BUREAU OF THE CENSUS

STATISTICAL RESEARCH DIVISION REPORT SERIES

SRD Research Report Number: CENSUS/SRD/RR-88/20

CALCULATION OF THE VARIANCE OF POPULATION FORECASTS

by

William W. Davis

Statistical Research Division
Bureau of the Census
Washington, D.C. 20233

Recommended by: Nash J. Monsour

Report completed: August 16, 1988

Report issued: August 16, 1988

Report revised: October 1, 1988

CALCULATION OF THE VARIANCE OF POPULATION FORECASTS

William W. Davis
Statistical Research Division
Bureau of the Census
Washington, D. C. 20233

June 1988

Abstract

In this paper we propose a recursive method for forecasting populations and calculating the uncertainty in the estimate. In a recent survey paper Land (1986) discusses three widely used classes of methods for national population forecasts. They are demographic accounting equations, statistical time series, and structural modeling methods. This paper combines the first two of these techniques. An advantage of the time series approach is that forecast variances can be derived.

In this paper we assume that time series methods have been used to model and to forecast fertility, survival, and migration rates. Either univariate (Box and Jenkins, 1970) or multivariate (Tiao and Box, 1981) autoregressive-integrated-moving average (ARIMA) time series models could be employed.

Assuming that means and variances have been calculated for future values of the fertility, mortality, and migration series this paper shows how to use the accounting equations to calculate the mean and the variance of future populations.

1. Introduction

We assume that data are available for n years (labelled i=1,2,..,n) and d+1 age categories (labelled j=0,1,..,d) for both sexes (labelled s = f (female) or s = m (male)). We follow the demographic convention of using sex as a subscript before the variable and age and category as standard subscripts. The notation for the raw data is given in Table 1.

Table 1: Definitions of the Yearly Statistical Data

symbol	number with sex=s, category=j, period=i
$_{s}y_{ij}$	population
$_{s}b_{ij}$	births
$_{s}d_{ij}$	deaths
$_{s}u_{ij}$	emigrants
$_{s}w_{ij}$	immigrants

The problem of interest is to use the data to forecast future values of population. In the notation of Table 1 we wish to forecast $_{s}y_{n+k,j}$ for k=1,2,...,L, j=0,...,d and s=f and m, where L is the forecast horizon.

We assume that the data are collected at equally spaced intervals. We choose the category length to be the time between collections. For definiteness we will assume 1 year intervals. The jth age category is defined to include all people who were j

years old at the beginning of the period.

Most of the time series modelling of annual demographic data has focused on fertility and/or births (see, e.g., Land, 1986, Section 3.2). The use of univariate Box and Jenkins' analysis for modelling fertility rates goes back, at least, to Lee (1974). Lee models an aggregate of the different fertility rates and concludes that the series is not stationary. Then he derives confidence limits for future forecasts. Since the confidence limits grow so quickly, he rejects the model! A statistician might draw quite different conclusions from this analysis.

Demographers have long stratified populations by age and sex (and frequently race). Sample survey research has shown that stratification is beneficial in estimation if populations with consistently different behavior on the quantity of interest can be identified (e.g.,Cochran (1963, Chapter 5)). The same principle applies to forecasting populations.

The classical demographic approach for forecasting the total number of births in a population is to represent it as the sum over mutually exclusive exhaustive categories (i.e., strata), which are formed using the mother's age. The number of births in each age category are then forecast and summed to give an aggregate forecast. Of course, the number of births in the individual age categories may be of interest in their own right.

The forecast in each age category can be made in several ways. We will discuss the following two methods:

(i) time series approach: model and forecast the births in a category using data on the number of births in the category for previous time periods.

(ii) demographic approach: forecast both the number of women and the birth rate in the category. The product of these two gives the forecast for the number of births in the category.

The change in the number of women in a cohort may be obtained using the accounting identities

population change = immigrants - (deaths + emigrants).

The time series forecast is based solely on information obtained from cohorts <u>other</u> than one to be forecasted. The demographic approach uses information on previous cohorts to forecast the birth rate, but uses information on the cohort in question to forecast the number of women. Since the death rate is low for women of child bearing ages, the number of women in a category can be predicted quite accurately. The time series forecasts of the number of births may be slow to respond to the changes, which are solely caused by changes in the distribution of women in the child bearing years (and not by fertility changes). We will follow the demographer's approach, but we assume that time series methods have been used to forecast birth rates in the strata.

These two approaches were contrasted in McDonald (1981), which was concerned with forecasting the number of (first) births to Australian women. He developed a transfer function model

using economic leading indicators, but the variables had little predictive capability. McDonald's approach was criticized by Lee (1981) and Long (1981) for failure to use any of the ideas of the demographic approach such as the accounting equations, birth rates, and stratification.

Keyfitz (1977, p.24) has studied empirically whether stratification is useful in forecasting populations. Although there is some gain, he concludes that "...we should not be under the illusion that projection by age and sex is a powerful technique for discerning the future."

Probably the most important single variable to forecast is the total population, which is the sum of the components of the population distribution. Cohen (1986) gave an approach to this problem, which is valid asymptotically under weak assumptions. In population forecasting the horizon may be longer than the time series so it is unclear whether asymptotic techniques are relevant.

The accounting equations can be used with both projections and forecasts to produce future population distributions. In projection, future values are <u>assumed</u> for fertility, mortality, and migration. In forecasting, the future values of the three rates are <u>predicted</u> based on previous data.

Currently, the U. S. Census Bureau projects high, medium, and low values for each of the three rate (see, e.g., Spencer and Long, 1983). These projections are based on birth expectations

from current surveys as well as information from historical data. The high and low values are not claimed to be percentiles of a probability distribution.

Long (1987) studied the accuracy of the last 40 years of Census Bureau projections. Over this period a naive method (use the current value to forecast the future) has performed as well as the Census Bureau's projections. This time period may have been unique in its large fertility variation since it included both the "baby boom" and the "baby bust."

In forecasting, the future rates would be selected automatically and objectively. For the method to be successful the process that generated the past must continue into the future.

We will assume that forecasts are made for mortality and fertility rates rather than deaths and births. The denominators of these rates are based on population, which is determined at a fixed time during the year, while the other statistics in Table 1 involve the whole year. Thus, the population of several categories and/or times must be combined to provide a valid rate estimate.

In this paper the problem of interest is to forecast the future probability distribution of population and to determine the variances of these forecasts. Forecasting of fertility, mortality, and migration rates is viewed as an intermediate step. Future population depends only on the present population and

future survival, fertility, and migration rates. Thus, future population predictions can be based on forecasts of these rates. Also, the prediction variance in these rates can be used to determine the prediction variance for the future population.

A new approach to forecasting fertility rates, which incorporates the joint behavior of all ages through time, was introduced by Bozik and Bell (1987). Their approach is to reduce the fertility rates to a lower dimensional vector and to forecast this vector using multivariate ARIMA model techniques. These forecasts are then converted to forecasts for the fertility rates and the estimated variances are calculated. In this approach only data on the rate in question (e.g.,fertility) is used in modelling and forecasting. This simplifies the analysis considerably and is optimal under reasonable assumptions, which are given in Section 4.

An important problem is to determine interval estimates for future values (called prediction intervals here). Of course, if the distribution of future values can be approximated by the Gaussian distribution, the prediction intervals can be obtained from the mean and the variance. It is not clear that the Gaussian approximation will be very good in this case.

Empirical work indicates that the fertility series is non-stationary. Bozik and Bell determine that the first difference of the logarithm of the fertility series is stationary. If this series is Gaussian, the predictive

distribution of future values is the lognormal.distribution.

In order to determine the predictive distribution of population, it is necessary to model the mortality, fertility, amd migration series. Since the recursive equations for population involve products of the components, it may be impossible to determine the predictive distribution of population analytically. The distribution could be approximated by simulating the component series and using the accounting equations to calculate future population values. Quantiles of the simulated distribution can be used as prediction intervals. The current population distribution can be used as a starting value for the recursion.

In Section 2 we give a model for the fertility and mortality rates. In Section 3 we state the demographic accounting equations for one data collection system. In Section 4 we state our assumptions about the forecasts of fertility, mortality, and migration, and in Section 5 we provide the recursive formulae for propogation of the mean and the variance of future values of the population. These values can be used to obtain an approximate predictive distribution and also can be used to check the moments obtained by simulation. In Section 6 an alternative method for updating the mean and the variance is given using the extended Kalman filter (e.g., Ljung and Soderstrom, 1983, Section 2.2.3). Also, the assumptions and the algorithm of the extended Kalman filter are compared with those of Section 5.

2. Estimation of Rates

For three of the components (fertility, mortality, and emigration) we will assume that models and forecasts have been developed for the rates (rather than the events). Following the traditional demographic notation we will estimate each rate by

$$\text{rate} = \text{events/exposure} \qquad (1)$$

The events are assumed to be broken down by age, sex, and time interval as in Table 1. Exposure is the number of person-years lived in that category during that time period. It is not a measured quantity, so it must be estimated.

Willekens (1985) defines four data collection plans and points out that the plan impacts the estimation of rates. We will assume a period-cohort collection plan. Analogous results can be obtained using other plans.

We will make the simplifying assumption that births, deaths, immigration, and emigration happen uniformly throughout the year. For the period-cohort collection system this allows the exposure ${}_sL_{ij}$ to be estimated by

$${}_sL_{ij} = ({}_sy_{ij} + {}_sy_{i-1,j-1})/2 \qquad (2)$$

Estimates of fertility, mortality, and emigration rates can be obtained from (1), (2), and Table 1. The influence on national population of emigration and death are the same so they may be combined. Thus, we define fertility rates ${}_sr_{ij}$ and mortality rates ${}_sm_{ij}$ (which includes emigration) by

$$ {}_{s}r_{ij} = {}_{s}b_{ij}/{}_{f}L_{ij} \qquad (3) $$

and

$$ {}_{s}m_{ij} = ({}_{s}d_{ij} + {}_{s}u_{ij})/{}_{s}L_{ij} \qquad (4). $$

The rates defined in (3) and (4) are essentially estimates of probabilities (if multiple births in a year are ignored). For example, ${}_{s}r_{ij}$ is an estimate of ${}_{s}p_{ij}$ which is the probability a woman in class (i,j) gives birth to a child with sex=s. For a woman, who is in class (i,j) for a fraction λ of the year, we assume the probability that a child of sex=s is born in that portion of the year is ${}_{s}p_{ij}\lambda$. There are $2{}_{f}L_{ij}$ women in class (i,j) for some portion of the year.

Using this notation the numerator of ${}_{s}r_{ij}$ can be represented

$$ {}_{s}b_{ij} = \sum_{k=1}^{2{}_{f}L_{ij}} \text{Bernoulli}(\lambda_k \; {}_{s}p_{ij}) \qquad (5) $$

where λ_k is the random fraction of the year that the kth woman is in class (i,j). From previous assumptions the random variables $\{\lambda_k\}$ are a sample from a distribution uniformly distributed between 0 and 1. Thus, the mean and variance of λ_k are 0.5 and 1/12 respectively. If we also assume independence, then from (5)

$$ E({}_{s}b_{ij}) = {}_{f}L_{ij} \; {}_{s}p_{ij} \qquad (6) $$

and

$$ \text{Var}({}_{s}b_{ij}) = {}_{f}L_{ij} \; {}_{s}p_{ij} \cdot (1 - {}_{s}p_{ij}/2) \qquad (7) $$

It follows from (6) and (7) that ${}_{s}r_{ij}$ is an unbiased estimate of ${}_{s}p_{ij}$ with

$$\mathrm{Var}({}_{s}r_{ij}) = {}_{s}p_{ij}\ (1-{}_{s}p_{ij}/2)/{}_{f}L_{ij} \qquad (8)$$

A similar definition can be made for the survival probabilities and the same method can be used to demonstrate that the ratio estimates are unbiased.

Since (8) is small, we can use ${}_{s}r_{ij}$ as a surrogate for ${}_{s}p_{ij}$. The simplest model for ${}_{s}p_{ij}$ is that it is constant over time (for each j). Due to changes in the role of women in American society this will not be the case, but the changes should be gradual. This gradual change may be captured in some ARIMA model.

The standard (e.g., Box and Jenkins, (1970)) ARIMA model formulation assumes a constant variance for the series. Because of (8) we know that this is false, but small deviations from constant variance may not cause problems.

The number of live births (${}_{I}b_{ij} = {}_{f}b_{ij} + {}_{m}b_{ij}$) may not be broken down by sex - rather the total may be reported. Previous demographic studies have shown that the ratio of male to female births is approximately constant across age groups, time, and cultures; and the value of this ratio is g = 1.05. Thus, if only the total births are given, the male and female rates can be estimated by ${}_{m}r_{ij} = {}_{T}r_{ij}g/(g+1)$ and ${}_{f}r_{ij} = {}_{T}r_{ij}/(g+1)$, where ${}_{T}r_{ij}$ is the overall fertility rate ignoring the sex.

3. The Accounting Equations

By specializing Willeken and Drewes' results (1984) to the case of no internal migration, the demographic accounting equations can be stated

$$\underline{y}_i = X_i \, \underline{y}_{i-1} + F_i \, \underline{w}_i \qquad (9)$$

where $\underline{y}_i^T = ({}_f\underline{y}_i^T, {}_m\underline{y}_i^T)$, $\underline{w}_i^T = ({}_f\underline{w}_i^T, {}_m\underline{w}_i^T)$, with ${}_s\underline{y}_i^T = ({}_sy_{i0}, \cdots, {}_sy_{id})$ and ${}_s\underline{w}_i^T = ({}_sw_{i0}, \cdots, {}_sw_{id})$. Equation (9) is formally identical to the state equation of the Kalman filter (1960), but as in many social science problems the matrices are unknown. Thus, the Kalman updating equations cannot be used without some modification.

Now we state the equations which are used to determine X_i and F_i. The equations are most easily stated in three categories: newborns (j=0), open-ended category (j=d), and middle categories ($1 \leq j \leq d-1$).

Categories ($1 \leq j \leq d-1$)

All people in category j were either in category j-1 in the previous period or immigrated during the period. The basic recursive equation can be stated

$${}_sy_{ij} = {}_ss_{ij} \, {}_sy_{i-1,j-1} + {}_sf_{ij} \, {}_sw_{ij} \qquad (10)$$

where ${}_ss_{ij}$ and ${}_sf_{ij}$ are both functions of the mortality rate. In fact, ${}_ss_{ij} = (1-{}_sm_{ij}/2)/(1+{}_sm_{ij}/2)$ and ${}_sf_{ij} = (1+{}_ss_{ij})/2$. By a Taylor expansion we have that

$${}_ss_{ij} = 1 - {}_sm_{ij} + O({}_sm_{ij}^2) \qquad (11)$$

Ignoring the remainder term in (11) we see that ${}_{s}s_{ij}$ can be interpreted as a survival probability. Since the immigrants are only present for half a year on the average, their survival function is larger than the natives.

Category 0

All survivors in this category were born either to natives or to immigrants. In either case three things must occur simultaneously

(i) the woman must survive until the birth

(ii) the woman must give birth to a child of a specified sex

(iii) the baby must survive the rest of the year

For native women the probabilities of the three events are ${}_{f}f_{ij}$, ${}_{s}r_{ij}$, and ${}_{s}f_{i0}$ respectively. We define ${}_{s}q_{ij}$ as the product of these three probabilities

$$ {}_{s}q_{ij} = {}_{f}f_{ij}\ {}_{s}r_{ij}\ {}_{s}f_{i0} \qquad (12) $$

which measures the frequency of surviving infants arising from native women in category j. A similar expression is valid for frequency of infants arising from immigrants with a multiplier of 0.5 to reflect their random entry into the country. The number of survivors in category 0 can be written as a sum over the categories of women

$$ {}_{s}y_{i0} = \sum_{j} {}_{s}q_{ij}\ ({}_{f}y_{i-1,j} + {}_{f}w_{ij}/2) \qquad (13) $$

Category d

The open ended category is the only one for which people can remain in the same category in successive years. Using similar

arguments to the middle categories the basic equation can be shown to be

$$ {}_s y_{id} = \sum_{j=d-1}^{d} ({}_s s_{ij}\ {}_s y_{i-1,j} + {}_s f_{ij}\ {}_s w_{ij}) \qquad (14) $$

Equations (10)-(14) determine the matrices X_i and F_i. In fact,

$$ X_i = \begin{bmatrix} {}_f\underline{q}_i^T & \underline{0}^T \\ {}_f S_i & 0 \\ {}_f\underline{q}_i^T & \underline{0}^T \\ 0 & {}_m S_i \end{bmatrix} \qquad (15) $$

$$ F_i = \begin{bmatrix} {}_f\underline{q}_i^T/2 & \underline{0}^T \\ {}_f F_i & 0 \\ \underline{0}^T & {}_m\underline{q}_i^T/2 \\ 0 & {}_m F_i \end{bmatrix} \qquad (16) $$

where $\underline{q}_i^T = ({}_f\underline{q}_i^T, {}_m\underline{q}_i^T)$ and ${}_s\underline{q}_i^T = ({}_s q_{i0}, {}_s q_{i1}, \cdots, {}_s q_{id})$, and the d x d+1 dimensional matrices ${}_s S_i$ and ${}_s F_i$ are given by

$$ {}_s S_i = \begin{bmatrix} {}_s s_{i1} & 0 & 0 & 0 \\ 0 & {}_s s_{12} & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & {}_s s_{i,d-1} & {}_s s_{id} \end{bmatrix} \qquad (17) $$

$$ {}_s F_i = \begin{vmatrix} {}_s f_{i1} & 0 & 0 & 0 \\ 0 & {}_s f_{12} & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & {}_s f_{i,d-1} & {}_s f_{id} \end{vmatrix} \qquad (18) $$

In the sequel it will be useful to have the representation

$$K_i^T = K^T(\underline{r}_i) = \begin{bmatrix} X_i^T \\ F_i^T \end{bmatrix} = (\underline{k}_{1i} \cdots \underline{k}_{Di}) \qquad (19)$$

where $\underline{r}_i=(\underline{r}_i,\underline{s}_i)$ with $\underline{r}_i = ({}_f\underline{r}_i, {}_m\underline{r}_i)$ and ${}_s\underline{r}_i=({}_s r_{i0}, {}_s r_{i1}, \cdots, {}_s r_{id})$, and a similar definition for $\underline{s}_i$, where $D=2(d+1)$. Also we define

$$\underline{z}_i^T = (\underline{y}_{i-1}^T, \underline{w}_i^T) \qquad (20)$$

so that (9) can be stated

$$\underline{y}_i = K_i \underline{z}_i \qquad (21).$$

Equation (9) shows that $\underline{y}_{n+k}$ can be calculated if $\underline{y}_n$ and $\underline{\theta}_{n+1}, \ldots, \underline{\theta}_{n+k}$ are known where $\underline{\theta}_i = (\underline{r}_i, \underline{s}_i, \underline{w}_i)$. The use of $\underline{s}_i$ rather than $\underline{m}_i$ or $\underline{f}_i$ is somewhat arbitrary since the three are related by 1-1 transformations. Since X_i and F_i are nonlinear functions of $\underline{m}_i$, it will prove convenient to use one of the other two vectors. The survival model could be obtained by combining models for emigration and mortality.

After models are obtained for the three components of $\{\underline{\theta}_n\}$, simulation can be used to approximate the distribution of $\underline{y}_{n+k}$. One can generate M independent replicates of $\{\underline{\theta}_{n+1}, \ldots, \underline{\theta}_{n+k}\}$ and calculate $\underline{y}_{n+k}$ using (9) for each replicate. The empirical distribution of these M replicates should be close to the (unknown) distribution of $\underline{y}_{n+k}$.

4. Component Forecasts

In this section we will assume that forecast means and variances are available for the next L values of the components of $\underline{\theta}_i$. The notation is given in Table 2 where $\underline{\theta}^{(n)}=(\underline{\theta}_1,\underline{\theta}_2,..,\underline{\theta}_n)$ and similar definitions are made for $\underline{r}^{(n)}$, $\underline{s}^{(n)}$, $\underline{w}^{(n)}$, and $\underline{y}^{(n)}$.

Table 2: Notation for Forecasts

quantity $(1\leq i\leq L)$	Mean	Variance	Data Used
fertility: $\underline{r}_{n+i}$	$\bar{\underline{r}}_{n+i}$	$\Sigma_{\underline{r}_{n+i}}$	$\underline{r}^{(n)}$
survival: $\underline{s}_{n+i}$	$\bar{\underline{s}}_{n+i}$	$\Sigma_{\underline{s}_{n+i}}$	$\underline{s}^{(n)}$
immigration: $\underline{w}_{n+i}$	$\bar{\underline{w}}_{n+i}$	$\Sigma_{\underline{w}_{n+i}}$	$\underline{w}^{(n)}$
population: $\underline{y}_{n+i}$	$\bar{\underline{y}}_{n+i}$	$\Sigma_{\underline{y}_{n+i}}$	$\underline{y}^{(n)}$

We forecast future values using data $\underline{D}^{(n)} = (\underline{y}^{(n)},\underline{\theta}^{(n)})$ available at time n. If we assume a square error loss function, the optimal forecast of a parameter is its conditional expectation given $\underline{D}^{(n)}$. Thus, for example, the optimal forecast of $\underline{r}_{n+i}$ is $E(\underline{r}_{n+i}|\underline{D}^{(n)})$. Table 2 shows that the forecast of $\underline{r}_{n+i}$ has been obtained from $\bar{\underline{r}}_{n+i} = E(\underline{r}_{n+i}|\underline{r}^{(n)})$. This assumption eliminates complicated modelling of joint multivariate time series and is valid under the following two assumptions

(A1) $\underline{\theta}_{n+i}$ and $\underline{y}^{(n+i-1)}$ are conditionally independent given $\underline{\theta}^{(n)}$ for all $i\geq 1$ for all n.

(A2) The three component series of $\underline{\theta}^{(n)}$ are independent.

Under these assumptions we have, for example,

$$E(\underline{r}_{n+i}|\underline{D}^{(n)}) \overset{(A1)}{=} E(\underline{r}_{n+i}|\underline{\theta}^{(n)}) \overset{(A2)}{=} E(\underline{r}_{n+i}|\underline{r}^{(n)}) = \bar{\underline{r}}_{n+i} \quad (22)$$

where the assumptions used to establish the equality have been shown. If $\bar{\underline{\theta}}_{n+i} = E(\underline{\theta}_{n+i}|\underline{D}^{(n)})$ then a similar argument establishes

$$\bar{\underline{\theta}}_{n+i} = (\bar{\underline{r}}_{n+i}, \bar{\underline{s}}_{n+i}, \bar{\underline{w}}_{n+i}) \quad (23)$$

Also, we define $\bar{K}_{n+i} = E(K_{n+i}|\underline{D}^{(n)})$ then by (19) and (23) this can be shown to be $K(\bar{\underline{r}}_{n+i})$. This equality would not be valid if $\underline{m}_i$ replaced $\underline{s}_i$ as a component in $\underline{\theta}_i$.

Now we give a brief discussion of the assumptions. We have assumed that the three components of $\underline{\theta}_n$ were independent in (A2). While fertility and mortality should be roughly independent, they have a small correlation with immigration. Since immigration influences the composition of the population, it could influence future fertility and mortality rates. Whether the immigration data is sufficiently accurate to model and to exploit this relationship is not clear.

Obviously, $\{\underline{y}_n\}$ and $\{\underline{\theta}_n\}$ are dependent. In fact, $\underline{y}_{n+k}$ can be calculated exactly from $\underline{y}_n$ and $\{\underline{\theta}_{n+1}, \ldots \underline{\theta}_{n+k}\}$. We expect that future values of $\{\underline{\theta}_n\}$ should be dependent on previous values. Assumption (A1) says that knowledge of future values of the population will not make future rates any more predictability than they would be using only the previous rates.

All the series of Table 2 are necessarily positive. An approach which utilizes this fact should produce better forecasts

as it will not forecast negative values even in the distant future. Although the Box-Cox (1964) family of transformations could be employed, we will use the log transformation for each of the components.

The component series are then modelled and forecast using standard (e.g., ARIMA model) techniques. The resulting mean and covariance matrix are appropriate for the transformed data. If the transformed data are approximately Gaussian, the following result can be used to obtain the moments of the untransformed series. If $\underline{y}=\log(\underline{x}) \sim N(\underline{\mu},\Sigma)$ where $\underline{\mu}^T = (\mu_1,,,\mu_d)$ and $\Sigma = (\sigma_{ij})$, then

$$E(x_i) = \exp(\mu_i + \sigma_{ii}/2) \tag{24}$$

$$Var(x_i) = E^2(x_i)(\exp(\sigma_{ii})-1) \tag{25}$$

$$Cov(x_i,x_j) = \exp(\sigma_{ij}) \tag{26}$$

Equations (24)-(26) can be applied to find the moments of $\underline{r}_{n+i}$, for example, with $\underline{\mu} = E(\underline{r}^*_{n+i}|\underline{r}^{(n)})$ and $\Sigma=Cov(\underline{r}^*_{n+i}|\underline{r}^{(n)})$ where $\underline{r}^*_n = \log(\underline{r}_n)$.

5. Mean and Variance Propagation

In this section we obtain recursive forecasts for $\bar{y}_{n+i}$ and $\Sigma_{y_{n+i}}$ using the forecasts of the three components in Table 2. The following lemma is useful in the derivation of the recursion.

Lemma 1: K_{n+i} and z_{n+i} are independent given $D^{(n)}$.

We use p as a generic symbol for a density or distribution, where the argument of p defines the random variable. From (A1) with i=1 we can show that

$$p(\theta_{n+i}|D^{(n)},y_{n+1},\cdots,y_{n+i-1}) = p(\theta_{n+i}|\theta^{(n)}) \qquad (27)$$

It follows from (27) that

$$p(\theta_{n+i}|D^{(n)},y_{n+i-1}) = p(\theta_{n+i}|\theta^{(n)}) \qquad (28)$$

and

$$p(w_{n+i}|D^{(n)},y_{n+i-1}) = p(w_{n+i}|\theta^{(n)}) \qquad (29)$$

Thus, by (28), (29), and (A2) we have

$$p(\tau_{n+i}|z_{n+i},D^{(n)}) = \frac{p(\theta_{n+i}|y_{n+i-1},D^{(n)})}{p(w_{n+i}|D^{(n)},y_{n+i-1})} = \frac{p(\theta_{n+i}|\theta^{(n)})}{p(w_{n+i}|\theta^{(n)})}$$
$$= p(\tau_{n+i}|\theta^{(n)}) \qquad (30)$$

Equation (30) shows that τ_{n+i} and z_{n+i} are independent given $D^{(n)}$. Since K_{n+i} is a function of τ_{n+i}, it is also independent of z_{n+i} given $D^{(n)}$, which completes the proof of the lemma.

The recursion for the mean update is given by the following corollary.

Corollary: $\bar{\underline{y}}_{n+i} = \bar{X}_{n+i}\bar{\underline{y}}_{n+i-1} + \bar{F}_{n+i}\bar{\underline{w}}_{n+i} = \bar{K}_{n+i}\bar{\underline{z}}_{n+i}$ for $i \geq 1$ with $\bar{\underline{y}}_n = \underline{y}_n$

Proof: Applying Lemma 1 we have that

$\bar{\underline{y}}_{n+i} = E(K_{n+i}\underline{z}_{n+i}|\underline{D}^{(n)}) = E(K_{n+i}|\underline{D}^{(n)})E(\underline{z}_{n+i}|\underline{D}^{(n)}) = \bar{K}_{n+i}\bar{\underline{z}}_{n+i}$

where $\bar{K}_{n+i} = K(\bar{\tau}_{n+i})$.

The next lemma will be used to calculate the recursive form of the covariance update.

Lemma 2: Let $K^T = (\underline{k}_1 \ldots \underline{k}_D)$ be independent of $\underline{z}$ with $\bar{\underline{z}} = E(\underline{z})$, $Cov(\underline{z}) = \Sigma_{\underline{z}}$, $\bar{K} = E(K)$, and $S_{uv} = Cov(\underline{k}_u, \underline{k}_v)$ for $1 \leq u,v \leq D$ then

$$Cov(K\underline{z}) = \bar{K}\Sigma_{\underline{z}}\bar{K}^T + H \qquad (31)$$

where $H = E(K-\bar{K})\Gamma(K-\bar{K})^T = (h_{uv})$ with $h_{uv} = Tr(\Gamma S_{uv})$, and $\Gamma = \bar{\underline{z}}\bar{\underline{z}}^T + \Sigma_{\underline{z}}$.

Proof: By the independence assumption $E(K\underline{z}) = \bar{K}\bar{\underline{z}}$. Also, the two summands in the decomposition $K\underline{z} - \bar{K}\bar{\underline{z}} = K(\underline{z}-\bar{\underline{z}}) + (K-\bar{K})\bar{\underline{z}}$ are uncorrelated and have covariance matrices $\bar{K}\Sigma_{\underline{z}}\bar{K}^T + H_1$ and H_2 respectively with $H_1 = (Tr(\Sigma_{\underline{z}}S_{uv}))$ and $H_2 = (\bar{\underline{z}}^T S_{uv}\bar{\underline{z}})$. Since $H = H_1 + H_2$, this completes the proof.

Now we give a recursive method to calculate the covariance matrix of the forecast. By Lemma 1 K_{n+i} is independent of $\underline{z}_{n+i}$ given $\underline{D}^{(n)}$. Thus, with $\Sigma_{\underline{z}_{n+i}} = diag(\Sigma_{\underline{y}_{n+i-1}}, \Sigma_{\underline{w}_{n+i}})$ from Lemma 2 we have that

$$\Sigma_{\underline{y}_{n+i}} = \bar{K}_{n+i}\Sigma_{\underline{z}_{n+i}}\bar{K}_{n+i}^T + H_{n+i} \qquad (32)$$

where

$H_{n+i}=E(K_{n+i}-\bar{K}_{n+i})\Gamma_{n+i}(K_{n+i}-\bar{K}_{n+i})^T=(h_{uv,n+i})=(Tr(\Gamma_{n+i}S_{uv,n+i}))$

with $\Gamma_{n+i} = \Sigma_{\underline{z}_{n+i}}+\bar{\underline{z}}_{n+i}\bar{\underline{z}}_{n+i}^T$ and $S_{uv,n+i} = Cov(\underline{k}_{u,n+i},\underline{k}_{v,n+i}|\underline{D}^{(n)})$.

From (15), (16), and (19) the DxD dimensional matrix K_{n+i} is given by

$$K_{n+i} = \begin{bmatrix} {}_f\underline{q}_i^T & \underline{0}^T & {}_f\underline{q}_i^T/2 & \underline{0}^T \\ {}_fS_i & 0 & {}_fF_i & 0 \\ {}_f\underline{q}_i^T & \underline{0}^T & \underline{0}^T & {}_m\underline{q}_i^T/2 \\ 0 & {}_mS_i & 0 & {}_mF_i \end{bmatrix} \qquad (33)$$

where D=2(d+1) and $\bar{K}_{n+i}$ is calculated by replacing τ_{n+i} with $\bar{\tau}_{n+i}$ in (33). Now we sketch the calculation of $S_{uv,n+i}$ for $1\leq u,v\leq D$.

For all rows of K_{n+i} except the first and the (d+2)nd the maximum number of non-zero elements is 2 ,and the elements are linear functions of ${}_sf_{n+i,j}$. Thus, covariances of the elements of any two pairs of these D-2 rows can be obtained easily from $\Sigma_{\underline{f}_{n+i}}$. Because the two complicated rows (1 and d+2) involve products of elements of $\underline{f}_{n+i}$, it is convenient to use it rather than $\underline{s}_{n+i}$. Their covariance matrices are related by $\Sigma_{\underline{f}_{n+i}}=\Sigma_{\underline{s}_{n+i}}/4$.

The elements of rows 1 and d+2 are products of elements of $\underline{\tau}_{n+i}$. Since ${}_sq_{n+i,j}$ involves a product ${}_ff_{n+i,j}\ {}_ff_{n+i,0}$, $Var({}_fq_{n+i,j})$ will in general involve mixed moments of order four. The mixed moments of order three or more are all 0 for the multivariate Gaussian distribution, which we assume in the sequel

for the distribution of the mortality rates. We will use the following well known results on quadratic forms (e.g., Searle, 1971, Chap. 2) to derive Lemma 3. If $\underline{x} \sim N(\underline{\mu},\Sigma)$,

(a) $E(\underline{x}^T P \underline{x}) = Tr(P\Sigma) + \underline{\mu}^T P \underline{\mu}$

(b) $Var(\underline{x}^T P \underline{x}) = 2\ Tr(P\Sigma)^2 + 4\underline{\mu}^T P \Sigma P \underline{\mu}$

(c) $Cov(\underline{x}^T P_1 \underline{x}, \underline{x}^T P_2 \underline{x}) = 2\ Tr(P_1 \Sigma P_2 \Sigma) + 4\underline{\mu}^T P_1 \Sigma P_2 \underline{\mu}$

Using an obvious notation the following are obtained from (a)-(c)

$$E(x_0 x_j) = \mu_0 \mu_j + \sigma_{0j} \tag{34}$$

$$Var(x_0 x_j) = \sigma_{00}^2 + 2\sigma_{0j}^2 + \sigma_{jj}^2 + \mu_0^2 \sigma_{jj} 2\mu_0 \mu_j \sigma_{0j} + \mu_j^2 \sigma_{00} \tag{35}$$

$$Cov(x_0 x_j, x_0 x_k) = 2(\sigma_{0j}\sigma_{0k} + \sigma_{00}\sigma_{jk}) + \mu_0^2 \sigma_{jj} + \mu_0(\mu_k \sigma_{0k} + \mu_j \sigma_{0k}) + \mu_j \mu_k \sigma_{00} \tag{36}$$

Equations (34)-(36) are used in the derivation of Lemma 3. The lemma is stated for females, but a similar one can be obtained for males.

Lemma 3. Under the assumption of normality of $\underline{f}_{n+i}$ we have

$$Var({}_f q_{n+i,j}) = \sigma_{r_{jj}}(\sigma(f_{0j}) + \bar{f}_0 \bar{f}_j)^2 + (\bar{r}_j^2 + \sigma(r_{jj}))Var({}_f f_{n+i,0}\ {}_f f_{n+i,j})$$

with $$Var({}_f f_{n+i,0}\ {}_f f_{n+i,j}) = \sigma_{f_{00}}^2 + 2\sigma_{f_{0j}}^2 + \sigma_{f_{jj}}^2 + \bar{f}_0^2 \sigma_{f_{jj}} + 2\bar{f}_0 \bar{f}_j \sigma_{f_{0j}} + \bar{f}_j^2 \sigma_{f_{00}}$$

$$Cov({}_f q_{n+i,k}, {}_f q_{n+i,j}) = \sigma_{r_{jk}}(\bar{f}_j \bar{f}_0 + \sigma_{f_{0j}})(\bar{f}_k \bar{f}_0 + \sigma_{f_{0k}}) + (\sigma_{r_{jk}} + \bar{r}_j \bar{r}_k)$$
$$[2(\sigma_{f_{j0}}\sigma_{f_{k0}} + \sigma_{f_{00}}\sigma_{f_{jk}}) + \bar{f}_0^2 \sigma_{f_{jk}} + \bar{f}_0(\bar{f}_k \sigma_{f_{0j}} + \bar{f}_j \sigma_{f_{0k}}) + \bar{f}_j \bar{f}_k \sigma_{f_{00}}]$$

Proof: Since we state the results for females, it is possible to suppress the sex variable. We do this in places to simplify the notation. Using the standard conditioning argument

$\mathrm{Var}({}_f q_{n+i,j}) = E(\mathrm{Var}({}_f q_{n+i,j} | {}_f r_{n+i,j})) + \mathrm{Var}(E({}_f q_{n+i,j} | {}_f r_{n+i,j}))$

It follows from (34) that

$$({}_f q_{n+i,j} | {}_f r_{n+i,j}) = {}_f r_{n+i,j} E({}_f f_{n+i,0}\ {}_f f_{n+i,j})$$
$$= {}_f r_{n+i,j} ({}_f \bar{f}_0\ {}_f \bar{f}_j + \sigma_{f_{0j}})$$

so $\mathrm{Var}({}_f q_{n+i,j} | {}_f r_{n+i,j}) = \sigma_{r_{jj}} ({}_f \bar{f}_0\ {}_f \bar{f}_j + \sigma_{f_{0j}})^2$

Also, $\mathrm{Var}({}_f q_{n+i,j} | {}_f r_{n+i,j}) = {}_f r^2_{n+i,j} \mathrm{Var}({}_f f_{n+i,j}\ {}_f f_{n+i,0})$ so the variance assertion is obtained by applying (35). Similarly for the covariance we have

$$\mathrm{Cov}({}_f q_{n+i,j}\ {}_f q_{n+i,k}) = \mathrm{Cov}(E({}_f r_{n+i,j} | {}_f r_{n+i,j}), E({}_f q_{n+i,k} | {}_f r_{n+i,k}))$$
$$+ E(\mathrm{Cov}({}_f q_{n+i,j}, {}_f q_{n+i,k}) | {}_f r_{n+i,j}, {}_f r_{n+i,k}) =$$
$$\mathrm{Cov}({}_f r_{n+i,j} ({}_f \bar{f}_j\ {}_f \bar{f}_0 + \sigma_{f_{0j}}), {}_f r_{n+i,k} ({}_f \bar{f}_k\ {}_f \bar{f}_0 + \sigma_{f_{0k}})) +$$
$$E({}_f r_{n+i,j}\ {}_f r_{n+i,k}\ \mathrm{Cov}({}_f f_{n+i,j}\ {}_f f_{n+i,0}, {}_f f_{n+i},\ {}_f f_{n+i,0}))$$

Since $\mathrm{Cov}({}_f r_{n+i,j}, {}_f r_{n+i,k}) = \sigma_{r_{jk}}$, $E({}_f r_{n+i,j}\ {}_f r_{n+i,k}) = \sigma_{r_{jk}} + {}_f \bar{r}_j\ {}_f \bar{r}_k$ the covariance assertion follows.

6. Extended Kalman Filter

In this section we give an approximate solution to the forecasting problem using the extended Kalman filter (EKF). In this approach the non-linear system (9) is linearized using a first order Taylor expansion. The Kalman filter algorithm can be applied to the resulting equations. The approach is simpler and the assumptions are weaker than that of Section 5, but the means and variances are only approximations.

As in Section 4 we assume that independent multivariate models have been developed for the three components of $\underline{\theta}_n$. Also, we assume that each component model can be represented in state space form. This will be the case if ARIMA models are developed for the components (e.g., Akaike, 1974).

Let $\underline{v}_{nk}$ denote the state variable representation for the k^{th} component of $\underline{\theta}_n$ for k=1,2,3 so that

$$\underline{\theta}_n = \begin{bmatrix} \underline{r}_n \\ \underline{s}_n \\ \underline{w}_n \end{bmatrix} = \left[\begin{array}{ccc|ccc|ccc} I & 0\ldots 0 & & 0 & 0\ldots 0 & & 0 & 0\ldots 0 \\ 0 & 0\ldots 0 & & I & 0\ldots 0 & & 0 & 0\ldots 0 \\ 0 & 0\ldots 0 & & 0 & 0\ldots 0 & & I & 0\ldots 0 \end{array}\right] \begin{bmatrix} \underline{v}_{n1} \\ \underline{v}_{n2} \\ \underline{v}_{n3} \end{bmatrix} = B\underline{v}_n \qquad (37)$$

The state equation for $\underline{v}_n$ is

$$\underline{v}_{n+1} = A\underline{v}_n + W\underline{e}_{n+1} \qquad (38)$$

where $A=\text{diag}(A_1,A_2,A_3)$, $W=\text{diag}(W_1,W_2,W_3)$, $\underline{e}_n^T=(\underline{e}_{n1}^T,\underline{e}_{n2}^T,\underline{e}_{n3}^T)$, and $\{\underline{e}_n\}$ are independent and identical $N(\underline{0},\Sigma)$ random errors with $\Sigma=\text{diag}(\Sigma_1,\Sigma_2,\Sigma_3)$. It is clear from (37) and (38) that the components of $\underline{\theta}_n$ are independent (e.g., (A2) of Section 4)).

6.1 Linearization of the Population Equations

In order to apply the Kalman algorithm it is necessary to linearize (9). We expand these in a first order Taylor series in $\underline{\phi}_n=(\underline{y}_{n-1},\underline{\theta}_n)$ about $\bar{\underline{\phi}}_n=(\bar{\underline{y}}_{n-1},\bar{\underline{\theta}}_n)$. These equations can be written

$$\underline{y}_n = \bar{K}_n\,\bar{\underline{z}}_n + \bar{G}_n(\underline{\phi}_n-\bar{\underline{\phi}}_n) \tag{39}$$

where

$$\bar{G}_n = (\bar{X}_n \quad \bar{D}_n) \tag{40}$$

and

$$D_n = D(\underline{\theta}_n) = (D_{nr}\ D_{ns}\ F_n) = \left(\frac{\partial}{\partial \underline{r}_n}(K_n\underline{z}_n) \quad \frac{\partial}{\partial \underline{s}_n}(K_n\underline{z}_n) \quad F_n\right) \tag{41}$$

where F_n is given by (16). Equation (39) can be written

$$\underline{y}_n = -\bar{D}_n^*\,\bar{\underline{\phi}}_n + \bar{G}_n\,\underline{\phi}_n \tag{42}$$

where

$$\bar{D}_n^* = (0\ \bar{D}_{nr}\ \bar{D}_{ns}\ 0) \tag{43}$$

The matrices D_{ir} and D_{is} are determined in an obvious fashion and are given by equations (44)-(52)

$$D_{ir} = \begin{bmatrix} {}_f\underline{\alpha}_i^T & \underline{0}^T \\ 0 & 0 \\ \underline{0}^T & {}_m\underline{\alpha}_i^T \\ 0 & 0 \end{bmatrix} \tag{44}$$

where ${}_s\underline{\alpha}_i^T=({}_s\alpha_{i0},{}_s\alpha_{i1},\cdots,{}_s\alpha_{id})$ with

$${}_s\alpha_{ij}={}_sf_{i0}\ {}_ff_{ij}\ {}_f\chi_{ij} \tag{45}$$

and

$${}_s\chi_{ij} = {}_sy_{i-1,j-1} + {}_sw_{ij}/2 \tag{46}$$

$$D_{is} = \begin{bmatrix} {}_f\underline{\gamma}_i^T & \underline{0}^T \\ {}_f\chi_i & 0 \\ {}_m\underline{\gamma}_i^T & \underline{\xi}^T \\ 0 & {}_m\chi_i \end{bmatrix} \tag{47}$$

with

$${}_s\chi_i = \begin{bmatrix} {}_s\chi_{i1} & 0 & 0 & 0 \\ 0 & {}_s\chi_{12} & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & {}_s\chi_{i,d-1} & {}_s\chi_{id} \end{bmatrix} \tag{48}$$

$${}_f\underline{\chi}_i^T = (\sum_j {}_f\gamma_{ij}, {}_f\gamma_{i1}, \cdots, {}_f\gamma_{id}) \tag{49}$$

and

$${}_m\underline{\chi}_i^T = (0, {}_f\gamma_{i1}, \cdots, {}_f\gamma_{id}) \tag{50}$$

where

$${}_s\gamma_{ij} = {}_s f_{i0}\ {}_s r_{ij}\ {}_f\chi_{ij} \tag{51}$$

and

$$\underline{\xi}^T = (0.5\ {}_f f_{i0} \sum_j {}_m r_{ij}\ {}_f\chi_{ij}, 0, \ldots 0) \tag{52}$$

6.2 Recursions for Forecasting Using the EKF

Combining (38) and (42) the equations for the extended Kalman filter can be written

$$\underline{\phi}_{n+1} = \begin{bmatrix} 0 & \bar{D}_{nr} & \bar{D}_{ns} & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \bar{\underline{y}}_{n-1} \\ \bar{\underline{v}}_n \end{bmatrix} + \begin{bmatrix} X_n & \bar{D}_n \\ 0 & A \end{bmatrix} \underline{\phi}_n + \begin{bmatrix} 0 \\ W \end{bmatrix} \underline{e}_{n+1} \tag{53}$$

which can be stated in matrix form as

$$\underline{\phi}_{n+1} = -\mathbb{D}_n^* \bar{\underline{\phi}}_n + \mathbb{R}_n \underline{\phi}_n + W^* \underline{e}_{n+1} \tag{54}$$

A recursion for the forecast mean and covariance follows easily from (54). In fact,

$$\bar{\underline{\phi}}_{n+i+1} = (\mathbb{R}_{n+i} - \mathbb{D}_{n+i}^*)\bar{\underline{\phi}}_{n+i} \tag{55}$$

and

$$\Sigma_{\underline{\phi}_{n+i+1}} = \mathbb{R}_{n+i} \Sigma_{\underline{\phi}_{n+i}} \mathbb{R}_{n+i}^T + W^* \Sigma_{\underline{e}_{n+i+1}} W^{*T} \tag{56}$$

Comparison of (55) and (56) with the results of Section 5 indicates that

(i) the updating formula for $\bar{\underline{y}}_{n+i}$ using the EKF coincides with the procedure of Section 5.

(ii) the updating formula for $\Sigma_{\underline{y}_{n+i}}$ using the EKF does not coincide with procedure of Section 5.

(iii) $\mathrm{Cov}(\underline{y}_{n+i-1}, \underline{\phi}_{n+i})$, which is computed from (56), will not in general be 0. This shows that assumption (A1) is not required using the EKF.

Now we briefly discuss the starting values for the recursion using data $\underline{D}^{(n)}$. Some of the components of $\underline{\phi}_1, \ldots, \underline{\phi}_n$ will be missing (e.g., $\underline{y}_{n+1}$). A useful approximation to the best starting values for the recursions, which are obtained using the EKF algorithm, may be obtained using

$$\bar{\underline{\phi}}_n = \begin{bmatrix} \underline{y}_n \\ A_n \underline{y}_n \end{bmatrix} \quad \text{and} \quad \Sigma_{\underline{\phi}_n} = \begin{bmatrix} 0 & 0 \\ 0 & W\Sigma_{\underline{e}_{n+1}} W^T \end{bmatrix} \tag{57}$$

Acknowledgment

I wish to thank Bill Bell, Jim Bozik, John Long, Nash Monsour, and Franz Willekens for useful comments on a previous version of this manuscript.

7. References

Akaike, H. (1974), "Markovian representation of stochastic processes and its application to the analysis of autoregressive moving average processes," Ann. Inst. Statist. Math., 26, 363-387.

Box, G.E.P. and Cox, D.R. (1964), "An analysis of transformations," J. Roy. Statist. Soc. Ser. B, 26, 211-243.

Box, G.E.P. and Jenkins, G.M. (1970), Time Series Analysis: Forecasting and Control, Holden-Day, San Francisco.

Bozik, J.E. and Bell, W.R. (1987), "Forecasting age specific fertility using principal components,", Proc. of Social Statist. Sec.of Amer. Statist. Assn., 396-401.

Cochran, W.G. (1963), Sampling Techniques, John Wiley, New York.

Cohen, J.E. (1986),"Population forecasts and confidence intervals for Sweden: a comparison of model-based and empirical approaches," Demography, 23, 105-126.

Kalman, R.E. (1960), "A new approach to linear filtering and prediction problems," J. Basic Engineering, 82, 35-45.

Keyfitz, N. (1977), Applied Mathematical Demography, John Wiley, New York.

Land, K.C. (1986), "Methods for national population forecasts: a review," J. Amer. Statist. Assn., 81, 888-901.

Lee, R. D. (1981), Comment on "Modelling demographic relationships:an analysis of forecast functions for Australian births," by J. McDonald, J. Amer. Statist. Assn.,.76, 793-795.

Lee, R.D. (1974),"Forecasting births in post-transition populations, J. Amer. Statist. Assn. , 69, 607-617.

Ljung, L. and Soderstrom, T. (1983), Theory and Practice of Recursive Estimation, MIT Press, Cambridge, Mass.

Long, J. F. (1981), Comment on "Modelling demographic relationships: an analysis of forecast functions for Australian births," by J. McDonald, J. Amer. Statist. Assn., 76, 796-798.

Long, J. F. (1987), "The accuracy of population projection methods at the U. S. Census Bureau," paper presented to the the Annual Meeting of the Population Association of America.

McDonald, J. (1981), "Modelling demographic relationships: an analysis of forecast functions for Australian births" J. Amer. Statist. Assn., 76,782-792.

Searle, S.R. (1971), Linear Models, John Wiley, New York.

Spencer, G. and Long, J. F. (1983), "The new Census Bureau

projections," Amer. Demographics, April , 24-31.

Tiao, G.C. and Box, G.E.P. (1981),"Modelling multiple time series with applications,"J. Amer. Statist. Assn., 76,802-816.

Willekens, F. (1985), Discussion on "Improving national population projection methodology," in Bureau of the Census, First Annual Research Conference Proceedings, 288-297.

Willekens, F. and Drewe, P. (1984), "A multiregional model for regional demographic projection," in Demographic Research and Spatial Policy: The Dutch Experience, H. ter Heide and F. Willekens (editors), London: Academic Press, 309-334.