International Journal of Forecasting 9 (1993) 187–202

0169-2070/93/$06.00

Modeling and forecasting the time series of US fertility: Age distribution, range, and ultimate level

Ronald D. Lee

Demography and Economics, 2232 Piedmont Ave, University of California, Berkeley, CA 94720, USA

Abstract

To develop high-low bounds for population projections, traditional demographic forecasts assume fertility is always high or always low. Such bounds are related to bounds on average annual fertility up to year t rather than to individual year bounds in time series forecasts, and so the autocorrelation of time series forecast errors is important in practical applications. This paper develops methods for using time series methods to make constrained long term forecasts of fertility. Specifically, age–time variations in fertility are modeled with a single time-varying parameter, or fertility index; upper and lower bounds on the total fertility rate are imposed by forecasting an inverse logistic transform of the fertility index; the long run level of the fertility forecast is also constrained to equal a prespecified level. The principal interest is in the variance and the autocorrelation structure of the forecast errors. Based on these for the USA we conclude: (1) the probability interval for average fertility up to time t begins to contract after about 50 years, but only very slightly: (2) the probability interval for average fertility up to year 2065 is about three-fifths as wide as that for single year fertility in 2065, but is still far wider than the band for official forecasts; (3) realizations of the simple ARMA (1,0,1) forecast model exhibit long fluctuations something like actual fertility in industrial nations; (4) the model of fertility age patterns fits poorly at older ages, but may be adequate for present purposes.

Keywords: Fertility; Demographic; Projection; Forecasts; Constrained

1. Introduction

Many methods have been used to forecast fertility–trend extrapolation of period or cohort fertility rates, fertility expectations gathered through surveys, the application of social and economic theory, and statistical time series analysis (which might be viewed as a variant of extrapolative methods). Many of these methods and their accompanying refinements are ingenious, and there have been some isolated successes.[1] On the whole, however, the state of our knowledge and understanding in this important

This research report was prepared as part of a project on 'Modeling and Forecasting Demographic Time Series', supported by a grant from NICHHD: RO1-HD24982.

[1] For example, Easterlin's socio-economic theory successfully predicted the baby bust, and also predicted a substantial upturn in the eighties. It is too early to tell whether the upturn of 1988–1990, nipped by the recession, is consistent with his prediction. In the late 1950s, Freedman et al. (1959) predicted the baby bust based on their analysis of survey data on fertility expectations. With the possible exception of Easterlin's theory, however, none of the methods has proved consistently useful.

area is discouraging, despite the substantial resources that have been devoted to it. This paper discusses some of the difficulties encountered by time series methods, and suggests some new ways to deal with them. It emphasizes long term forecasts, for which the problems in using time series methods are acute.

Demographers are interested in forecasting fertility mainly as an input for forecasts of population size and age composition. It is increasingly recognized that population forecasts should provide well-based indications of their own reliability. Therefore it is important not only to develop sound point forecasts of fertility, but also to develop probability intervals for these forecasts.

One might think that it would be sufficient to provide high and low bounds, along with their accompanying probability coverage. Indeed, such information would be valuable. However, identical probability bands could have very different implications for the uncertainty of population forecasts. Consider two examples. In traditional demographic forecasts, the high and low bounds on fertility or mortality represent levels that are maintained every year over the range of the forecast. The corresponding high and low population scenarios are then calculated by consistently applying either the high or the low vital rate forecasts. This traditional approach represents one polar extreme, in which forecast errors are perfectly correlated over time: if the fertility forecast is wrong in some way one year, it is assumed to be wrong in the same way every year thereafter. For this reason, there is absolutely no cancellation of errors over time, and the probability bounds for the average fertility over the first T forecast periods will be just as wide as the original bounds. It is a very different matter if errors are completely independent from one year to the next, since there would then be substantial cancellation of errors over time.[2] For an average of the fertility forecast over T periods, the probability bounds would be reduced by a factor of one over the square root of T, and uncertainty about the average will go to zero as T increases.[3]

As long term population growth depends (approximately) on long run average fertility rather than on its level in any particular year, it is very important to find out which characterization of errors in fertility forecasts is more accurate. Time series methods are the natural way to approach this problem.

Furthermore, the shape of the fertility interval in the traditional forecasts is almost certainly wrong. That interval initially opens wider as the horizon lengthens, as our uncertainty grows rapidly as we move away from the present. It then reaches a plateau where it remains thereafter. If the intervals were intended to bound fertility itself in each individual year, then intervals of this shape would be appropriate. However, they are actually intended to provide for reasonable *population size intervals* in the future. As I argue in detail below, this means that traditional fertility levels must instead refer to something like the average level of fertility up to each given horizon. Uncertainty about such an average will first increase, then stabilize, *then decrease*. The decrease occurs because the number of terms being averaged increases with the horizon length, and so the average should approach more closely the true mean level of fertility.[4] Time series methods can also shed light on this issue.

Use of modern time series methods to forecast fertility dates back to the early 1970s, and there have been many contributions in the intervening years. Some analysts chose to model and forecast the series of annual births directly [see Saboia (1977)]. Lee (1974, 1981) argued for instead forecasting some summary fertility index such as the Total Fertility Rate. The forecast

[2] At the other polar extreme, forecast errors would be perfectly negatively correlated, and cancel perfectly. This case appears even more unlikely than the other two considered, however.

[3] This assumes that the forecast is known to be unbiased; if not, the uncertainty does not go to zero.

[4] The error in the traditional forecast of fertility for any particular year has two components: an error in the guess of the ultimate level or the true mean (bias), and an error due to the deviation of fertility in that particular forecast year from the ultimate or true mean level (intrinsic error). In the forecast of the level of fertility averaged over the first T years of the forecast, the intrinsic error will go to zero as T gets large, and the overall error will therefore come to equal only the error in forecasting the ultimate or true mean level, that is the bias. For this reason, the error band should eventually narrow. If fertility is a white noise process, then the period of initially increasing uncertainty does not occur, because the last observed value provides no information about future values. Realistically, however, fertility is autocorrelated rather than white noise.

level of fertility could then be distributed by age according to some rule based on past patterns. McDonald (1981), while taking the event count (births) as the entity to forecast, introduced the use of transfer function models, and a greater level of sophistication in the time series analysis itself. Alho and Spencer (1985) introduced a number of new features, particularly more careful attention to several aspects of the correlation structure of errors, and integration of time series methods with use of expert opinion. Carter and Lee (1986) developed joint forecasts of a nuptiality index and a marital fertility index, leading to forecasts of births and marriages with joint confidence intervals. The US Census Bureau (1989) projections from base year 1988 used time series methods for the first 10 or 15 years of the long run fertility forecasts. Their method is a blend of time series analysis and principal components, and involves forecasting the level and age distribution of the rates in an integrated fashion [see Bozik and Bell (1987, 1989)].

Any approach to forecasting fertility, including time series methods, must confront a broad range of difficulties and requirements. (1) The fertility of each age group behaves idiosyncratically over time, yet when the age-specific time series are viewed together, there is a striking similarity in their behavior. (2) Biology bounds fertility above at a TFR of about 16, and its definition bounds it below at zero. Forecasts should respect these constraints. (3) Our prior knowledge about fertility in the USA, and the institutions, customs and preferences which affect it, place far narrower bounds of plausibility on fertility. This information should to some degree be incorporated in the forecast, while recognizing the limitations of our knowledge and imaginations. (4) These last two considerations are at risk as soon as we propose a stochastic model with a normally distributed error term, as such distributions allocate some probability to every positive and negative outcome, even though some are known to be impossible. (5) Fertility rates in one year are rarely very different from those in the preceding or following year, yet the level may change substantially over the course of a decade. Such features of the series should also be expressed in the forecast. (6) The time series of US fertility exhibits an apparent cycle with period about 40 years. We must decide whether to view this as an aberration or an integral part of the series. (7) Some trends that may have contributed to the secular decline in fertility have largely run their course. This is true of declining infant and child mortality and perhaps of improving contraceptive technology. For this reason, extrapolation of past trends may be inappropriate. (8) At the same time, some of these changes, such as that in mortality and contraceptive technology are very unlikely to be reversed, so levels of fertility observed in the past should not automatically be viewed as plausible or even possible in the future–contrary to the statistical assumption of stationarity. Many efforts to forecast fertility ignore some of these aspects of the problem.

In addition to these general problems facing any forecasting procedure, there are also some which afflict time series methods in particular. Box and Jenkins (1976) time series methods are well-suited for short run forecasts, but not for forecasts over long horizons, such as the 75 year horizon of many official US forecasts. Standard procedures typically lead us to fit ARMA models to first differences of the basic series. But when integrated for forecasting purposes, such models generate forecast intervals which widen rapidly and without limit. The forecast may also drift steadily away from the initial value without limit, if the fitted (differenced) model includes a constant term. Yet other models, which are often statistically indistinguishable, involve a first order autoregressive term slightly less than unity, and behave radically differently in the long run: probability intervals stabilize, and an equilibrium level is approached. Furthermore, the constant term in such fitted models may have a large standard error and be insignificantly different from zero (the case for the USA). Yet the constant term determines the trend or equilibrium in these models, and therefore plays a critical role in shaping the long run behavior of the forecasts. Attempts to use time series models to make long term forecasts of fertility face these difficulties in addition to those listed earlier.

This paper develops several new methods for dealing with these problems [see Lee (1992) for a related discussion]: (i) a simple way to construct and estimate a one parameter family of fertility schedules which allows the age pattern of fertility to change; (ii) the straightforward derivation of forecasts and associated confidence intervals based on this family; (iii) a simple way to incorporate prior information about bounds on fertility into the forecast through use of a logistic transform; and (iv) a way to introduce prior information about the ultimate level of fertility into the time series model.

2. A one parameter model of changing age patterns of fertility

It is important that fertility forecasts be disaggregated by age, so that they can be used in conjunction with population age distributions to generate forecasts of numbers of births. Yet we do not want to generate independent age-specific forecasts, which would be time-consuming and would overlook their strong statistical interdependence. The approach taken here is to model and fit a one-parameter family of fertility age distributions, and then to forecast that single parameter. From its forecasts, the forecasts of the individual age specific rates can be recovered using the model. In Lee and Carter (1992), this approach is applied successfully to model and forecast mortality.

An alternative approach is to use two or more parameters to model the fertility age distribution and then to forecast each parameter singly or all parameters jointly. Knudsen et al. (1993) fits a four parameter model schedule. Such an approach involves fitting fewer age varying parameters, and is in that sense more parsimonious. However, it also requires forecasting more time varying parameters, in principle jointly, and in that sense is less parsimonious. Because the parametric representations of the age schedules are usually highly non-linear, there is some danger that the forecasts will be unstable over long forecast horizons. To derive probability intervals for the fertility forecasts, it is necessary to take into account the non-linear effects of their joint probability distributions, and this is not generally done.

Let the matrix $\{f_{x,t}\}$ represent the collection of fertility rates for age x at time t. The problem is to fit this matrix of data with a model in which only one parameter varies with t. In earlier work [see Lee (1974)], it was assumed that the matrix could be modeled as: $f_{x,t} = f_t \times m_x$, where f_t was the total fertility rate (TFR) and m_x was a proportional age distribution which summed to unity. In this model, all age specific fertility rates move up and down together proportionately, so the proportional age pattern of fertility is invariant.

The model used in this paper is a slight generalization, in which the age pattern of fertility can change when the level changes. In particular

$$f_{x,t} = a_x + f_t \times b_x + e_{x,t} \,. \tag{1}$$

Deviations of age specific fertility from the base schedule a_x are now assumed to move in proportion to one another, rather than the rates themselves. A logarithmic version of the same model will also be considered:

$$\ln(f_{x,t}) = a_x + f_t \times b_x + e_{x,t} \,. \tag{2}$$

In this latter model, negative birth rates cannot occur, regardless of the level of the index f_t. It also has the advantage of producing assymetric probability bands, bounded below by 0 but not above.

These models, particularly the logarithmic one, are very similar in structure to the Coale–Trussell (1974) 'Big M, little m' model, where b_x is similar to their age-pattern-of-control variable, and f_t is similar to their variable for strength of control, m.

Note that these representations of $f_{x,t}$ are not unique. For any given a_x, b_x, and f_t, there is an identical representation with a_x, cb_x, f_t/c for arbitrary c. Likewise, $a_x - cb_x$, b_x, $f_t + c$ form an equivalent representation. In what follows, the b_x will be standardized to sum to unity, and the f_t to sum to zero (over the sample range). With this choice, it can be shown that the a_x will simply equal the average age-specific values over time of fertility, $f_{x,t}$, and the sum of the a_xs, call it A, equals the average value of the TFR over the sample period. Summing both sides of Eq. (1), we find that $\text{TFR}_t = A + f_t + E_t$, where E_t is the sum across age of the $e_{x,t}$s. E_t should be close to zero, but will not generally equal zero. Consequently, to a close approximation, the fertility index in the linear model, f_t, is the deviation in period t of the TFR from its long term average, A.

3. Fitting the model to US data

The models of Eqs (1) and (2) cannot be estimated using ordinary regression procedures, because there are no observed variables on the right hand side. The singular value decomposition [SVD; see Good (1969); Wilmoth (1990)] can be used to find a least squares fit for the matrix of $f_{x,t}$ after the age-specific means, a_x, have been subtracted. The first right and left vectors and leading value of the SVD provide a family of solutions, from which the normalization described above selects a unique one.[5]

It is usually desirable to do a second pass estimate of f_t, by choosing the value of f_t which, given a_x, b_x and a population age distribution for each year, implies exactly the right number of aggregate births. In this way, values of f_t can be calculated for years in which age specific fertility data are not available, while total birth counts and population age distributions are. In the USA the second pass estimate extends the f_t series to the years 1909–1916, and also up to the most recent year for which total births are available—typically several years later than the latest age specific fertility data. This second pass estimation of f_t also has the advantage that it produces fertility schedules which are consistent with the total birth and population age distribution data. For the linear model, second pass estimates of f_t should differ only slightly from the first pass SVD estimates. However, if the log version of the model is used, second stage estimates of f_t may differ substantially and are therefore mandatory.

For the USA, age specific fertility rates for ages 15–45 are available from 1917 to 1987.[6] Both the linear and the logarithmic versions of the model were estimated. The logarithmic version in effect fits proportional derivations in the data, and therefore gives equal emphasis to variations in rates that are absolutely large and small. Consequently, the fitted f_t has a strong downward secular trend in order to fit the secular trend in fertility at older ages, resulting in a poor fit at other ages. For the logarithmic model, $f_t + A$ gives a very poor fit to the TFR over the sample period. For these reasons, the remainder of the analysis is limited to the linear model.

Table 1 shows the SVD estimates of a_x and b_x. As noted earlier, the sum of the a_x, which here is 2.6325, is simply the average level of the TFR over the period 1917–1987. It is informative to plot the b_x coefficients together with a_x/A. If the size of the shared component of period variations in age specific fertility, b_x, is proportional to the mean level of fertility, a_x, then the two plots would be identical. These plots are shown in Fig. 1, and it can be seen that b_x and a_x/A are indeed very similar, with fertility at ages 20–24 somewhat more than proportionately volatile, while fertility at ages 25–29 is somewhat less than proportionately volatile. No consistent pattern is evident, and the proportionality assumption is surprisingly good in this instance. For a Third World population with data covering the fertility transition, the results would surely be quite different.

The fitted model tracks fertility at some ages quite well, and at other ages rather poorly. The proportions of variance accounted for in each of the age specific birth rates, for age groups 15–19 through 40–44, are 0.73, 0.85, 0.96, 0.88, 0.55, and 0.31. Combining fertility at all ages, the model accounts for 0.82 of the total variance about age specific means. Evidently the fit is best in the prime reproductive years, and worst for the youngest and oldest ages. With this data set,

Table 1
Estimated values of a_x and b_x coefficients for US women, 1917-1987, by SVD

Age group	a_x	b_x
15–19	65.05	0.1038
20–24	159.70	0.3482
25–29	143.11	0.2421
30–34	90.75	0.1598
35–39	50.17	0.1083
40–44	17.72	0.0379

[5] An alternative which yields a close approximation, should software to estimate the SVD be lacking, is first to estimate the a_xs as the age specific means of $f_{x,t}$ [or $\ln(f_{x,t})$]. Then subtract these from the matrix of $f_{x,t}$, to form a matrix of age specific rates with zero-means. An estimate of f_t is obtained by summing these rates across ages for each period t of the matrix $\{f_{x,t} - a_x\}$. Then the b_x are obtained by regressing, separately for each age x, the series $f_{x,t} - a_x$ on f_t, where the regression should contain no constant term.

[6] Age specific Current Population Survey data are available through 1990, but the registration data have a longer publication lag.

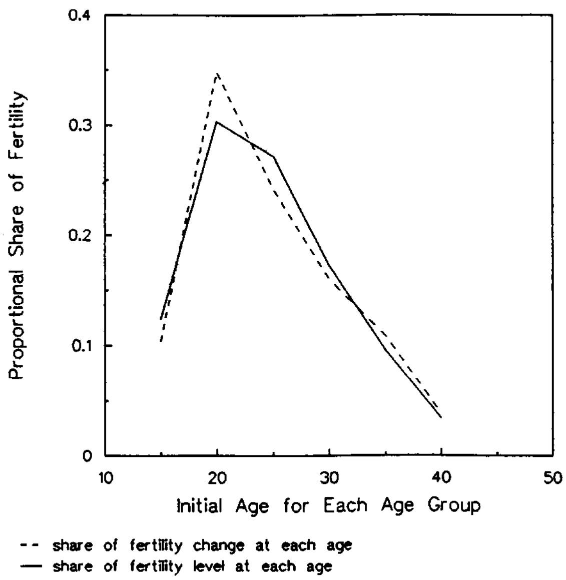

Fig. 1. Estimated age pattern of fertility change compared with level (US data, 1909–1987).

the poor fit at some ages is unavoidable for any one-parameter family of fertility schedules, because similar general levels of fertility have occurred in conjunction with very dissimilar birth rates at the older ages—as in the 1930s and more recent years, or the baby boom years and earlier in the century. Put differently, fertility at the older ages has been trending downward independently of the general level of fertility, until the mid-1970s, when it began to trend upward while the general level remained quite flat. More than one time-varying parameter is required to describe such changes, but that would involve new complexities and risks.[7]

Although fits for these individual ages are only so-so, the sum of the fitted age specific rates accounts for 0.999 of the variance in the actual Total Fertility Rate. This is not surprising, given the way the fertility index, f, is calculated in the second pass.

[7] Bozik and Bell (1987) use the first few principal components to forecast fertility; this amounts to using a several parameter approach. The difficulty is that one must then forecast the several parameters jointly, and derive confidence bands that depend on their joint variability. Also, there is a risk of projecting transitory aspects of the age patterns into the future. The recent up-turn in old age fertility might be extrapolated, even though it may, in reality, reflect catching-up behaviour of a temporary sort. Over the shorter horizons of Bozik and Bell, this last problem would not arise, however.

4. Incorporating prior constraints by means of a transform

The next step is to model f_t as a time series, and then to forecast it. In the remainder of the paper, it will be convenient to work with $F_t = f_t + A$, where A is the sum of the a_x values. F_t is the fitted value of the TFR, while f_t fluctuates around 0. The difficulty with forecasting F_t directly, however, is that nothing in the usual modeling procedure ensures that negative fertility rates do not occur. Nor is there any guarantee that forecasts will be demographically plausible, or even biologically possible. As an example of the difficulty, consider the forecast arising from standard Box–Jenkins procedures. After first-differencing the TFR, we are led to the model (0,1,1) with no constant. The point forecast is constant at 2.03, which is plausible. However, by 2065 the 95% probability intervals is −1 to +5 children per woman. The negative rates are, of course, impossible, and the rates over 3.5 seem far more unlikely that this model indicates.

Because of these difficulties, it is helpful to incorporate prior information and constraints directly in the modeling process. In this section, I will incorporate prespecified lower and upper bounds on the TFR, denoted L and U, using a procedure first suggested and applied by Alho (1990, p. 524) in his measure of fertility 'volatility', and subsequently developed independently for forecasting fertility in this paper. The bounds L and U for the TFR could be chosen very liberally as 0 and 16 births per woman, or more structure could be imposed by choosing narrow bounds, such as 1.0 and 3.5, for example.

This approach could be put on a more rigorous statistical basis by assuming that our prior knowledge of the constraints is not absolute, but rather leads to some probability distribution for the constraints. In this case, estimation could be carried out using Bayesian methods. The final outcome would then be a probability-weighted mixture of the prior information and the new information from the statistical analysis. If the evidence in the actual fertility data were sufficiently strong, it could overrule our prior information. Bayesian methods for ARMA models were developed in Monahan (1983). Here, however, I have taken the simpler approach of assuming fixed prior constraints.

Define a transformed fertility index, g_t, by

$$g_t = \ln\{(F_t - L)/(U - F_t)\}\,. \qquad (3)$$

The series g can then be modeled and forecast. The forecast of F is obtained by transforming back to F from g. Solving eq. (3) for F_t, we find

$$F_t = (U \times \exp(g_t) + L)/(1 + \exp(g_t)) \qquad (4)$$

which is a logistic transform. It can be seen that as g goes to infinity, F goes to U, the upper bound on fertility. As g goes to negative infinity, F goes to L, the lower bound on fertility. The forecast and its probability interval will, by construction, fall within these limits. This ensures that the long run forecasts will be relatively sensible, and also allows us to assume that errors in the transformed process g are normally distributed, although the implied errors in the F process will not be.

If the error in the forecast of g is symmetrically distributed, then the point forecast of g will be both the expected value of g and its median value for the relevant year. However, unless the point forecast of g happens to be midway between L and U, the error distribution of F will not be symmetric. When the point forecast of F is derived as the inverse transformed point forecast of g, then it will still be the median value of the distribution of F for that year. However, the point forecast of F will not typically be the expected or average value of F. Then we have the choice of redefining the point forecast of F to be the expected value of F, and calculating it; or retaining the inverse transform of the forecast of g as the point forecast. Here we will pursue the second option, which is simpler, although a case could certainly be made for calculating the expected value of F. Fortunately, simulation results to be reported below indicate that in practice, there is very little difference between the two.

5. Constraining the ultimate level of the forecast

In practice, even imposing bounds of the sort just described may be inadequate, in the sense that fitted models may still produce point forecasts for the distant future that are inconsistent with our best guesses. For example, the otherwise preferred time series model, constrained as just described, predicts an ultimate level of 2.2 for the TFR in the middle of the next century. A model which is nearly indistinguishable in performance over the sample range generates a forecast of 1.5 for the TFR. Differences of this magnitude in the point forecast of fertility exert a powerful influence on the point forecasts of population size and age distribution in the next century. This experience suggests that time series methods, even constrained in this way, cannot provide us with very useful point forecasts of fertility far in the future.

It is a simple matter to constrain a model to equilibrate at a prespecified level, which is the level to which the forecast will converge in the long run. Consider a general ARMA model for the fertility process F_t. The equilibrium value, F^*, is given by the constant term divided by one minus the sum of the autoregressive coefficients, provided this sum is not unity. Therefore, the constant term must equal F^* times one minus the sum of the autoregressive coefficients. After making this substitution, the model can be rewritten as a process for $y_t = F_t - F^*$, with a constant term equal to zero. The suggested procedure, then, is to form the new series $y_t = F_t - F^*$; model this series as an ARMA with no constant term; forecast y_t over the desired horizon; and then recover the forecasts of F_t by adding F^* to the forecasts of y_t.

In the forecasts reported below, this device is combined with the logistic transform. Lower and upper limits on the TFR are taken to be 0 and 4.0, and F^*, the ultimate level of the TFR, is taken to be 1.85, which is the average of the USCB assumption of 1.8, and the Social Security Actuary's 1.9.[8] These were chosen in the belief that they reflect a consensus view among demographers.

One might well wonder about the point of fitting a stochastic model, when so much of the outcome is determined by assumption. There are two important benefits from doing so. First, we get an estimate of the uncertainty of the forecast.

[8] This is done by solving for the value of g, G^*, which is the appropriate transform of F^*, where $G^* = \ln[(F^* - L)/(U - F^*)]$.

Note that in this application, the uncertainty is considerably less than the upper and lower bounds on fertility, and so it is not simply imposed by the assumptions chosen. Second, we get an estimate of the degree of autocorrelation in the errors in forecasting fertility. Taken together, these two benefits have many implications and amply justify the enterprise. When fertility forecasts such as this one are used as inputs to a stochastic population projection [see Lee and Tuljapurkar (1991)] the autocovariance information is an extremely important ingredient. It can be used to assess the extent to which the traditional High–Low Scenario bounds can be expected to differ from consistent bounds on annual variations. It can be used to assess the extent to which the traditional High–Low bounds should be expected to narrow in later years of a forecast. It can be used to calculate the degree to which quinquennial or decadal probability bands should be narrower than those for annual rates. It can be used to examine the extent to which random long swings are to be expected in fertility, and their amplitudes.

While it might be thought that the uncertainty would be understated by imposing a prespecified equilibrium value on the fertility model, owing to uncertainty about the prespecified level, the matter is not so simple. It is true that the long run forecast of the *average* value of fertility will eventually collapse on the assumed equilibrium level, as the variance of the long run average forecast goes to zero with an increasing forecast horizon. As that assumed equilibrium level will generally be wrong, the true error variance will not collapse to zero, and so the forecast constructed in this way will eventually, at long enough horizons, understate the true uncertainty. However, prespecification of the 'mean' eliminates one fitted coefficient, and therefore inevitably leads to a worse fit over the sample period, and larger residual errors. For this reason, forecasts for individual years must have somewhat larger estimated standard errors relative to forecasts based on an empirically estimated equilibrium value. It is therefore likely that the estimated standard errors of the forecast will initially be larger if the forecast is based on an assumed equilibrium level rather than on an empirically estimated mean. Thus forecasts based on prespecification of the mean overstate the true intrinsic uncertainty, but understate the uncertainty arising from bias.

6. Extensions of this approach

Rather than assuming that fertility is constantly tending towards a fixed prespecified equilibrium as above it may be more appropriate to assume that the equilibrium level of fertility is itself a stochastic process trending over time. (In this context, the equilibrium level is to be interpreted as the normal level at any instant, as distinct from the actual level which might reflect various kinds of temporary distorting influences.) 'Structural' time series models of this sort are coming into wider use [see Harvey (1989)]. Any attempt to model and forecast fertility in populations which have not completed their fertility transitions prior to the sample period must certainly incorporate a trending equilibrium, and almost all Third World populations fall into this category. Even in industrial nations, 20th century trends in mortality, contraceptive technology, women's roles and education indicate that trending equilibria may be necessary. For Third World populations, we might specify a logistic function for the equilibrium of the process, starting at around six children per woman and ending at perhaps two. For industrial nations, it might be enough to allow for an exponential decline in equilibrium fertility toward a lower asymptote of about two children, as the earlier plateau period for fertility is so long past (unless one used data series stretching back into the 19th century). There is a deterministic version of this approach, in which fertility is assumed to follow a prespecified pattern of decline to a prespecified ultimate level, typically at a net reproduction rate of unity. This approach is used in international population projections by the United Nations (1992, pp. 6–13) and the World Bank [see Vu et al. (1988, pp. 6–8)], and apparently originated with Frejka (1973).[9] The model developed in this section might be viewed as a stochastic version of this approach.

The equilibrium level of fertility, c_t, can be

[9] The World Bank [see Vu et al. (1988, pp. 6–8)] assumes various kinds of 'geometric curve' change in fertility to an ultimate level of replacement, for example.

viewed as a first order autoregressive process

$$c_t = \alpha + \beta c_{t-1} + u_t \, .$$

This process, with $0 < \beta < 1$, will tend to the ultimate level $\alpha/(1-\beta)$. This can be taken equal to the prespecified final equilibrium of, say, 2.0.

Actual fertility will not typically equal the equilibrium, however. A variety of short term influences cause it to deviate, including economic and political conditions, shifts in fashions regarding childbearing, and wars, among others. The 'Easterlin Hypothesis' gives one account of how long term fluctuations about equilibrium might arise. Each of these influences just mentioned tends to persist over time, but nonetheless changes more rapidly than does the equilibrium itself. These deviations from current equilibrium, given by $F_t - c_t$, can themselves be viewed as forming a first order autoregressive process, so that

$$(F_t - c_t) = \gamma(F_{t-1} - c_{t-1}) + e_t \, .$$

From inspection, it is clear that c_t can be re-expressed as an infinite sum of its disturbances, u_{t-i}, with weights equal to β^i, plus $\alpha/(1-\beta)$. Likewise, it is clear that F_t can be written as c_t plus the infinite sum of e_{t-i} with weights γ^i. Combining these, we can write

$$F_t = \alpha/(1-\beta) + \sum [\beta^i u_{t-i} + \gamma^i e_{t-i}],$$

where the sum is from 0 to infinity.

Thus in this specification, fertility is the sum of two infinite moving averages of disturbances, one with a shorter memory and one with a longer memory. The unconditional expectation of f is, of course, just $\alpha/(1-\beta)$. Our hypothesis is that the variance of u will be relatively small, and indeed might be zero; in this case, the equilibrium would decline exactly exponentially towards its long term level.

We can derive a different representation of the same process in single equation form as follows. Note that $c_t - \beta c_{t-1} = \alpha + u_t$, while $(F_t - c_t) - \gamma(F_{t-1} - c_{t-1})$ equals e_t. Subtracting $\beta(F_{t-1} - c_{t-1})$ from $F_t - c_t$, and substituting from the above expressions, this yields

$$F_t = \alpha(1-\gamma) + (\gamma + \beta)F_{t-1} - \gamma\beta F_{t-2} + u_t - \gamma u_{t-1} + e_t - \beta e_{t-1} \, .$$

As before, we can constrain this long run equilibrium to be some value we judge to be plausible, such as two children per woman, or more generally, F^*. To do so, we need only subtract F^* from each F_t, and omit the constant from the equation. We then have

$$(F_t - F^*) = (\gamma + \beta)(F_{t-1} - F^*) - \gamma\beta(F_{t-2} - F^*) + u_t - \gamma u_{t-1} + e_t - \beta e_{t-1} \, .$$

In this equation, there are only two coefficients to be estimated. In addition, however, we must estimate the variances of e and u. The equation cannot be estimated using ordinary procedures, because β and γ occur in the moving average terms for the errors. Kalman filter techniques are called for.

7. The fitted model

In this paper, estimates will be presented for the equilibrium-constrained and logistic-transformed model discussed earlier, but not for the structural model just discussed. A time series model is estimated for the transformed index g, constrained in such a way that the total fertility rate must always lie between 0 and 4, with an equilibrium median value of 1.85. The fitted model for g, the transformed index (standard errors in parentheses), is

$$(g_t - \ln(1.85/2.15)) = \underset{(0.104)}{0.9701} \times (g_{t-1} - \ln(1.85/2.15)) + u_t + \underset{(0.013)}{0.4042} \times u_{t-1} \, ,$$

$$R^2 = 0.96 \quad \text{S.E.} = 0.1618 \text{ Ljung–Box prob} = 0.28 \, .$$

Using this equation, it is straightforward to generate forecasts and probability intervals for g. Using Eq. (4) these can easily be converted into forecasts and intervals for F_t and $f_t = F_t - A$. These can then be used to generate forecasts and intervals for the individual age specific rates, if desired, using eq. (1).

8. Fertility forecasts

Figure 2 shows the point fertility forecasts, with probability intervals, from 1990 to 2065. The point forecasts decay rapidly to the prespecified equilibrium level of 1.85; this decay would, of course, be more obvious if the initial fertility level had not been so close to 1.85 in 1989. Without the constraint, the forecast would have risen asymptotically to an equilibrium level of about 2.3 children per woman. The probability intervals can be seen to grow very rapidly. A visual comparison with the pace of past episodes of change, during the 1920s and 1930s, the baby boom, and the baby bust, confirms that this rapid widening of the intervals is consistent with the past volatility of fertility. The shape of the intervals indicates that most useful information from initial fertility levels is used up within 20 years of the base period. The intervals eventually converge to the range 0.5–3.4 children.[10] In other words, the forecast asserts that towards the middle of the next century, there is less than a 2.5% chance that fertility will attain baby boom levels *in any given year*, and likewise less than a 2.5% chance that it will fall as low as 0.5 *in any given year*. Of course, the chance that it will surpass these bounds in *some* year during this period is far greater.

To explore the effects of constraints on U, L and F^*, alternatively specified models were estimated, and the results are shown in Table 2. The model just discussed, which is the preferred model, is the first listed. The other versions differ with respect to U, L, or F^*, including variants in which some or all of the constraints are dropped entirely. It can be seen that the (approximate) standard errors range from 0.56, for the model in which the TFR is constrained to lie between 1 and 4, to 0.73 for the preferred model. The completely unconstrained ARMA model (1,0,1) falls in the middle of this range, while the two partially constrained models lie at the upper end. In the completely unconstrained ARMA, the fertility level is 2.27 after 75 years; when the range is constrained to (0,4), with unspecified equilibrium, fertility after 75 years is 2.0. Otherwise, fertility simply moves to the prespecified level. These results suggest that the width of the probability interval is not highly sensitive to the L, U constraints unless these are made quite narrow. One of the narrower intervals is found for the case when $U-L$ is greatest, and the interval for the unconstrained ARMA is

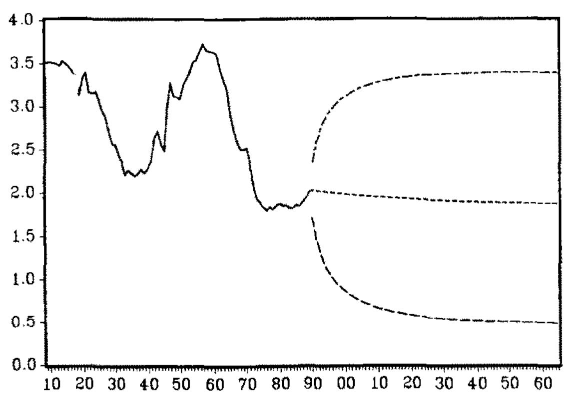

Fig. 2. Forecast of the US TFR based on logistic transform with $U=4$, $L=0$, and $F^*=1.85$. The forecast model is (1,0,1) with no constant.

Table 2
Comparison of forecasts of TFR and their approximate standard errors for different model specifications

Model spec			Forecast horizon (years)			
L	*U*	*F*	5	10	25	75
4	0	1.85				
TFR			2.00	1.98	1.93	1.87
S.E.			0.43	0.55	0.67	0.73
4	0	2.50				
TFR			2.12	2.20	2.35	2.49
S.E.			0.42	0.52	0.61	0.62
4	1	1.85				
TFR			2.01	1.99	1.94	1.87
S.E.			0.33	0.42	0.52	0.56
5	0	1.85				
TFR			2.01	1.99	1.94	1.87
S.E.			0.32	0.43	0.56	0.63
4	0	–				
TFR			2.08	2.12	2.21	2.30
S.E.			0.42	0.53	0.63	0.66
–	–	1.85				
TFR			2.01	1.99	1.94	1.87
S.E.			0.33	0.45	0.61	0.71
–	–	–				
TFR			2.07	2.11	2.18	2.27
S.E.			0.33	0.44	0.56	0.62

In models with U and L specified, the standard error of the forecast of the TFR is approximated by dividing by four the difference between the upper and lower 95% probability band.

[10] Changing the prespecified value of F^* changes the ultimate value of the point forecast, but does not have much effect on the confidence intervals.

narrower than any of the (0,4) versions.[11] It is clear from examining the fits of the ARMA models that imposing an equilibrium of 1.85 causes very little deterioration. Imposing limits of (0,5) causes further deterioration which is still very slight, while imposing the (0,4) constraint causes more substantial deterioration.[12]

From these comparisons it appears that if one only needed a 0.95 probability interval, simply constraining the equilibrium would be adequate, even though unrealistic or impossible values receive positive probability (see the second to the last model). For long term stochastic population forecasts or stochastic simulations, however, where the effects of impossible outliers are not well known, it would be wiser to use the more complicated nonlinear models.

9. Forecasts of age specific fertility

The forecasts of age specific rates can be calculated from the forecasts of the fertility index, f, using Eq. (1) and the coefficients in Table 1. Table 3 presents the results for selected years, through 2065.[13] Because the fitted model did not do a good job of tracking age pattern changes in past years, it is helpful to compare the age patterns forecast here with those forecast by the USBC (1989, p. 21). USBC forecasts an ultimate mean age of fertility of 26.9 years, while the rates for 2065 in Table 3 imply a mean age of 27.0 years, essentially the same. However, comparison with the age pattern of the USBC (1989, p. 129, middle assumption) shows that the rates forecast in this paper are higher in the younger and older ages, and lower in the middle years of childbearing. The Social Security Office of the Actuary's ultimate age pattern (Wade, 1989, p. 6, alternative II) is slightly younger, with a mean age of 26.4 years.

The most striking contrast in Table 3 is between the birth rates forecast for older women in this paper, on the one hand, and the far lower rates forecast by both USBC and Social Security, on the other. However, even given the context of socioeconomic change in the USA, it is not clear how age patterns of fertility will evolve in the future. It could well be that some women will find it optimal to complete their childbearing very early, and then have an uninterrupted work career, while others find it optimal to postpone childbearing until later ages, in order first to

[11] This would definitely not be true for a (–,1,–) model with a constant.

[12] Let the relative size of $(1 - R^2)$ for the completely unconstrained model be 1.00. Then the other models, with their relative sizes for $(1 - R^2)$, are as follows: (–, –, 1.85) 1.00; (0, 5, 1.85) 1.03; (0, 4, 1.85) 1.25; (0, 4, –) 1.25. Clearly imposing 1.85 makes little difference, while imposing an upper limit of 4 rather than 5 or none, makes a substantial difference.

[13] Forecast errors for the age groups arise from errors in estimating a_x and b_x, from the equation error $e_{x,t}$, and from errors in forecasting F_t. For a detailed analysis of the additional uncertainty arising from other sources [see Lee and Carter (1992, appendix 2)].

Table 3
Forecasts of age specific fertility rates and the total fertility rate for the USA, 1990–2065 (per 1000 women)

Year	15–19	20–24	25–29	30–34	35–39	40–44	TFR
1990	52	117	114	71	37	13	2024
2000	51	114	111	70	36	13	1978
2010	51	112	110	69	35	13	1945
2020	50	110	109	68	35	12	1929
2030	50	109	108	67	34	12	1901
2040	50	108	107	67	34	12	1888
2050	49	107	107	67	34	12	1878
2065	49	106	106	66	34	12	1868
USBC	40	99	116	75	26	4	1800
Soc Sec	55	111	112	72	26	4	1900
1990 CPS	40	113	112	80	37	9	1958

See text for method of calculation. Ultimate level is assumed to be 1.85, with bounds of $U = 4$ and $L = 0$. 'USBC' figures are projections under the Middle assumption for years after 2050, taken from US Bureau of the Census (1989, p. 129). The 'Soc Sec' figures are projections under Alternative II for years after 2013, and are calculated from Wade (1989, p. 6). The 'CPS' figures are taken from US Bureau of the Census (1991, p. 12).

become well-established in their careers. The 1990 Current Population Survey provides fertility estimates for the 12 months preceding June 1990, shown on the last line of Table 3. The rates at ages 35–39 are actually higher than those forecast here, and those for 40–44 are midway between the USBC/Social Security forecasts and those presented in this paper.

10. Probability bounds for the TFR

It will seem to many that the probability intervals for fertility in these forecasts are so wide as to be virtually useless. They appear particularly wide in contrast to the usual 'high' and 'low' brackets provided by official forecasts. However, it is essential to keep in mind that the probability intervals given here apply to fertility levels in a single year, not to the average level over a long time period. Even over a 5 year period, there will be some cancellation of forecast errors, so that the confidence intervals should be somewhat tighter for forecasts of 5-yearly rates than for annual rates.

Consider a population growing from period 0 to period T with a stochastic annual growth rate in year t of n_t. Then the population at time T, P_T, would be given by: $P_T = P_0 \Pi(1 + n_t)$. Rewriting this in terms of logs: $\ln(P_T) = \ln(P_0) + \Sigma \ln(1 + n_t)$. From this we can see that P_T depends only on the average over $t = 0$ to $T - 1$ of the logs of $(1 + n_t)$. But the average of the logs is not the same as the log of the average of $(1 + n_t)$, so knowing the average rate of growth is not really sufficient. Nonetheless, to a first approximation it is the average rate of growth which is relevant. The average rate of growth can in turn be approximated from average fertility and average mortality. It is for this reason that this section emphasizes the distribution of the long run average level of fertility, which is far more relevant for the distribution of future population sizes than is the distribution of annual levels of fertility.[14]

Here, stochastic simulations are used to assess the uncertainty in forecasts of the *average* level of fertility.[15] Figure 3 shows the average result from ten sets each of 1000 simulations, each for

[14] In addition, we should take into account the fact that the distribution of the average population growth rate depends on both fertility and mortality, and their covariation. The 95% confidence band for population size should be derived from the joint probability distribution for long run average values of fertility, mortality and migration, and will not correspond to the 95% probability interval on fertility. The exact calculation would be very complex, and we would do better to turn to fully stochastic population forecasts as in Lee and Tuljapurkar (1991).

[15] These could be calculated analytically from the estimated model, but the calculation is cumbersome.

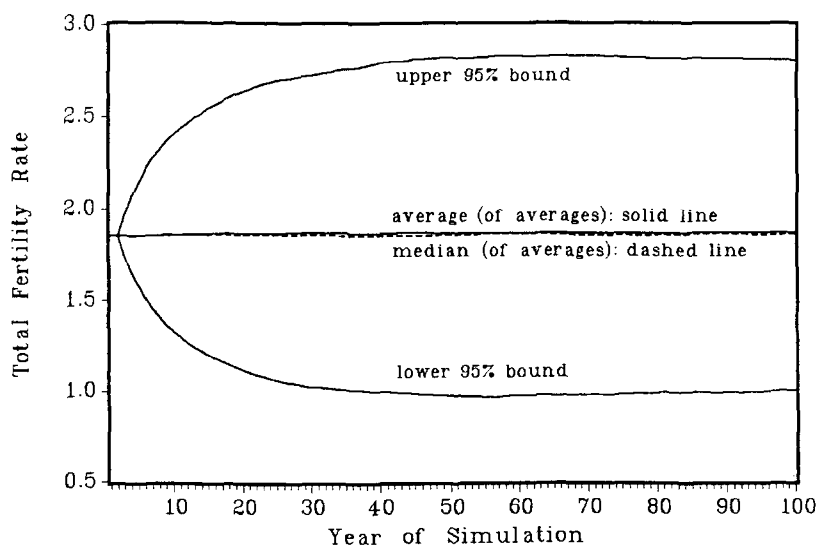

Fig. 3. Stochastic simulations of cumulated average TFRs, based on the fitted model (1,0,1) with bounds 0–4 and ultimate level 1.85. Based on the average of 10 sets of simulations, each with 1000 simulated series. The number plotted for year t refers to the average value of all simulated TFRs up to and including year t. See text.

100 years, of the estimated (1,0,1) model with no constant, constrained to an ultimate level of 1.85, and constrained to fall between 0 and 4. The figure shows the mean, median, and upper and lower 95% probability bounds for averages of the first t years of simulated fertility, for each t. For example, for $t = 73$, the figure shows the average of the first 73 values for 10 000 stochastic simulations. The plot shows the average, median and upper and lower 95% intervals for these 10 000 values of the average of the first 73 realizations.

Table 4 presents summary data for a 75 year horizon, corresponding to the forecast from 1990 to 2065. The intervals are clearly considerably narrower than for fertility in individual years (1.8 versus 2.9 years), consistent with the calculation above. However, these are still more than twice as wide as the intervals used by the Census Bureau (0.7) and by Social Security (0.6), as can be seen from the table. While the Census Bureau attaches no explicit probability to its high–low interval, Stoto (1983, p. 17) concludes from an analysis of their past record and that of other developed countries that these intervals should cover the true outcome about two-thirds of the time.[16] Most of the uncertainty in past projections has come from fertility rather than mortality or migration, so we can halve our 95% interval to get one more comparable with the Census

Table 4
Forecast intervals for long run average fertility

	Low	Medium	High	Hi–Lo
Social security	1.6	1.9	2.2	0.6
Census Bureau	1.5	1.8	2.2	0.7
Present model (Av.)	1.0	1.85	2.8	1.8
Present model (Ind.)	0.5	1.85	3.4	2.9

(Av.) refers to 95% probability intervals containing the average level of fertility from 1990 to 2065; (Ind.) refers to 95% probability intervals containing the level of fertility in the individual year 2065.

[16] Stoto's analysis of the US record alone suggested that the high–low intervals would be right substantially less than two thirds of the time; inclusion of other developed countries which had less volatile fertility, raises his appraisal of the accuracy of official projections. His analysis of US data covers far shorter horizons than the forecasts in this paper, so the comparison made in the text is shaky. For a further discussion, see US Bureau of the Census (1989, pp. 14–15).

interval. This gives us an interval of 0.9 children, which is much closer to the Census interval of 0.7.

It was pointed out earlier that intervals for long run average values should actually contract at longer horizons, as more terms are being averaged together. In Fig. 3 we can see that this occurs. However, we can also see that the tendency is very slight, because the autocorrelation in fertility is so high. The fan width narrows from 1.86 after 60 years to 1.80 after 100 years. For most practical purposes, there is no harm in the traditional assumption that interval fans are non-decreasing. The same simulations also show that there is very little difference between the median and the average level of fertility, with the median typically less by about 0.01 children per woman. For practical purposes, this is negligible.

11. Dynamic behavior of the fitted model

During the sample period 1909–1989, fertility in the USA exhibited long and deep swings, similar to those in many industrial populations. By contrast, the forecast of fertility shown in Fig. 2 is almost perfectly flat: because the estimated time series model is first order autoregressive, it has no tendency to cycle. However, specific realizations of models of this sort do exhibit long wave-like fluctuations. In a sense, the fitted model does predict long swings for the future, but it makes no prediction about when they will occur; consequently, they average out to zero in the point forecasts.

Figure 4 shows three different realizations of the estimated model, each 100 years long, generated by the same kind of stochastic simulation used for Fig. 3. The actual US TFR series as shown in Fig. 2 is smoother, and its cycle more regular than these simulations. However, the simulations imitate the historical behavior far better than the point forecast does. When the fitted fertility model is incorporated in a full-blown stochastic population forecast, as in Lee and Tuljapurkar (1991), these realistic features are critically important in generating the estimates of population and age group size dispersion, and the probability distributions of quantities such as the old age dependency ratio. These features derive from the autocovariance

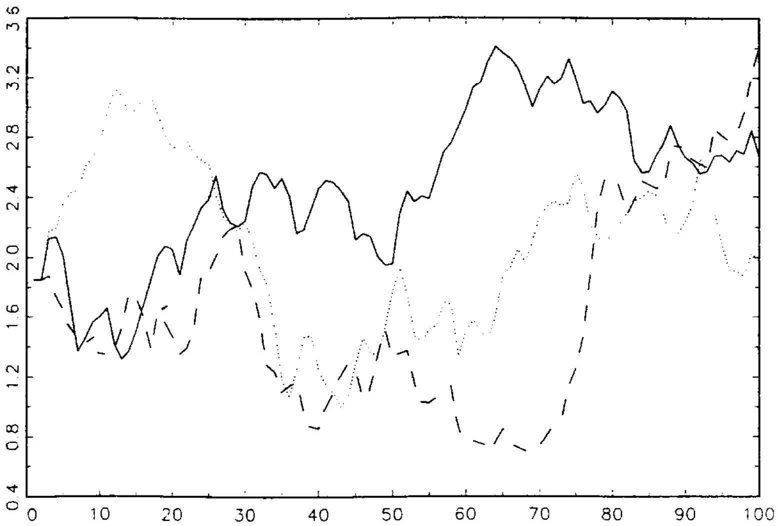

Fig. 4. Three stochastic simulations of the fitted fertility model. The model is the same one used for Fig. 5 and 6; that is, it uses the logistic transform with $U = 4$ and $L = 0$; the equilibrium level is constrained to be 1.85; and the model is (1,0,1) with no constant. See text for the actual equation.

structure of the modeled series, which is invisible in the plot of point forecasts and probability intervals shown in Fig. 2.

12. Discussion

Statistical time series analysis provides a powerful set of proven tools for modeling and forecasting many kinds of series. The generic approach is sufficiently flexible to generate good short term forecasts for series arising from many kinds of substantive processes, including fertility. However, the methods are not well suited to longer run forecasting of the sort that demographers are often called upon to perform. Also, these methods are most naturally applied to a single series or only a few series, and the multiplicity of highly correlated age-specific rates requires special treatment.

Here it was shown how first order regularities in the age–time variations of fertility may be efficiently captured by two parameters for each age group, with only a single time-varying parameter. Modeling and forecasting can then be restricted to that single parameter. In practice, the fit to US age-specific fertility variations was rather poor at older ages, where variations were not closely linked to the overall level of fertility. In applications to Third World populations still in the process of transition, this approach may capture more of the variance as the typical age pattern of natural fertility moves toward the typical pattern of controlled fertility, while the overall level of fertility declines.

In forecasts of mortality, the age distribution of rates is critical to the value of life expectancy at birth, and so uncertainty about the age specific rates is important and affects the projected population age distributions. In forecasts of fertility, however, what matters most is to get the implied number of births right, and this depends far more on the TFR than on its allocation over age. The fitted model accounted for 82% of the variance in age specific fertility rates about their average levels. However, age specific errors were strongly correlated in such a way that 99.9% of the variance in the TFR was accounted for. The fact that the model did not account better for age variations in the US does not greatly reduce its value in forecasting fertility.

It was also shown how fertility outcomes generated by the model can be constrained to lie within a range chosen by the analyst based on prior information, so that neither negative nor otherwise implausible values can occur. Finally, it was shown how the model can be constrained to move ultimately to a prespecified level of fertility chosen by the analyst, again based on prior information. Point forecasts generated in these ways are unlikely to deviate widely from the middle or medium trajectory of official forecasts, since similar information is used in similar subjective ways. The payoff comes in the probability intervals attached to the forecasts, and in the autocovariance structure of the forecast errors, without which the probability intervals are useless. These estimates can be used, in conjunction with other methods and models, to produce fully stochastic population forecasts.

The probability intervals for the fertility forecasts reveal once again how little we know about the future evolution of fertility over the medium and long run. Within 8 years, the standard error in the preferred model is 0.5 years, so that the 95% probability interval ranges from one to three children per woman. This interval may seem to overstate the uncertainty, but it merely reflects the rapidity of changes that have occurred in the sample period. Planners should keep in mind the volatility of fertility, and its past history of long swings.

Earlier, two extreme situations were contrasted: if forecast errors are perfectly correlated over time, then there would be no tendency for forecast errors to be offsetting, and fertility averaged over a number of years would have the same probability bounds as fertility in a single year. This is one way of viewing the basis for the traditional approach to scenario-based population projection. The fan of probability bounds for fertility averaged up to time T would then have horizontal edges. By contrast, if forecast errors are independent, then errors cancel with averaging, and the width of the fan would decline in proportion to the square root of T. According to the analysis here, the fan for the average of fertility forecasts narrows so slightly at longer horizons, that its edges are nearly horizontal. At the say time, there is substantial averaging of errors, so that the width of the fan for fertility averaged up to time T is only about 60% of the width for individual forecast years. It was also shown that specific realizations of the simple model fitted here exhibit long, deep fluctuations rather like those observed in the fertility of many industrial populations including that of the USA, even though the trajectory of mean fertility is very flat. Under the traditional scenario-based procedure, the consequences of such fluctuations are overlooked when possible variations in dependency ratios are calculated, for example. It is critically important that their probability be accurately reflected in any stochastic population forecast.

Time series methods are unlikely to add much accuracy to our long run point forecasts of the level of fertility. They are too sensitive to minor alterations in model specification, and they ignore too much extraneous information about fertility and its biological and socio-economic context. However, they can tell us a great deal about the uncertainty of our forecasts, and whether forecast errors can be expected to cancel over time or to accumulate. This information is essential to the development of stochastic population forecasts, and to our understanding of the way that errors are propagated in population projections.

References

Alho, J.M., 1990, "Stochastic methods in population forecasting", *International Journal of Forecasting*, 6, 521–530.

Alho, J. and B. Spencer, 1985, "Uncertain population forecasting", *Journal of the American Statistical Association*, 80, 306–314.

Box, G.E.P. and G.M. Jenkins, 1976, *Time Series Analysis: Forecasting and Control*, rev. ed. (Holden-Day, San-Francisco).

Bozik, J.E. and W.R. Bell, 1987, "Forecasting age specific fertility using principal components", *Proceedings of the American Statistical Association*, Social Statistics Section, 396–401.

Bozik, J.E. and W.R. Bell, 1989, "Time series modeling for the principal components approach to forecasting age-specific fertility", Paper presented at the *1989 Meetings of the Population Association of America*.

Carter, L. and R.D. Lee, 1986, "Joint forecasts of US marital fertility, nuptiality, births and marriages using time series models", *Journal of the American Statistical Association*, 81, 902–911.

Coale, A. and J. Trussell, 1974, "Model fertility schedules: variations in the age structure of child bearing in human populations", *Population Index*, 44, 185–258.

Freedman, R., P.K. Whelpton and A.A. Campbell, 1959, *Family Planning, Sterility, and Population Growth* (McGraw-Hill, New York).

Frejka, T., 1973, *The Future of Population Growth: Alternative Paths to Equilibrium* (Wiley, New York).

Good, I.J., 1969, "Some applications of the singular decomposition of a matrix", *Technometrics*, 11, 823–831.

Harvey, A.C., 1989, *Forecasting, Structural Time Series Models and the Kalman Filter* (Cambridge University Press, Cambridge).

Knudsen, C., R. McNown and A. Rogers, 1993, "Forecasting fertility: An application of times series methods to parameterized fertility schedules", *Social Science Research*, 22, 1–23.

Lee, R.D., 1974 "Forecasting births in post-transition populations: Stochastic renewal with serially correlated fertility", *Journal of the American Statistical Association*, 69, 607–617.

Lee, R.D., 1981, "Comment on McDonald: Modelling demographic relationships: An analysis of forecast functions for Australian live births", *Journal of the American Statistical Association*, 76, 793–795.

Lee, R.D., 1992, "Stochastic demographic forecasting", *International Journal of Forecasting* 8, 315–327.

Lee, R.D. and L. Carter, 1992, "Modeling and forecasting the time series of US mortality", *Journal of the American Statistical Association*, September.

Lee, R.D. and S. Tuljapurkar, 1991, "Stochastic population projections for the US: Beyond high, medium and low", Paper presented at the *1991 Annual Meetings of the Population Association of America*, Washington DC.

McDonald, J., 1981, "Modeling demographic relationships: An analysis of forecast functions for Australian births", *Journal of the American Statistical Association*, 76, 782–792.

Monahan, J.F., 1983, "Fully Bayesian analysis of ARMA time series models", *Journal of Econometrics*, 21, 307–331.

Saboia, J.L.M., 1977, "Autoregressive integrated moving average (ARIMA) models for birth forecasting", *Journal of the American Statistical Association*, 72, 264–270.

Stoto, M., 1983, "The accuracy of population projections", *Journal of the American Statistical Association*, 78, 13–20.

United Nations, Department of International Economic and Social Affairs, 1992, *Long-range World Population Projections: Two Centuries of Population Growth 1950–2150*, United Nations Sales Number E.92.XIII.3 (New York).

United States Bureau of the Census, 1989, Current Population Reports, Series P-25, No. 1018, *Projections of the Population of the United States, by Age, Sex and Race: 1988 to 2080*, US Government Printing Office, Washington, DC.

United States Bureau of the Census, 1991, Current Population Reports, Series P-20, No. 454, *Fertility of American Women: June* 1990, US Government Printing Office, Washington, DC.

Vu, M.T., E.Bos and R. Bulatao, 1988, *Europe, Middle East, and North Africa Region Population Projections* (1988–89 edition), Working Papers of the Population and Human Resources Department of the World Bank (October) WPS 117.

Wade, A., 1989, *Social Security Area Population Projections 1989*, Actuarial Study No. 105 (US Department of Health and Human Services, Social Security Administration, Office of the Actuary).

Wilmoth, J. R., 1990, "Variation in vital rates by age, period, and cohort" in: C. Clogg, ed. *Sociological Methodology: 1990*, 20, 295–336.