to live in their children's homes. They enjoy better recreational amenities, a better climate and have more entertainment options, such as golf and TV. Costas (1998) documents an increasing number of retirees that cite a preference for leisure as their main motivation for retirement.

2.2. The Role of Women

As documented by papers in Smith (1980) and Blau (1998), women's participation in the labor force changed fundamentally in the latter part of the twentieth century. Unlike men, the trend of labor force participation rates for women continues to increase, although rates of increase for periods 1990–1998 are lower than in prior years (Fullerton 1999b).

Increases in female labor force participation rates can be seen on both a period and a cohort basis (Blau 1998; Goldin 1990; Fullerton 1999b). The cohort of women born during 1926–1945 seems to have had the largest increase in labor force participation. From 1960 to 1970, they were in the 16–24 and 25–34 age groups. They had 8.5% (16–24) and 9% (25–34) increases from 1960 to 1970. From 1970 to 1980, they were in the 25–35 and 35–44 age groups. They had 20.5% (25–34) and 14.4% (35–44) increases from 1970 to 1980.

There are several explanations for increased female labor force participation. These include young women postponing and reducing fertility, reduction of marriage and increases in divorce. Further, there seems to be a substantial rise in attachment to the labor force among new mothers, particularly married women (Olsen 1994; Blau 1998).

Other important determinants of female labor force participation rates include education and the presence of young children. It is well documented that levels of education affect labor force participation rates. Although the rates for less educated males have fallen over the last twenty-five years, corresponding female rates have not risen as quickly as other education level groups (Blau 1998). There is also strong evidence that the presence of young children in the household tends to reduce the labor participation of women (Killingsworth and Heckman 1986).

3. Data

The data for this article were provided by the SSA; these data are used by the SSA for their projections of Social Security financial status. The data are compiled by the U.S. Bureau of Labor Statistics, which also makes short-range projections of the labor force as summarized in Fullerton (1999a, 1999b) and the *BLS Handbook of Methods* (1997, Chapter 13). The U.S. Bureau of Labor Statistics also uses a demographic cell approach although it uses age, gender and race origin groups, compared to the age, gender, marital status and presence of young children groups used by the SSA. This section summarizes the key features of the data; additional summary information can be found in Fullerton (1999a, 1999b). Specifically, this section demonstrates the strong patterns of LFPRs by age, gender and the presence of children and underscores the distinct differences between males and females.

The labor force participation rates are the civilian labor force divided by the civilian noninstitutional population. The data are broken down by gender and 29 age categories corresponding to 16–17, 18–19, 20–24, 25–29, . . . , 45–49, 50–54, 55, 56, . . . , 73, 74. For the quinquennial age brackets from ages 20 through 54, there are three categories of marital status, corresponding to never married, married with spouse present and married with spouse absent. Moreover, for females age 20–44, we also have breakdowns according to whether their own child under the age of six is present in the household. Thus, we consider a total of 101 demographic cells in this article, 43 for males and 58 for females. For the quinquennial age brackets from ages 20 through 54, the civilian labor force and civilian noninstitutional population come from the Current Population Survey for the month of March. For the other data, the civilian labor force and civilian noninstitutional population come from the U.S. Bureau of Labor Statistics Annual Household Survey. We consider data from 1968 through 1998.

Because numerical details of the data are readily available (Fullerton 1999a, 1999b), we focus on graphical summaries. Figure 1 summarizes the total labor force participation rate for each year during the period 1968–1998. Here, we see that total LFPR has increased due to the increased labor force participation rates of females. Male LFPRs have declined over the period, from 81% in 1968 to 78% in 1998.

Figure 2 shows the labor force participation rate as a function of age, by gender, with a comparison of the first year 1968 to the most recent year 1998. It is useful to compare the shape of the curve between years, for males and females. For males, we see that the difference between the two LFPR curves, that measures the drop in labor force participation, decreases as age grows over the working years, 20–55, and is constant after age 55. In contrast, for females we see that the difference between the two LFPR curves is large and constant over the primary working years, 20–55, and declines after age 55.

Figures 3a–4b provide further information on the change in LFPRs over time, by age. Figures 3a and 3b focus on the teen and working careers ages 16–54. Figure 3a shows that labor force participation rates for male teens, 16–17 and 18–19, has been relatively constant. This is also true of the 20–24 age group. All other male age groups exhibit a small but steady decline in labor force participation rates. In contrast, Figure 3b shows a steady increase in female LFPRs for all age groups, except for teens ages 16–17 and 18–19.

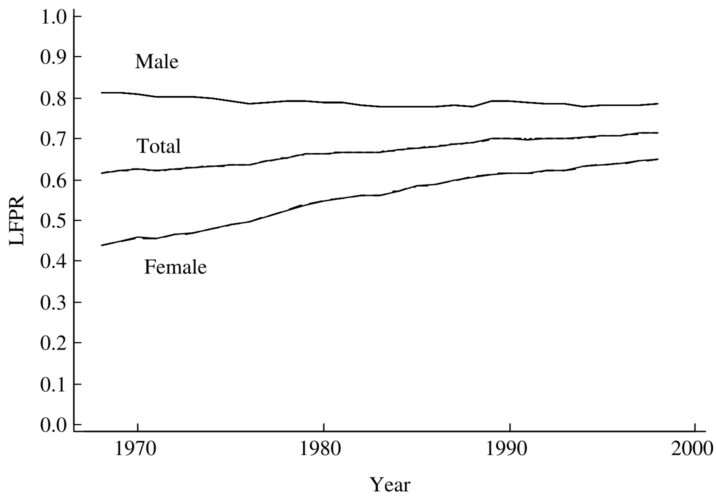

Fig. 1. Multiple time series plot of the Total Labor Force Participation Rate (LFPR), with breakdowns by gender

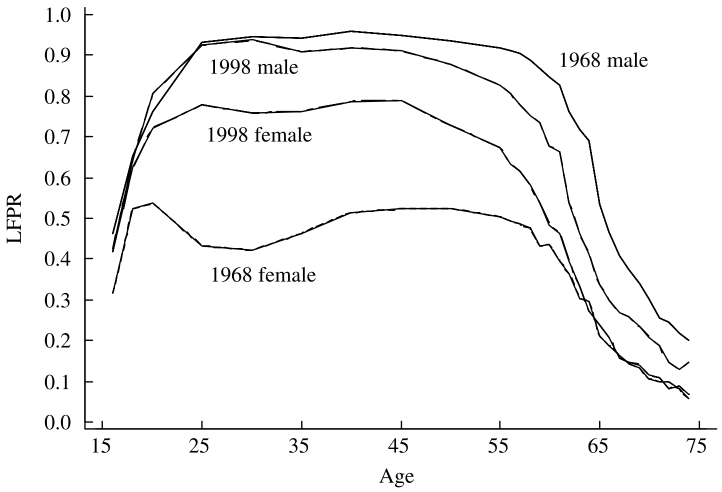

Fig. 2. Labor Force Participation Rate (LFPR) by age, for 1968 and 1998 male and females

Figures 4a and 4b focus on the later working years, 55–74. Figure 4a shows a steady decline in male LFPRs for those prior to eligibility for Social Security early retirements benefits, ages 55–61, with more rapid declines for ages 62–64. Although difficult to see in this figure, there is also a decline at age 65. For ages 66–74, labor force participation rates are relatively stable. In contrast, Figure 4b shows steady increases in female LFPRs for ages 55–61. Further, female LFPRs for ages 62–74 appear stable over the period 1968–1998.

Most of our analyses examine movements of rates from calendar year to calendar year, known as the "period" basis in the demography literature. It is also possible to follow rates

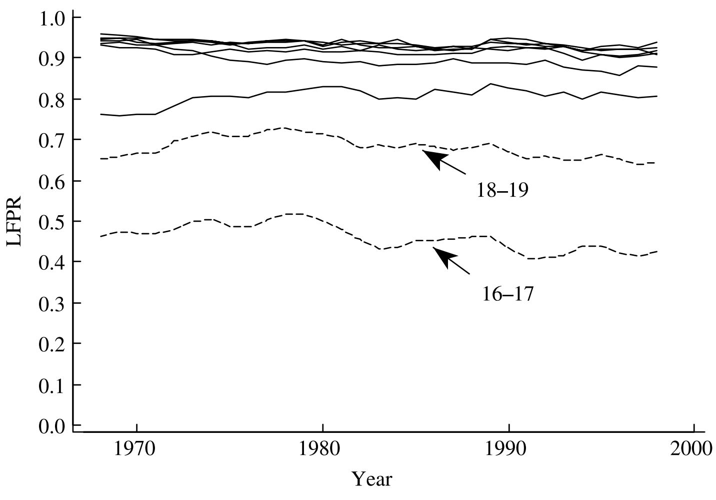

Fig. 3a. Multiple time series plot of male Labor Force Participation Rate (LFPR) by age. Each line corresponds to an age group. Groups 16–17 and 18–19 are marked; the remaining are the quinquennial age groups 20–24, 25–29, . . ., 50–54

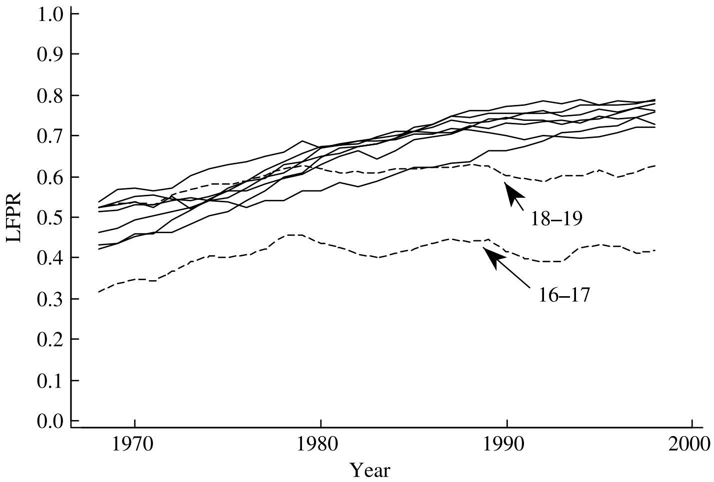

Fig. 3b. Multiple time series plot of female Labor Force Participation Rate (LFPR) by age. Each line corresponds to an age group. Groups 16–17 and 18–19 are marked; the remaining are the quinquennial age groups 20–24, 25–29, . . ., 50–54

according to year of birth, known as the "cohort" basis. Figure 5 illustrates this approach. Here, each line follows a group of workers over the ages 55–74, where the group is defined by the workers' year of birth. However, long histories of labor force participation are required for a complete picture of a cohort's working career. For example, with 31 years of data, we can observe the complete primary working experience over ages 25–54 for only two cohorts; moreover, no cohorts are completely observed over ages 25–74. Further, although not presented here, the figures corresponding to Figures 1–4 on a cohort

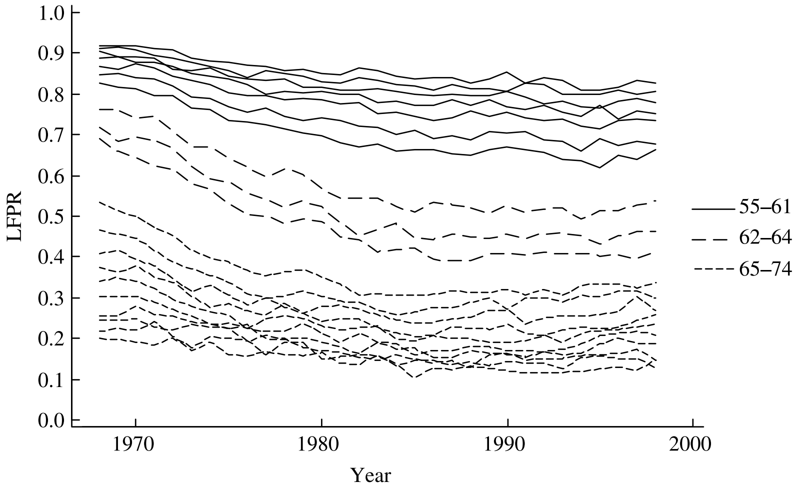

Fig. 4a. Multiple time series plot of male Labor Force Participation Rate (LFPR) by individual ages 55–74. Each line corresponds to an age. Ages 55–61 are graphed with solid lines, ages 62–64 are graphed with dashed lines and 65–74 are graphed with dotted lines

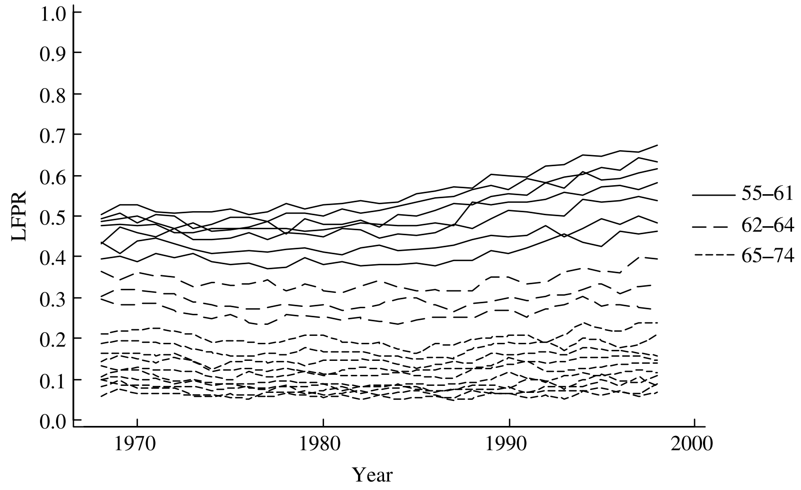

Fig. 4b. Multiple time series plot of female Labor Force Participation Rate (LFPR) by individual ages 55–74. Each line corresponds to an age. Ages 55–61 are graphed with solid lines, ages 62–64 are graphed with dashed lines and 65–74 are graphed with dotted lines

basis are more erratic than on a period basis. This lack of stability makes the forecasting problem more difficult. For these two reasons, we follow the method of analysis that is customary in the demography literature and examine our data on a period basis.

Figures 6 and 7 suggest the importance of examining labor force participation rates by marital status and presence of children under the age of six within a household. Figure 6 shows female labor force participation by marital status and presence of children, summarized over age groups 20–24, 25–29, . . . , 40–44. This figure underscores the

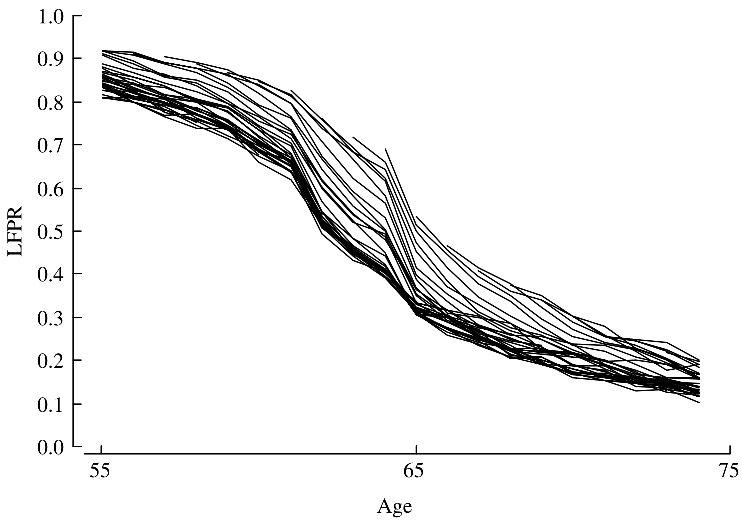

Fig. 5. Plot of male Labor Force Participation Rate (LFPR) versus age. Each line follows a birth cohort

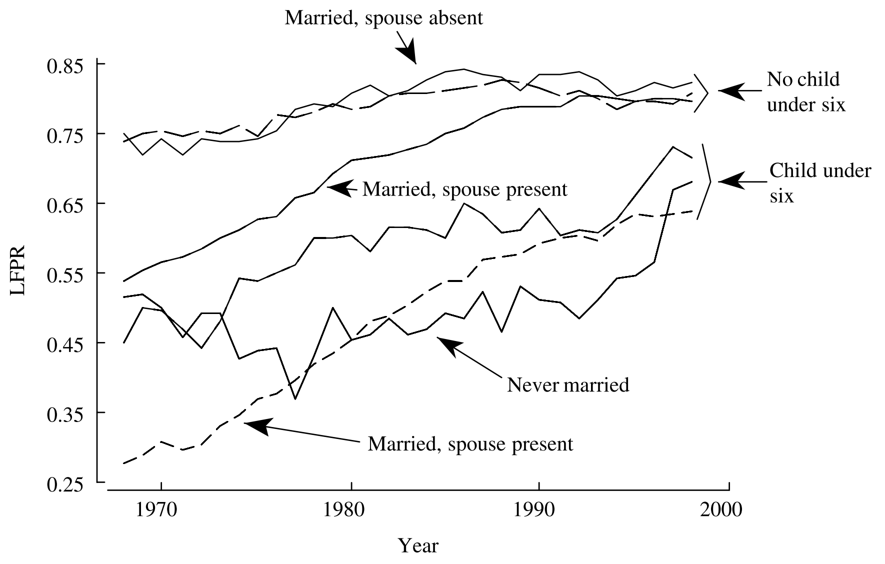

Fig. 6. Plot of female Labor Force Participation Rate (LFPR) versus year, by marital status and presence of children under the age of six in the household, ages 20-44

importance of considering the presence of children under the age of six within a household as a separate category. Further, Figure 6 shows that females that are married with the spouse present have historically behaved differently than females that were never married or with the spouse absent although, beginning in the 1990's, they now behave similarly to these other categories. However, for a larger number of age categories, Figure 7 suggests that females married with the spouse present still have different, and lower, participation

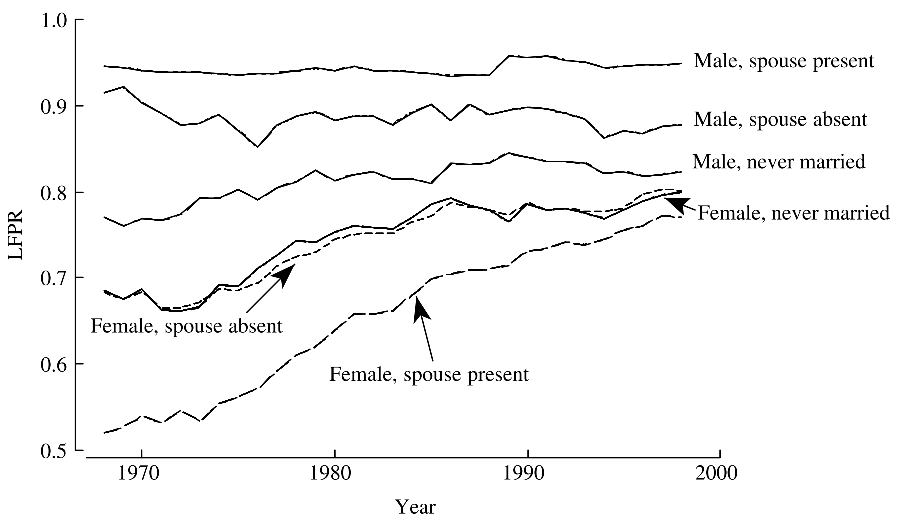

Fig. 7. Plot of Labor Force Participation Rate (LFPR) versus year, by gender and marital status, ages 20–54

rates than females that were never married or whose spouse is absent. Figure 7 also emphasizes the importance of marital status for male LFPRs.

In the demography literature, it is customary to summarize age-specific rates to provide an overview of the behavior of several rates simultaneously. For example, one often summarizes mortality rates using an expectation of life at birth. Similarly, one summarizes fertility using the "total fertility rate." Because of our interest in Social Security, we summarize labor force participation rates using the "expected number of working years," defined as follows. Let a denote the initial age and let b denote the last age under consideration. Further let ${}_tp_a$ denote the probability that a life aged a survives an additional t years. Then, we define

$$E_{a,b} = \sum_{t=0}^{b-a} {}_tp_a LFPR_{a+t}$$

to be the expected number of years that a life aged a works over the next $b - a$ years. Here, $LFPR_{a+t}$ denotes the probability that a life aged $a + t$ works during the year. Thus, for example, $E_{16,24}$ denotes the expected number of years that a life age 16 will work over the next nine years. To estimate $E_{a,b}$, we use a constant 1995 U.S. Population mortality (by gender, as appropriate), so that we summarize the changing effect of labor force participation rates, not mortality. Most calculations of $E_{a,b}$ will use labor force participation rates for a single year; thus $E_{a,b}$ provides a summary statistic on a period basis. We will also follow cohorts over time.

To illustrate, Figure 8 shows $E_{25,54}$, $E_{55,64}$ and $E_{65,74}$ for both males and females. Each calculation of $E_{a,b}$ is done for a calendar year and thus summary statistics are available for each year 1968–1998. Figure 8 shows the convergence of male and female expected working years. The greatest strides in narrowing the gap, though it is still the largest gap, are for the 25–54 age group. The expected number of working years for males aged 55–64

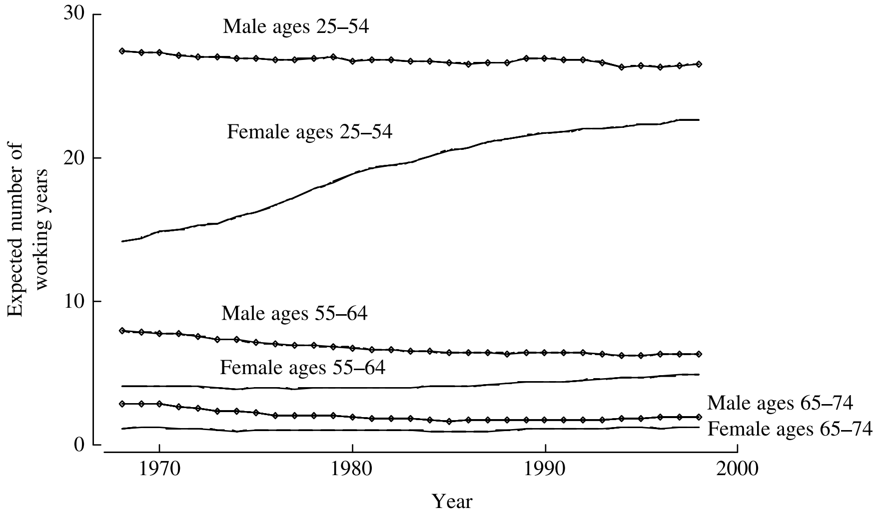

Fig. 8. Plot of expected number of working years versus year. Female age groups are indicated by a solid line, male age groups are indicated by a solid line plus a circle

has dropped considerably, from 7.94 in 1968 to 6.36 in 1998, yet has increased for females aged 55–64, going from 4.07 in 1968 to 4.86 in 1998. The difference for the elderly labor force participants, aged 65–74, is small and relatively constant. Although not shown in Figure 8, a similar pattern holds for young workers age 16–24.

Figure 9 summarizes the expected number of working years on a cohort basis, not the period basis that was used in Figure 8. We do not present the calculation of $E_{25,54}$; this calculation requires following a cohort over 30 years. With 31 years of data, complete information was available for only two cohorts. Instead, Figure 9 presents $E_{16,24}$, $E_{55,64}$ and $E_{65,74}$ for both males and females. This figure again shows that the gap between males and females is becoming narrower for each age group.

4. Models and Approach

Labor force participation rates, like mortality and fertility rates, are count data, expressed as a proportion of some total population. Count data often displays heteroscedasticity, with the variability as a function of the mean. For forecasting, it is customary to consider transformations of the rates to mitigate the heteroscedasticity. Logarithmic transforms are often considered for mortality and fertility data. However, mortality and fertility are infrequent events compared to labor force participation; thus, we consider alternatives to the logarithmic transform.

Specifically, we consider models where the response is the labor force participation rate ($y_{it} = LFPR_{it}$), the logarithmic rate ($y_{it} = \log LFPR_{it}$) and the logistic transform of the rate ($y_{it} = \log(LFPR_{it}/(1 - LFPR_{it}))$). To handle time trends, we also consider differencing each transform. Specifically, we consider the difference of rates ($y_{i,t} = LFPR_{i,t} - LFPR_{i,t-1}$), the difference of logarithmic rates ($y_{i,t} = \log LFPR_{i,t} - \log LFPR_{i,t-1}$) and the difference of the logistic transform, ($y_{i,t} = \log(LFPR_{it}/(1 - LFPR_{it}) - \log(LFPR_{i,t-1}/(1 - LFPR_{i,t-1}))$. In

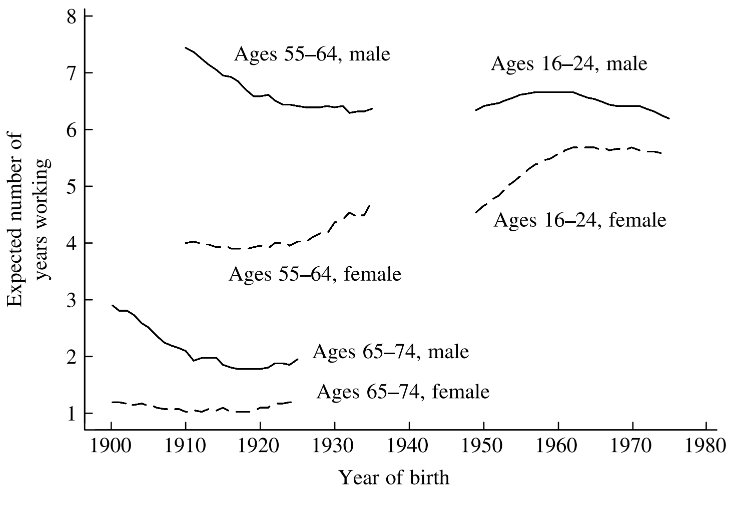

Fig. 9. Plot of expected number of working years versus year of birth. Calculations are done on a cohort basis. Female age groups are indicated by a dotted line, male age groups are indicated by a solid line

each case, a model will be used to predict the response. Then, the appropriate transformation will be made to convert the predicted responses into predictions for labor force participation rate. In our models, we use the symbol y_{it} to denote response for the ith cell at time t. These cells are formed using age, marital status and the presence of children under six within the household. As described in Section 3, we have 43 categories for males and 58 categories for females. Because of the distinct gender patterns discussed in the Section 2.2 literature review and observed in the Section 3 data summaries, we run models separately for males and females.

Government projections of labor force participation rates by the U.S. Bureau of Labor Statistics and SSA consider each cell separately. In the context of fertility and mortality, Bell (1997) notes the drawbacks in forecasting each cell separately. First, the long-run forecast curves may show irregular shapes that are not evident in historical data. Second, it is difficult to assess the forecast uncertainty for summary measures of the forecast curves since forecasting individual cells separately does not directly provide estimates of the correlations among different ages. Given the smooth shape of both fertility and mortality curves, one would expect correlations to be high. Similarly, in our labor force participation context, we would expect rates for cells that share similar gender, age, marital status and children status characteristics to be highly correlated.

4.1. Longitudinal Data Models

In contrast to modeling each cell separately, Section 4.1 and the Appendix review techniques that can be used to forecast all cells simultaneously. One set of forecasting techniques is based on longitudinal data models of the form

$$y_{it} = \mathbf{z}'_{it}\boldsymbol{\alpha}_i + \mathbf{x}'_{it}\boldsymbol{\beta} + \varepsilon_{it} \tag{4.1}$$

for demographic cell i and time t; see, for example, Frees (2004). The term $\boldsymbol{\alpha}_i$ is a vector of cell-specific parameters and $\mathbf{z}_{it}$ is the corresponding vector of explanatory variables that are known when y_{it} is observed. The term $\boldsymbol{\beta}$ is a vector of parameters common to all cells and $\mathbf{x}_{it}$ is the corresponding vector of explanatory variables. For example, the model where $\mathbf{z}_{it} = 1$ without any common parameters is

$$y_{it} = \alpha_i + \varepsilon_{it} \tag{4.2}$$

Assuming that the errors $\{\varepsilon_{it}\}$ are i.i.d., the forecast of all future values for the ith cell is $\bar{y}_i$, the average response for the ith cell. When the responses are differences of labor force participation rates, Equation (4.2) represents a random walk model of the original rates with a drift term. When considering fertility and mortality, Bell (1997) found that no curve fitting or principal component model dominated this simple alternative using short-term out-of-sample criteria.

As another example without common parameters, we may take $\mathbf{z}_{it} = (1\ y_{i,t-1})'$, so that the model is:

$$y_{it} = \alpha_{i,1} + \alpha_{i,2}y_{i,t-1} + \varepsilon_{it} \tag{4.3}$$

This is an autoregressive model of order 1, $AR(1)$, for each cell. A version of this model is currently being used by the SSA for their forecasts (Frees 1999).

The model in Equation (4.3) has two regression parameters for each group, totaling $202(= 2 \times 101)$ for our case. A more parsimonious version of this model is

$$y_{it} = \beta_1 + \beta_2 y_{i,t-1} + \varepsilon_{it} \tag{4.4}$$

This model may be arrived at from Equation (4.1) by omitting cell-specific parameters and using $\mathbf{x}_{it} = (1\, y_{i,t-1})'$ for the explanatory variables associated with common parameters in the vector $\boldsymbol{\beta}$.

The models in Equations (4.2)–(4.4) constitute what we call the "basic" longitudinal data models. As an example of a more complex type of longitudinal data model, we also consider a version of the model in Equation (4.2) but with the cell-specific terms treated as a random variable, not an unknown parameter. This representation is known as a "random effects" model. This change affects the forecast through both the mean and the variance function. In particular, a random α_i that is common to all responses within the ith cell affects the serial correlation of the series $y_{i1}, \ldots, y_{iT}$. To allow a more flexible representation of the serial correlation structure, we also consider models such as in Equation (4.2) where α_i is random and the series $\{\varepsilon_{i1}, \ldots, \varepsilon_{iT}\}$ has an $AR(1)$ structure.

Each longitudinal data model that we have discussed so far has assumed that responses from different cells are independent. However, as with mortality and fertility data, it seems reasonable to suppose that the response from "neighboring" cells will in some way derive from similar experiences. A companion article that focuses on modeling variance components, Frees (2003), finds large positive correlations among teen cells and among older worker cells. Section 7 introduces a model where parameters from closely related cells are taken to be identical. Alternatively, the principal components model incorporates this feature by considering the appropriate linear combinations of cells most suitable for forecasting.

4.2. *Model Selection Criteria*

As described in the introductory part of this section, we allow the response variable y to be either the original rates *LFPR*s, the logged rates, the logistic transform of the rates or the corresponding differenced versions. Each unit of measurement has advantages relative to other units of measurement. The original measurement unit is the easiest to explain and interpret. The logarithmic measurement is useful because negative forecasts occasionally appear; the anti-logarithmic transformation (exponential) forces the forecasts to be positive. Similarly, the logistic transformation forces the forecasts to lie in the unit [0,1] interval. The difference transformation accounts for trends in the data; further, the difference of logarithmic values can be interpreted as proportional changes.

To compare results from different models and different units of measurement, we use two out-of-sample summary statistics. The first is based on the mean absolute percentage error, defined as

$$MAPE_t = \sum_i \left| \frac{LFPR_{i,t} - LF\hat{P}R_{i,t}}{LFPR_{i,t}} \right|$$

Here, $LFPR_{i,t}$ is the (actual) cell i labor force participation rate ($LFPR$) at time t, Σ_i represents the sum over all cells (that may include combinations of age, marital status and presence of a child under six) and $LF\hat{P}R_{i,t}$ is the forecast of $LFPR_{i,t}$. For example, when the response was logarithmic *LFPRs*, we estimated the model using past logarithmic *LFPRs*, forecast future logarithmic *LFPRs*, and converted these forecasts by taking exponents to get the forecasts $LF\hat{P}R_{i,t}$. The second out-of-sample summary statistic is the mean absolute error, defined as

$$MAE_t = 100 \times \frac{\sum_i \left| LFPR_{i,t} - LF\hat{P}R_{i,t} \right| N_{i,t}}{\sum_i N_{i,t}}$$

This statistic is weighted by $N_{i,t}$, the number of people in cell i at time t. See for example, Ahlburg (1982) for a discussion of advantages and disadvantages of these and alternative summary measures of forecasting accuracy.

5. Short-Range Out-of-Sample Results

This section summarizes the results of the model estimation and forecasting over a limited number of years, that we call "short-range." Forecasts from different models tend to be highly correlated with one another when based on a single in-sample period; thus, we used different in-sample periods to see the effects of the model adoption at different points in time. Specifically, we consider several in-sample periods, each beginning with 1968, and ending in the years 1982, 1986, 1990, and 1994, respectively. For each in-sample period, we forecast $t = 1, 2, 3, 4$ periods in advance and computed the out-of-sample statistics, $MAPE_t$ and MAE_t. When comparing models, we report the average *MAPE* and *MAE* for the four-year period.

To illustrate, Table 1 reports the mean absolute errors, for $t = 1, 2, 3, 4$, using female labor force participation rates as the in-sample data from 1968–1994, inclusive. Four models are considered, the model corresponding to Equation (4.2) that uses the cell-specific average value as the forecast, the $AR(1)$ model for each cell, corresponding to Equation (4.3), and the $AR(1)$ model with common parameters, corresponding to Equation (4.4). We also present the forecast using the most recent observation from a cell; this forecast, a constant over time, corresponds to a random walk model with known drift parameter equal to zero. Table 1 shows the MAE_t for each value of t, as well as the average over four years. In subsequent tables, we present only the four-year average. Models with low errors are desirable; based on the four-year average MAE criterion in Table 1, we would choose the model that uses the most recent cell observation as the constant (over time) forecast.

Tables 2a and 2b summarize the results for the "basic" longitudinal data models. These are the same models considered in Table 1, yet including male and female, *MAPE* and *MAE*, as well as the six units of measurement: the original percent units, logarithmic units, logistic transforms, differences, differences of logarithmic units and differences of the logistic transforms.

Several conclusions are evident from Tables 2a and 2b. The "constant" forecasts, based only on the most recent observations, do well for males, where rates have been relatively stable over the 1980s and 1990s, yet do not fare as well for females, where rates have

Table 1. Mean Absolute Errors (MAE) by year

Females. In-sample period is 1968–1994.

	1995	1996	1997	1998	Four-year average
Constant	1.466	1.217	1.927	1.982	1.648
Average	8.075	8.146	8.749	8.761	8.433
AR(1)	1.442	1.722	2.383	2.737	2.071
AR(1) Common	1.473	1.396	1.997	2.195	1.765

increased for many combinations of age, marital status and presence of young children. Using the time series average for each cell to forecast fares poorly, for the original percent units as well as the logarithmic and logistic transforms. One must use either differences or autocorrelations to handle the time trends; these two techniques seem to work equally well. The transformations display few strong advantages or disadvantages for short-range forecasting. In subsequent discussions, we focus on the autoregressive model with a common parameter, estimated using logistic transformation of *LFPRs*. This is a parsimonious model that will always have forecasts that lie within the unit interval and that seems to perform competitively in relation to alternatives based on the out-of-sample statistics in Tables 2a and 2b.

Tables 3a and 3b summarize the results for models based on principal components estimation. As in Tables 2a and 2b, we consider both male and female rates using the six units of measurement and summarize results with the four-year averages of *MAPE* and *MAE*. The principal components models include the model with one principal component, using this component with an adjustment for the mean of the series and for the most recent value of the series. The principal component was forecast using an integrated *AR*(1) model for responses that were not differenced and was forecast using an *AR*(1) model for differenced data. We experimented with other forecasting models for the time-varying components $\{k_t\}$ that are not reported here. The results reported are the best of those that we examined.

From Tables 3a and 3b, we conclude that use of only one principal component, without adjustment, on the original or transformed data, is a poor choice. For models where the responses were not differenced, the adjustment of the principal components using the most recent value of the series outperformed the adjustment based on the mean of the series for short-term forecasting. For models based on differenced data, the three principal component forecasting techniques performed equally well. The simplest model that does well is the one principal component model, without adjustments, based on either the differences or the differences of the transformed units. These models performed well for both male and females using both the *MAE* and *MAPE* basis. The performance of either of these models is comparable to the autoregressive model with a common parameter, estimated using logistic transforms of *LFPRs,* that was reported in Tables 2a and 2b.

Tables 4a and 4b summarize the results for another type of longitudinal data models, the so-called “random effects” type described in Section 4.1. For these models (without autocorrelation), the forecast is a weighted average of the cell-specific average response and the overall average response. Thus, it is not surprising that the forecast did poorly on both the original and transformed units because, in Table 2, we saw that the average was

Table 2a. Short-term out-of-sample statistics for basic longitudinal data models

Female

Response units	Forecasting technique	Mean absolute error Most recent in-sample year				Mean absolute percentage error Most recent in-sample year			
		1994	1990	1986	1982	1994	1990	1986	1982
Percent	Constant	1.648	1.816	2.374	2.176	3.123	3.124	4.261	3.325
Percent	Average	8.433	8.123	8.228	7.033	7.440	6.424	6.584	6.645
	AR(1)	2.071	1.868	1.830	1.863	3.651	3.305	3.685	3.687
	AR(1) Common	1.765	1.872	2.082	2.252	4.628	4.787	4.215	5.343
Logarithmic	Average	8.879	8.489	8.516	7.222	7.848	6.759	6.827	6.624
	AR(1)	2.140	1.958	1.889	1.920	3.857	3.529	3.847	3.627
	AR(1) Common	1.544	1.677	2.117	1.978	3.155	3.144	3.900	3.362
Logistic	Average	8.278	8.024	8.161	6.999	7.450	6.442	6.562	6.543
	AR(1)	1.967	1.769	1.793	1.783	3.652	3.327	3.690	3.596
	AR(1) Common	1.615	1.703	1.992	1.987	3.469	3.412	3.605	3.646
Difference	Average	2.045	1.997	1.950	1.469	3.275	3.507	4.603	3.115
	AR(1)	2.077	2.072	1.990	1.362	3.084	3.559	4.556	3.468
	AR(1) Common	1.488	1.871	2.003	1.782	2.980	3.583	3.803	4.233
Difference of logarithms	Average	2.413	2.408	2.100	1.667	3.461	3.748	4.657	3.206
	AR(1)	2.449	2.516	2.184	1.587	3.240	3.789	4.605	3.062
	AR(1) Common	1.598	1.929	2.160	1.764	2.849	3.118	3.988	3.107
Difference of logistics	Average	1.865	1.834	1.910	1.422	3.215	3.408	4.499	3.066
	AR(1)	1.883	1.877	1.939	1.323	3.035	3.440	4.450	3.300
	AR(1) Common	1.422	1.757	1.993	1.757	2.757	3.042	3.702	3.554

Table 2b. Short-term out-of-sample statistics for basic longitudinal data models

Male

Response units	Forecasting technique	Mean absolute error Most recent in-sample year				Mean absolute percentage error Most recent in-sample year			
		1994	1990	1986	1982	1994	1990	1986	1982
Percent	Constant	1.134	1.240	1.194	1.206	1.377	1.418	1.189	1.861
Percent	Average	2.367	2.666	2.742	2.863	2.728	4.193	4.636	5.459
	AR(1)	1.295	1.367	1.280	1.254	1.516	1.411	1.637	2.074
	AR(1) Common	1.690	0.996	1.968	1.266	1.871	1.473	2.028	1.655
Logarithmic	Average	2.301	2.597	2.674	2.814	2.496	3.912	4.370	5.261
	AR(1)	1.287	1.371	1.279	1.259	1.490	1.412	1.616	2.101
	AR(1) Common	1.867	0.991	2.120	1.358	1.854	1.343	1.936	1.611
Logistic	Average	2.389	2.666	2.761	2.891	2.684	4.113	4.570	5.433
	AR(1)	1.276	1.293	1.267	1.240	1.491	1.369	1.607	2.060
	AR(1) Common	1.573	0.969	1.846	1.189	1.572	1.286	1.635	1.615
Difference	Average	1.435	1.422	1.812	1.385	2.067	1.727	2.195	1.796
	AR(1)	1.442	1.442	1.805	1.535	2.121	1.731	2.271	1.869
	AR(1) Common	1.619	1.011	2.145	1.326	1.940	1.317	1.939	1.552
Difference of logarithms	Average	1.380	1.393	1.734	1.349	1.899	1.559	1.902	1.674
	AR(1)	1.411	1.413	1.740	1.499	2.008	1.567	1.995	1.730
	AR(1) Common	2.189	1.416	3.121	2.225	1.928	1.269	1.996	1.873
Difference of logistics	Average	1.503	1.353	1.858	1.405	2.059	1.636	2.114	1.820
	AR(1)	1.517	1.383	1.821	1.544	2.144	1.638	2.166	1.895
	AR(1) Common	1.407	0.990	1.921	1.078	1.946	1.283	1.935	1.452

Table 3a. Short-term out-of-sample statistics for principal component models

Female

Response units	Forecasting technique	Mean absolute error Most recent in-sample year				Mean absolute percentage error Most recent in-sample year			
		1994	1990	1986	1982	1994	1990	1986	1982
Percent	Basic PC	6.057	6.089	5.425	4.879	5.161	5.266	5.508	6.704
	PC with mean adj	3.044	3.578	2.527	1.960	4.536	5.019	4.763	3.599
	PC with recent adj	1.710	1.832	1.955	1.756	3.125	3.450	4.412	3.149
Logarithmic	Basic PC	6.886	7.046	8.570	7.193	6.077	6.204	6.900	6.610
	PC with mean adj	3.553	3.740	3.013	2.282	4.766	5.276	5.189	3.490
	PC with recent adj	1.737	1.971	2.127	1.966	3.137	3.594	4.363	3.227
Logistic	Basic PC	8.132	7.714	7.538	6.452	6.951	6.839	7.695	6.154
	PC with mean adj	2.769	3.125	2.477	1.904	4.480	4.789	4.680	3.407
	PC with recent adj	1.645	1.684	1.952	1.758	3.107	3.352	4.321	3.157
Difference	Basic PC	1.652	1.796	2.376	2.221	3.037	3.235	4.262	3.644
	PC with mean adj	2.035	1.994	1.948	1.514	3.119	3.537	4.546	3.439
	PC with recent adj	2.262	2.803	2.187	1.535	3.568	5.978	4.940	3.605
Difference of logarithms	Basic PC	1.648	1.809	2.386	2.214	3.037	3.220	4.267	3.255
	PC with mean adj	2.407	2.407	2.088	1.691	3.292	3.736	4.620	3.055
	PC with recent adj	2.971	3.489	2.187	1.699	4.887	7.052	4.950	3.225
Difference of logistics	Basic PC	1.651	1.794	2.379	2.207	3.036	3.222	4.255	3.581
	PC with mean adj	1.854	1.831	1.910	1.456	3.067	3.444	4.453	3.332
	PC with recent adj	1.953	2.462	2.041	1.473	3.225	5.489	4.650	3.491

Table 3b. Short-term out-of-sample statistics for principal component models

Male

Response units	Forecasting technique	Mean absolute error Most recent in-sample year				Mean absolute percentage error Most recent in-sample year			
		1994	1990	1986	1982	1994	1990	1986	1982
Percent	Basic PC	3.618	3.421	4.521	4.273	2.218	3.810	4.181	4.801
	PC with mean adj	1.586	1.609	1.693	1.759	2.360	1.583	1.916	1.573
	PC with recent adj	1.381	1.473	1.671	1.437	1.980	1.761	2.052	1.879
Logarithmic	Basic PC	1.823	2.047	2.811	2.323	2.005	1.671	2.222	1.920
	PC with mean adj	1.428	1.572	1.840	1.897	1.752	1.460	1.979	1.517
	PC with recent adj	1.260	1.419	1.636	1.390	1.698	1.578	1.866	1.742
Logistic	Basic PC	2.398	2.673	3.048	2.973	3.028	3.891	5.240	5.910
	PC with mean adj	1.729	1.579	1.765	1.671	2.707	1.583	1.859	1.617
	PC with recent adj	1.460	1.404	1.772	1.472	1.999	1.658	2.095	1.908
Difference	Basic PC	1.159	1.305	1.281	1.189	1.476	1.529	1.282	1.677
	PC with mean adj	1.445	1.424	1.830	1.446	2.089	1.728	2.225	1.843
	PC with recent adj	1.416	1.334	2.038	1.492	2.218	1.669	2.649	1.911
Difference of logarithms	Basic PC	1.196	1.349	1.399	1.225	1.691	1.593	1.558	1.676
	PC with mean adj	1.418	1.405	1.758	1.385	2.015	1.577	1.940	1.682
	PC with recent adj	1.487	1.350	1.775	1.461	2.068	1.582	2.178	1.779
Difference of logistics	Basic PC	1.134	1.207	1.205	1.189	1.377	1.396	1.198	1.758
	PC with mean adj	1.502	1.332	1.849	1.427	2.059	1.628	2.107	1.840
	PC with recent adj	1.916	1.306	1.832	1.498	2.214	1.624	2.290	1.902

a poor forecast. Use of the autocorrelation correction improved the performance for males although the performance for females remained poor. As with Tables 2a and 2b, differencing the data improved the forecasts markedly. Also as in Tables 2a and 2b the transformations did not adversely affect the performance of the forecasts, and have the advantage of restricting the range of forecasts to desirable intervals.

6. Long-Range Results

Unlike other government agencies, the Social Security Administration is deeply interested in the quality of the forecasts over a longer period. Specifically, this section considers forecasts 25 and 75 years into the future from the most recent available set of observations. Of course, we will not be able to compare the forecasts to actual held-out results, as in Section 5. We will, however, be able to make qualitative statements about the reasonableness of the forecasts; these statements will further influence the model selection.

For the accuracy of short-term forecasts, Section 5 showed that it is important to consider models based on differences of rates. Hence, this section only presents models based on differenced rates, although we still consider alternative transformations. As in Section 5, we consider the demographic cell to be the unit of analysis although we will often combine the forecasts from different cells to interpret results. Throughout, when combining forecasts based on different cells, we use the 1998 actual population. An alternative would be to use a forecast of the population. However, this would mean disentangling differences due to labor force participation and population. To keep the story simple, we use a single reference population with our forecasts, the 1998 actual population.

We begin by considering autoregressive models whose short-term performances were summarized in Tables 4a and 4b. Specifically, Figure 10 summarizes the 75-year forecasts of the autoregressive model with parameters that are cell-specific for males. Figure 10 shows that the forecasts of labor force participation rates without transformations can be both over one and less than zero. Further, forecasts based on logarithmic transformations can be over one. Thus, forecasts without any transformation or based on a logarithmic transformation require modification in order to be plausible. By using the logistic transformation, we constrain forecasts to lie between zero and one. Based on this logic, we henceforth consider only forecasts based on the logistic transformations.

Even with the logistic transformation, Figure 10 shows an undesirable feature of the autoregressive model with cell-specific parameters. That is, for the later working years, ages 55–74, we see that forecasts are jagged, in contrast to the smooth pattern in the actual 1998 rates. This erratic behavior is counterintuitive if one believes that labor force participation is a smooth function of ability to work as proxied by age. This point is emphasized in Figure 11 where this series is compared to forecasts based on an autoregressive model with common parameters. Intuitively, the jagged patterns in the cell-specific autoregressive forecasts are attributable to the fact that these forecasts use the time series pattern of each cell without regard to other cells that are close in age. In contrast, the forecasts based on common autoregressive parameters use the same estimates for updating each cell.

Thus, we see that the requirement of forecast rates lying between zero and one and the requirement that forecasts exhibit a smooth pattern across ages allow us to rule out some of the models presented in Section 5. Figures 12a and 12b summarize the forecasts from three

Table 4a. Short-term out-of-sample statistics for random effects longitudinal data models

Female

Response units	Forecasting technique	Mean absolute error Most recent in-sample year				Mean absolute percentage error Most recent in-sample year			
		1994	1990	1986	1982	1994	1990	1986	1982
Percent	Random effects	8.439	8.131	8.254	7.069	7.363	6.377	6.631	6.765
Logarithmic	Random effects	8.901	8.508	8.551	7.261	7.843	6.757	6.853	6.669
Logistic	Random effects	8.287	8.033	8.181	7.023	7.418	6.418	6.581	6.597
Difference	Random effects	1.518	1.685	1.974	1.830	3.274	3.504	3.727	3.709
Diff of logarithms	Random effects	1.604	1.746	2.096	1.800	3.004	3.181	4.037	3.219
Difference of logistics	Random effects	1.464	1.606	1.962	1.808	3.027	3.108	3.735	3.339
Percent	RE with AR(1)	3.828	4.277	5.116	5.456	4.177	3.911	4.793	5.540
Logarithmic	RE with AR(1)	5.329	5.714	6.172	5.896	5.015	4.722	5.437	5.530
Logistic	RE with AR(1)	3.832	4.280	5.046	5.339	4.185	3.891	4.815	5.249
Difference	RE with AR(1)	1.711	1.983	2.365	2.091	4.007	3.929	4.390	4.040
Diff of logarithms	RE with AR(1)	1.822	2.013	2.645	2.062	3.267	3.554	4.545	3.483
Difference of logistics	RE with AR(1)	1.589	1.884	2.388	2.090	3.248	3.429	4.175	3.594

Table 4b. Short-term out-of-sample statistics for random effects longitudinal data models

Male

Response units	Forecasting technique	Mean absolute error Most recent in-sample year				Mean absolute percentage error Most recent in-sample year			
		1994	1990	1986	1982	1994	1990	1986	1982
Percent	Random Effects	2.370	2.676	2.745	2.866	2.747	4.222	4.669	5.495
Logarithmic	Random Effects	2.300	2.610	2.670	2.810	2.503	3.928	4.388	5.278
Logistic	Random Effects	2.395	2.675	2.769	2.898	2.707	4.146	4.609	5.481
Difference	Random Effects	1.666	1.007	1.956	1.277	1.840	1.360	1.768	1.594
Diff of logarithms	Random Effects	2.140	1.252	2.403	1.635	1.825	1.282	1.827	1.761
Difference of logistics	Random Effects	1.518	1.002	1.769	1.140	1.837	1.337	1.768	1.559
Percent	RE with $AR(1)$	1.197	1.172	1.264	1.205	1.315	1.488	1.203	2.050
Logarithmic	RE with $AR(1)$	1.264	1.112	1.323	1.163	1.375	1.403	1.216	1.925
Logistic	RE with $AR(1)$	1.339	1.386	1.263	1.479	1.502	1.980	1.405	2.754
Difference	RE with $AR(1)$	1.748	1.068	2.029	1.231	1.910	1.392	1.909	1.653
Diff of logarithms	RE with $AR(1)$	2.131	1.183	2.305	1.483	1.865	1.320	1.912	1.778
Difference of logistics	RE with $AR(1)$	1.626	0.987	2.011	1.185	1.925	1.368	1.919	1.604

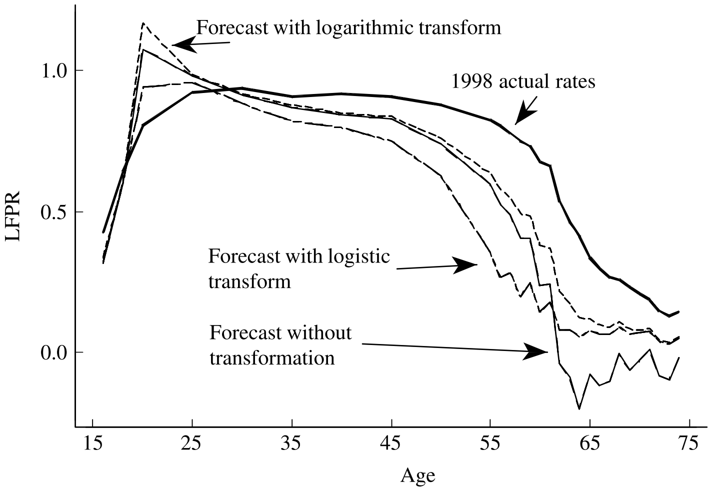

Fig. 10. Plot of 75-year forecasts of male Labor Force Participation Rates (LFPR) using the individual autoregressive model

competing models that satisfy these additional requirements: the autoregressive model with common parameters, the principal components method (without mean or bias adjustment) and the random effects model with a common autocorrelation parameter. We do not present the principal components method with mean adjustment; this model gives rise to the same pattern of jagged forecasts as the individual-cell autoregressive model because it is based on time series averages of individual cells. As noted earlier, the effect

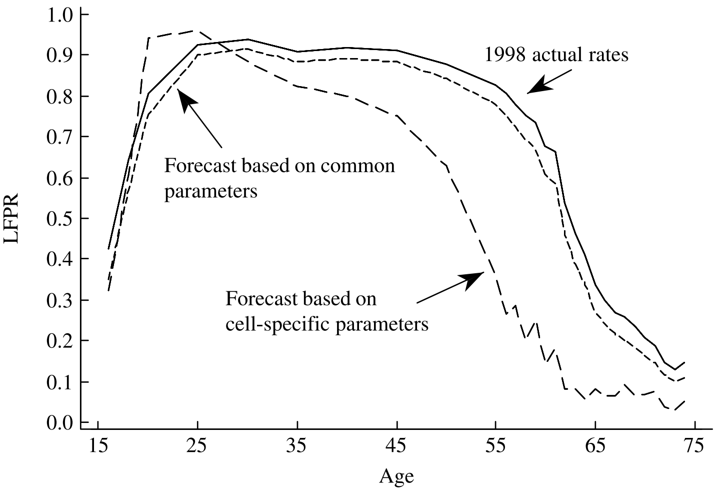

Fig. 11. Plot of 75-year forecasts of male Labor Force Participation Rates (LFPR). Both sets of forecasts are based on the autoregressive (AR) model. The AR model with cell-specific parameters exhibits jagged forecasts at the higher working ages when compared to actual 1998 rates and the forecasts based on the AR model with common parameters

of bias adjustment is minimal in models based on differenced data and thus we do consider this additional complication now. For the random effects model, we retained the autocorrelation parameter to allow for additional long-run patterns to emerge. Figure 12a summarizes the 75-year forecasts for females for each model; Figure 12b gives the corresponding forecasts for males.

Perhaps the most important conclusion that emerges from Figures 12a and 12b is that the choice of forecasting matters a great deal. Figure 12a shows that forecasts for the random effects and autoregressive models are close to one another, in contrast to the principal components forecasts. Conversely, Figure 12b shows that principal components and autoregressive model forecasts are close whilst the random effects model forecasts are more distant.

Table 5 summarizes the mean forecast of labor force participation for each forecasting method, over the 75-year period considered in Figures 12a and 12b, as well as 10-, 25- and 50-year forecasts. As before, the means are weighted by the 1988 actual population. From Table 5, we see that the principal components forecasts are flat over all forecast periods. Additional analysis shows that this is also true for each demographic cell, not just the overall mean. In contrast, both the autoregressive and random effects models extrapolate the female increasing and male decreasing labor force participation into the future, although at different rates.

Table 5 highlights another undesirable aspect of long-run projections. For forecasts based on the autoregressive and random effects models, male and female average labor force participation rates are about equal at the 25-year forecast horizon. As we saw in Section 3, male labor force participation rates have historically exceeded those for females, although the trends point to females closing the gap. The models with sex-specific parameters project closing the gap over the 25-year horizon and, continuing this trend, forecast female labor force participation rates exceeding male after this point. Although

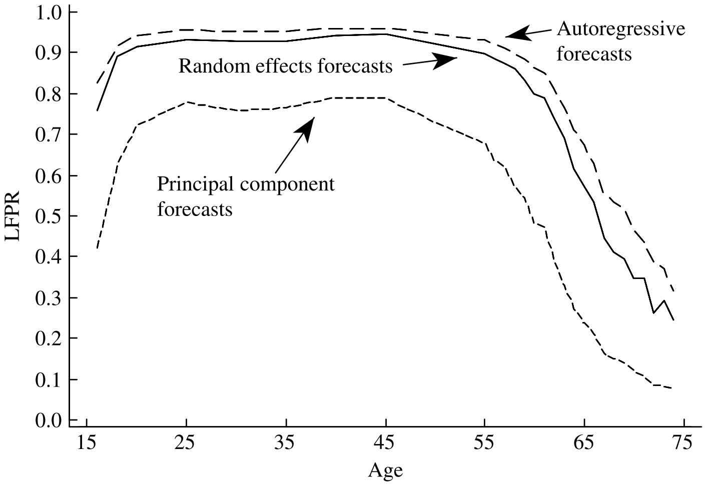

Fig. 12a. Plot of 75-year forecasts of female Labor Force Participation Rates (LFPR). The figure shows forecasts from each of the autoregressive, principal components and random effects models

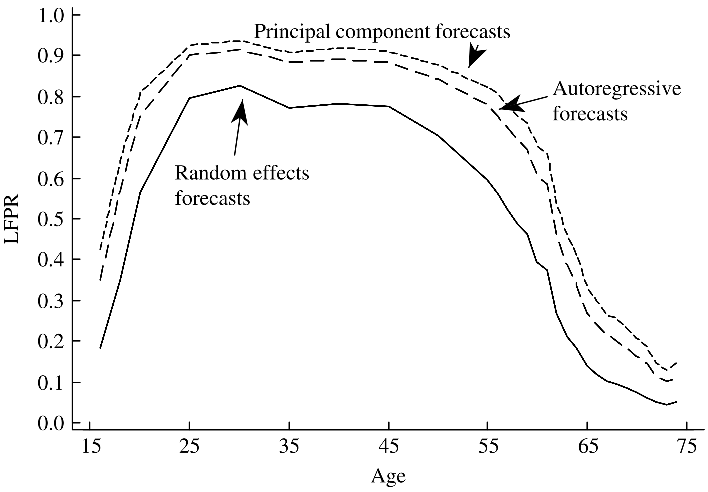

Fig. 12b. Plot of 75-year forecasts of male Labor Force Participation Rates (LFPR). The figure shows forecasts from each of the autoregressive, principal components and random effects models

this is certainly possible, the large differences in participation rates are certainly not a feature of the historical data upon which the forecasts are based.

As noted by Chu (1998), population dynamics derived solely from male vital rates and those derived solely from female vital rates will show ever-increasing differences with the passage of time. Of course, there are many possible modifications to each of the three basic models to mitigate this difficulty. The following section presents a modification of the autoregressive model that is simple and easy to apply.

Table 5. Mean Forecast Labor Force Participation Rate (Percentage)

Forecasting model	Forecast horizon			
	10 years	25 years	50 years	75 years
Females				
Autoregressive – common parameters	69.4	75.3	83.1	88.9
Principal components	65.0	65.0	65.1	65.1
Random effects – with autocorrelation parameter	68.4	73.2	79.9	85.2
Autoregressive – parameters varying by age group	66.0	67.4	69.5	71.4
Males				
Autoregressive – common parameters	77.9	77.2	75.9	74.5
Principal components	78.5	78.5	78.4	78.4
Random effects – with autocorrelation parameter	76.4	73.3	67.6	61.3
Autoregressive – parameters varying by age group	78.5	78.4	78.1	77.5

7. A Simple Model

This section describes the performance of a simple modification of the autoregressive model. Specifically, we now allow model parameters to vary by three age groups: less than 20, 20–54, and 55–74. Moreover, parameters vary by gender for the two older age groups, for a total of five sets of parameters. Section 3 showed that labor force participation rates for teens are currently similar, so the parameters (not the rates) were combined for the purposes of simplicity. Thus, the simple model combines features of the cell-specific and common parameter autoregressive models in Equations (4.3) and (4.4).

Table 5 summarizes forecasts from the simple model under the heading "Autoregressive – parameters varying by age group." Here, we see that the mean labor force participation rate increases for females and decreases for males as the forecast horizon increases. Figure 13 provides additional information about the vector of forecasts for the 75-year horizon. At 75 years, forecasts for male and female teens are nearly equal. Males in the primary working years, 20–54, have changed little since 1998 whereas females have increased, although still less than males. In the later working years, 55–74, the gap between males and females has closed so that differences are relatively small.

Figure 14 allows us to see the development of the forecasts over time. Figure 14 extends the Figure 8 times series plot of expected number of working years to include the forecast labor force participation rates (in addition to the 31 years of actual rates). Like Figure 8, Figure 14 shows the anticipated convergence in expected number of working years for each age group.

Thus, this simple model provides intuitively plausible long-range forecasts of future labor force participation rates. Although the details are not reported here, the model did as well on the short-range tests documented in Section 5 as the other models considered (although no model was uniformly superior to the others). The most compelling aspect of this model is its simplicity. Dynamics are handled through the differencing of rates and the autoregressive parameters. No arbitrary truncation of the rates is required because of the logistic transformation. The model does not require that we set male equal to female rates at some arbitrary time point in the future. Nor does it require that we stop using the data at some arbitrary time point in the future and use instead an "expert" opinion about future labor force participation rates. This is not to say that expert opinions are not desirable in many circumstances; we merely point out that a strength of the model is that it allows past experience to be our primary guide to forecasts of the future.

How reliable are the forecasts of this model? A companion article, Frees (2003), explores the forecasting accuracy of the model predictions. This article employs both a hierarchical (multilevel) model and a variant of a seemingly unrelated regression model to represent the complex variance structure of this high dimensional forecasting problem. With a model of both contemporaneous and dynamic sources of variation, the article develops forecasting prediction intervals. In part because of the natural bounds induced by the logistic transform, the forecast intervals turn out to be intuitively plausible for 10-year forecasts, such as those produced by the U.S. Bureau of Labor Statistics. Not surprisingly, for 75-year forecasts such as those required by the Social Security Administration (SSA), the bounds are uncomfortably wide (and are not explicitly provided in that article). Nonetheless, SSA is ultimately interested in providing accurate forecasts of the financial

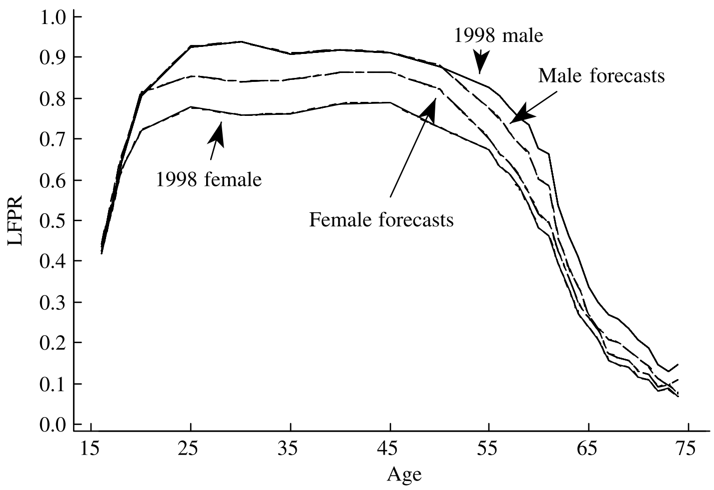

Fig. 13. *Plot of Labor Force Participation Rates (LFPR) versus age. The figure shows actual 1998 rates and 75-year forecasts. The forecasts are from the autoregressive model where parameters vary by age group and sex*

solvency of the Social Security system, not labor force participation rates. Of course, with other things being equal, a forecast of LFPRs with small forecast intervals is more useful than a forecast with large intervals. However, a forecast of LFPRs is just one of many components that are inputs to the forecasts of financial solvency, and the interaction of these forecasts with other forecasts would be critical for this type of forecasting exercise. The position of this article is that a stochastic model to forecast LFPRs is a critical input into a larger forecasting system for projecting Social Security funds; wide forecast

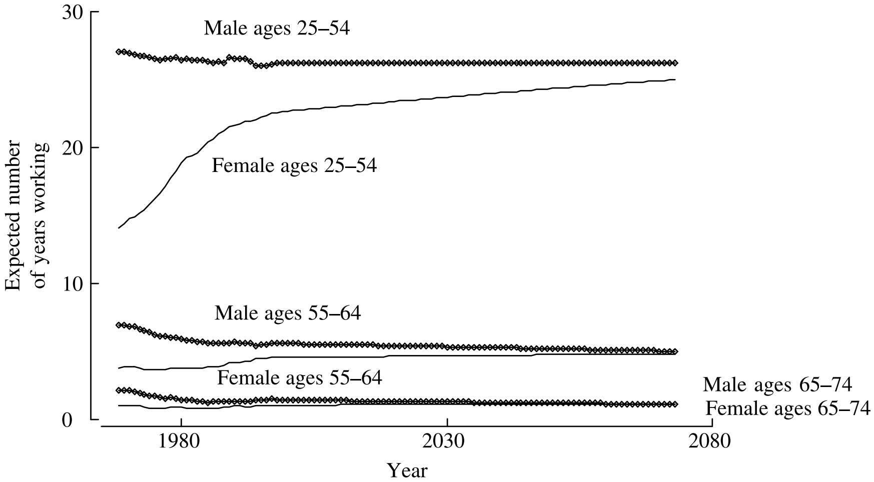

Fig. 14. *Plot of expected number of working years versus year. Based on actual LFPRs up to 1998 and forecasts after 1998. Female age groups are indicated by a solid line, male age groups are indicated by a solid line plus a circle*

intervals are a reflection of the data and the forecast horizon and do not mean that one should not produce stochastic forecasts. We note that the most recent government projection of the finances of the Social Security system (Board of Trustees 2003) does contain a stochastic projection model.

8. Summary and Conclusions

In this article, we have treated the problem of forecasting labor force participation rates as one might treat the forecasting of a demographic vital rate such as mortality or fertility. Although we recognize that labor force participation is an economic choice problem, one could make the same argument for fertility decisions and even, given the individual's lifestyle choices, mortality. The position of this article is to use statistical tools commonly employed in demography to forecast labor force participation rates by demographic cell; these forecasts can be used in Social Security as well as other government planning exercises.

By looking at short-term forecasting over alternative time horizons, we learned in Section 5 that it is useful to use the *change* in labor force participation rates as the unit of analysis. Section 6 on long-range forecasting showed us that a logistic transformation of the rates performs well; this transform ensures that forecast rates will lie between zero and one, thus avoiding any abrupt truncation of forecast rates. This article considered a variety of models based on past and current values of labor force participation rates, including autoregressive, random effects and principal components models. We documented that these models give a wide range of forecast values of labor force participation rates. Section 7 introduced an autoregressive model with parameters that are gender and age-group specific. This simple model relies only on past experience as a predictor of future behavior; thus, it is less sensitive to political intervention than alternatives.

Forecasting of cell-specific labor force participation rates is complex because the demographic cells are not independent of one another. To illustrate, labor force participation for females age 20–24 who were never married, one demographic cell, is not independent of the experience of females age 20–24 with the spouse absent, another demographic cell. Intuitively, this is one reason that cell-specific autoregressive models, which use individual parameters for each demographic cell, do not fare well in our analysis. In this article, we mitigated this problem by allowing different demographic cells to share common parameters. In a companion article, Frees (2003), complex models of the variance structure are developed in order to develop forecast intervals.

The dependence of demographic cells is also related to the issue of incorporating additional exogenous variables into the forecasting problem. That is, in Section 2, we reviewed the literature that suggests other potential variables that may influence future values of labor force participation rates. However, to assess the effect of these potential variables, we need to understand the standard errors of estimators that summarize this fit. Conventional standard errors are calculated assuming independence of demographic cells. Thus, an understanding of the relationships among demographic cells will also be useful in building a more complete forecasting model of labor force participation rates.

Appendix: Principal Component Models

A basic principal component model is of the form:

$$y_{it} = \beta_i k_t + \varepsilon_{it}, \quad i = 1, \ldots, C, t = 1, \ldots, T \tag{A.1}$$

Here, the parameter β_i describes the patterns of deviation by cell i and the term k_t is a time-varying stochastic process that does not depend on the index i. Unlike the longitudinal data models in Section 4.1, there are no observables on the right-hand side of Equation (A.1). For identifiability, we need to impose further restrictions and require that $\Sigma_i \beta_i = 1$ and $\Sigma_t k_t = 0$.

For additional notation, we let $\mathbf{y}_t = (y_{1t}, \ldots, y_{Ct})'$ be the vector of responses for time t. Let $\mathbf{Y} = (\mathbf{y}_1, \ldots, \mathbf{y}_T)'$ be the $C \times T$ matrix of responses. Further, let $\boldsymbol{\Lambda}$ be the $C \times J$ matrix corresponding to the first J eigenvectors of $\mathbf{Y}$. Specifically, the jth column of $\boldsymbol{\Lambda}$ is the eigenvector corresponding to the jth largest eigenvalue of $\mathbf{Y}$, for $j = 1, \ldots, J$. In our application, we use only the first principal component, so that $J = 1$.

With these restrictions on the model, the least squares estimators are easy to derive. It turns out that the least squares estimator of the vector $\{\beta_i\}$ is $\boldsymbol{\Lambda}$. Further, the least squares estimator of k_t is given by $\hat{k}_t = \boldsymbol{\Lambda}'\mathbf{y}_t$.

Given the least squares estimators, forecasting the series is now straightforward. We first use *ARIMA* time series methods to produce forecasts of $\{\hat{k}_t\}$; denote the L-step forecast as $\hat{k}_{T+L}$, for $L = 1, 2, \ldots$. We then use these forecast values to forecast the series with the Expression $\hat{\mathbf{y}}_{T+L} = \boldsymbol{\Lambda}\hat{k}_{T+L}$.

For estimating and forecasting age-specific mortality rates, Lee and Carter (1992) used the model

$$y_{it} = \alpha_i + \beta_i k_t + \varepsilon_{it}, \quad i = 1, \ldots, C, \ \ t = 1, \ldots, T \tag{A.2}$$

Adding the cell-specific intercept accounts for the average shape of the age profile. The least squares estimator of this parameter turns out be $\bar{y}_i$, the time series average over cell i. Thus, when forecasting, one first subtracts cell-specific means and then uses a one-dimensional principal components method. (This yields a new version of $\boldsymbol{\Lambda}$.) As another modification, Lee and Carter (1992) also required that the fitted numbers of deaths in year t equals the actual number of deaths when determining the estimator of k_t in lieu of the constraint that $\Sigma_t k_t = 0$. For forecasting labor force participation rates, we do not use a similar constraint in order to separate the LFPR forecasting from the population forecasting, although it is certainly a possibility.

A variation described by Bell (1997) involves subtracting the most recent series $\mathbf{y}_T$, in lieu of the cell-specific average. This produces no approximation error in the last year of data, implying a better position for short-run forecasting. Because this approach fared well in Bell's analysis of fertility and mortality rates, we document how this alternative fares in our Section 5 results section.

9. References

Ahlburg, D.A. (1982). How Accurate Are the U.S. Bureau of the Census Projections of Total Live Births? Journal of Forecasting, 1, 365–374.

Anderson, P.M., Gustman, A.L., and Steinmeier, T.L. (1999). Trends in Male Labor Force Participation and Retirement: Some Evidence on the Role of Pensions and Social Security in the 1970's and 1980's. Journal of Labor Economics, 17, 757–783.

Bell, W.R. (1997). Comparing and Assessing Time Series Methods for Forecasting Age Specific Demographic Rates. Journal of Official Statistics, 13, 279–303.

Blau, F.D. (1998). Trends in the Well-being of American Women, 1970–1995. Journal of Economic Literature, 36, 112–165.

Board of Trustees, Federal Old-Age and Survivors Insurance and Disability Insurance Trust Funds (2003). 2003 Annual Report of the Board of Trustees of the Federal Old-Age and Survivors Insurance and Disability Insurance Trust Funds, Washington, D.C.: Government Printing Office.

Carter, L.R. and Lee, R.D. (1992). Modeling and Forecasting U.S. Sex Differentials in Mortality. International Journal of Forecasting, 8, 393–412.

Chu, C. (1998). Population Dynamics: A New Economic Approach. Oxford University Press, New York.

Costas, D. (1998). The Evolution of Retirement. The University of Chicago Press, Chicago, IL.

Frees, E.W. (1999). Summary of Social Security Administration Projections of the OASDI System, Social Security Advisory Board. Paper available at http://www.ssab.gov/reports.html.

Frees, E.W. (2003). Stochastic Forecasting of Labor Force Participation Rates. Insurance: Mathematics and Economics, 33, 317–336.

Frees, E.W. (2004). Longitudinal and Panel Data: Analysis and Applications for the Social Sciences. Cambridge University Press, Cambridge.

Fullerton, H.N., Jr. (1999a). Labor Force Projections to 2008: Steady Growth and Changing Composition. Monthly Labor Review, November, 19–32.

Fullerton, H.N., Jr. (1999b). Labor Force Participation: 75 Years of Change, 1950–1998 and 1998–2025. Monthly Labor Review, December, 3–12.

Goldin, C. (1990). Understanding the Gender Gap. New York: Oxford University Press.

Killingsworth, M. and Heckman, J.J. (1986). Female Labor Supply: A Survey. In O. Ashenfelter and R. Layard (eds), Handbook of Labor Economics, Vol. 1, 103–205.

Lee, R.D. (2000). The Lee-Carter Method for Forecasting Mortality, with Various Extensions and Applications. North American Actuarial Journal, 4, 80–93.

Lee, R.D. and Carter, L.R. (1992). Modeling and Forecasting U.S. Mortality. Journal of the American Statistical Association, 87, 659–675.

Lee, R.D., Carter, L.R., and Tuljapurkar, S. (1995). Disaggregation in Population Forecasting: Do We Need It? and How to Do it Simply. Mathematical Population Studies, 5, 217–234.

Lee, R.D. and Miller, T. (2001). Assessing the Performance of the Lee-Carter Approach to Modelling and Forecasting Mortality. Working paper. Available at http://www.demog.berkley.edu/ ~ rlee.

Lee, R.D. and Tuljapurkar, S. (1994). Stochastic Population Forecasts for the United States: Beyond High, Medium and Low. Journal of the American Statistical Association, 89, 1175–1189.

Lee, R.D. and Tuljapurkar, S. (1997). Stochastic Forecasts for Social Security. In Frontiers in the Economics of Aging, D.A. Wise (ed.). Chicago, IL, The University of Chicago Press, 393–428.

Olsen, R.J. (1994). Fertility and the Size of the U.S. Labor Force. Journal of Economic Literature, March, 32, 60–100.

Perrachi, F. and Welch, F. (1994). Trends in Labor Force Transitions of Older Men and Women. Journal of Labor Economics. 12, 210–242.

Purcell, P.J. (2000). Older Workers: Employment and Retirement Trends. Monthly Labor Review, October, 19–30.

Quinn, J.F. and Burkhauser, R.V. (1990). Work and Retirement. In Handbook of Aging and the Social Sciences, R.H. Binstock and L.K. George (eds). San Diego, CA:Academic.

Rust, J. (1990). Behavior of Male Workers at the End of the Life Cycle: An Empirical Analysis of States and Controls. In Issues in the Economics of Aging, D.A. Wise (ed.). Chicago: The University of Chicago Press.

Smith, J.P. (ed.) (1980). Female Labor Supply: Theory and Estimation. Princeton, NJ: Princeton University Press,.

Smith, J.P. and Ward, M.P. (1985). Time Series Growth in the Female Labor Force. Journal of Labor Economics, January, 3, Supplement, S59–S90.

Thompson, P.A., Bell, W.R., Long, J.F., and Miller, R.B. (1989). Multivariate Time Series Projections of Parameterized Age-specific Fertility Rates. Journal of the American Statistical Association, 84, 689–699.

U.S. Bureau of Labor Statistics (1997). BLS Handbook of Methods. U.S. Bureau of Labor Statistics. Available at http://www.bls.gov/opub/hom/homtoc.htm.

Received July 2002
Revised March 2006