SFB 649 Discussion Paper 2008-052a

Modeling and Forecasting Age-Specific Mortality: A Bayesian Approach

Wolfgang Reichmuth*
Samad Sarferaz*

* Humboldt-Universität zu Berlin, Germany

This research was supported by the Deutsche Forschungsgemeinschaft through the SFB 649 "Economic Risk".

http://sfb649.wiwi.hu-berlin.de
ISSN 1860-5664

SFB 649, Humboldt-Universität zu Berlin
Spandauer Straße 1, D-10178 Berlin

SFB 649 ECONOMIC RISK BERLIN

Modeling and Forecasting Age-Specific Mortality: A Bayesian Approach[1]

Wolfgang H. Reichmuth
Humboldt-Universität zu Berlin

Samad Sarferaz
Humboldt-Universität zu Berlin

This Version: October 20, 2008

Abstract

We present a new way to model age-specific demographic variables, using the example of age-specific mortality in the United States, building on the Lee–Carter approach and extending it in several dimensions. We incorporate covariates and model their dynamics jointly with the latent variables underlying mortality of all age classes. In contrast to previous models, a similar development of adjacent age groups is assured, allowing for consistent forecasts. We develop an appropriate Markov chain Monte Carlo algorithm to estimate the parameters and the latent variables in an efficient one-step procedure. Via the Bayesian approach we are able to assess uncertainty intuitively by constructing error bands for the forecasts. We observe that in particular parameter uncertainty is important for long-run forecasts. This implies that existing forecasting methods, which ignore certain sources of uncertainty, may yield misleadingly sure predictions. To test the forecast ability of our model we perform in-sample and out-of-sample forecasts up to 2050, revealing that covariates can help improve the forecasts for particular age classes. A structural analysis of the relationship between age-specific mortality and covariates is conducted in a companion paper.

JEL classification codes: C11, C32, C53, I10, J11

Keywords: demography, age-specific mortality, Lee–Carter, stochastic, Bayesian state space models, forecasts

[1] A previous version of this paper has been circulated under the title "Bayesian Demographic Modeling and Forecasting: An Application to U.S. Mortality". We would like to thank Pooyan Amir Ahmadi, Alexander Rathke, Henning Weber, and especially Albrecht Ritschl and Ulrich Woitek for helpful comments. Financial support by the Deutsche Forschungsgemeinschaft through the SFB 649 Economic Risk is gratefully acknowledged by both authors. Samad Sarferaz also acknowledges financial support by the European Science Foundation via the Globalizing Europe Economic History Network and the Marie Curie Research Training Networks and thanks the Universität Zürich and the European University Institute in Florence for their hospitality. Contact: `reichmwo@staff.hu-berlin.de`, `sarferaz@wiwi.hu-berlin.de`

1 Introduction

Demographic issues are of general interest since they address the most fundamental attributes of human life. Their research takes place at the crossways of economics and sociology, medicine, and other academic disciplines, which in turn are often influenced themselves by demographic findings. This gives rise to a multidisciplinary scientific interest. Of course, such research is not only of interest to science, but also to many recipients in the domains of politics and business. Reliable forecasts of future mortality and a better understanding of the determinants of changing mortality are obviously of great importance in areas such as social security and public health. In the private sector such advancements of knowledge can have substantial monetary value, since they improve the calculation of life insurance rates and pension schemes for the insurance industry. Population forecasts that can be derived from demographic rates are another example of interest beyond pure science owing to their implications for investment decisions in the public and private sectors. All of these potential recipients benefit most from stochastic models, which yield distributional statements on the probabilities of outcomes instead of pure projections of some scenarios. For this purpose, stochastic models of age-specific mortality and other demographic variables are needed.

We present a new way to model age-specific demographic variables using the example of age-specific mortality. Existing parametric and nonparametric approaches to modeling and forecasting mortality suffer from different shortcomings in the embodiment of the age dimension. Our model avoids these drawbacks. Furthermore, it is very general and comprises both the well-known Lee–Carter model and the use of covariates as special cases. Advanced methods from the domain of Bayesian time series econometrics are used to set up the model and estimate the parameters. Unobserved or latent variables, which drive the common development of the observed age-specific variables, are complemented by observable covariates. We formulate two explicit laws of motion in the form of (vector) autoregressions (VARs), which ensure a relatively smooth development not only along the time but also the age dimension of the demographic variable. For the latter, this is usually neglected. The importance of this issue is demonstrated by the very smooth surface without jumps in Figure 1 representing U.S. mortality. We feel confident that a reasonable model of age-specific mortality should explicitly embody this feature and guarantee such smoothness across ages in forecasts too. By the use of VARs we also allow for mutual interactions between latent variables and all covariates in the model. Finally, we use Markov chain Monte Carlo (MCMC) methods to estimate the model with an efficient one-step procedure. By the choice of priors this Bayesian estimation approach also clearly reveals the assumptions made. Most notably, it yields not only point estimates but also distributional statements for the results in a very intuitive way.

Our approach is very flexible and can be applied to model all kinds of demographic variables, using different numbers of latent variables and different sets of covariates. In this paper, we present applications to U.S. mortality, with gross domestic product (GDP) and unemployment as important macroeconomic variables. Owing to our particular modeling approach, stochastic forecasts of the modeled variables are easily achieved and have the advantage of being fully consistent among adjacent age classes, unlike some parametric approaches or the

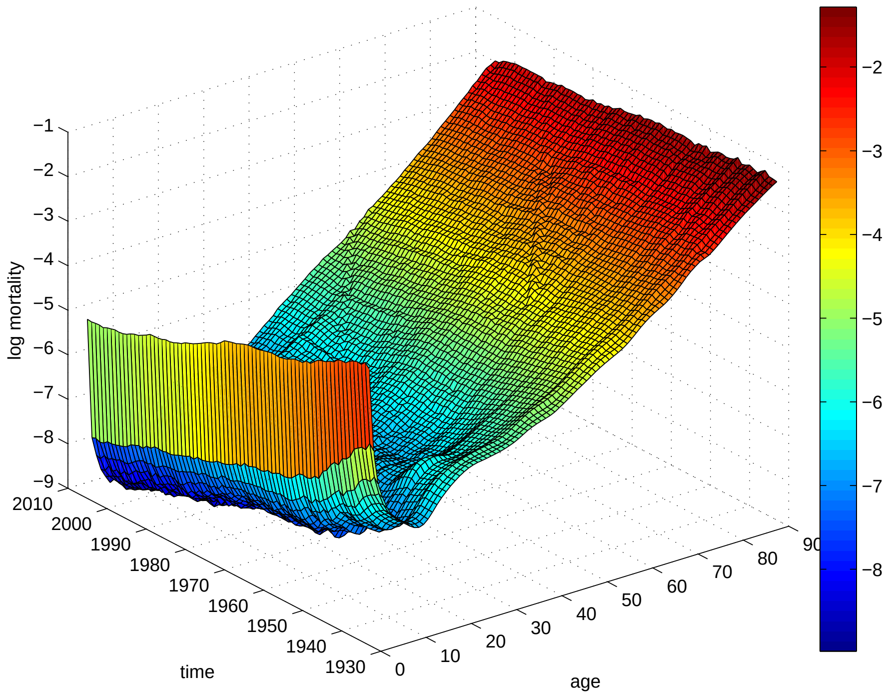

Figure 1: Mortality surface of the logarithmized age-specific total (female and male combined) mortality in the United States, 1933–2005.

popular Lee–Carter method. In addition to this important feature of age-related smoothness, we also can distinguish the impact of different sources of uncertainty on the forecast results. We show that the uncertainty associated with the random terms in the model is more important at the beginning, whereas the uncertainty associated with the estimation of parameters is very important in a longer perspective. This means that false confidence in forecasts may result from ignoring important sources of uncertainty by concentrating on the random term, such as in the Lee–Carter model. In-sample forecasts reveal that both versions of the model, either including covariates or not, perform accurately. We present out-of-sample forecasts of mortality with respective error bands for a longer horizon up to the year 2050 which show that covariates can help improve the forecasts for particular age classes. Moreover, the use of VARs, which is facilitated by the enormous reduction of the dimension with the help of latent variables, allows for further structural analyses of the interactions between the covariates and the demographic variable, revealing the full pattern of age-specific reactions to external influences. Such an extended analysis is presented in a companion paper.[2]

The presented approach can be applied to model, forecast, and analyze all kinds of age-specific variables. Mortality is just a prominent example owing to its great importance in general and to the fact that our model can be interpreted as a generalization of the established Lee–Carter model. Moreover, in addition to their own intrinsic value, forecasts of

[2]Cf. Reichmuth and Sarferaz (2008).

mortality also constitute an important part of the input needed for stochastic population forecasts with the cohort component method of stepwise interpolation of an initial population.

The rest of the paper is organized as follows: Section 2 provides a brief summary of the literature on modeling and forecasting mortality. Our model is stated in Section 3. Section 4 describes the predictive densities. Sections 5 and 6 address the priors and the estimation procedure, and the data are described in Section 7. The estimation and forecast results are presented in Section 8, which additionally provides some intuitively interpretable life table variables based on age-specific mortality. Finally, Section 9 presents our conclusions.

2 Literature on Modeling and Forecasting Mortality

We start with a short overview of some developments in modeling and forecasting age-specific mortality.[3] Models that map age to age-specific mortality take advantage of the obvious strong regularities in mortality's age pattern.[4] In the context of forecasting, these regularities have to be taken into account, because naive univariate forecasts of each age-specific time series separately would propagate too much noise, quickly leading to serious inconsistencies. Of course, such models also substantially reduce the dimensionality of the data to be handled.

2.1 Parametric Modeling of Age-Specific Mortality

Systematic patterns in mortality have been known since the development of the first life tables by Graunt (1662) and Halley (1693). In terms of a mathematical *law of mortality* for the observed age pattern, Gompertz (1825) first mentioned that mortality $m(x)$ at age x in adulthood shows a nearly exponential increase

$$m(x) = \alpha e^{\beta x} \, .$$

Among the many more sophisticated proposals for a formula of age-specific mortality since that time, Heligman and Pollard (1980) suggest a sum of three terms representing different components of mortality,

$$m(x) = A^{(x+B)^C} + De^{-E(\ln x - \ln F)^2} + GH^x / \left(1 + GH^x\right) \, ,$$

with eight time-dependent parameters $A_t, \ldots, H_t$. The rapidly falling first term accounts for mortality during childhood, the second term models the *accident hump* for young adults, and the third term picks up the Gompertz exponential for the senescent mortality of adulthood and old age. McNown and Rogers (1989) forecast the eight parameters of the Heligman–Pollard model using the univariate time series method of autoregressive integrated moving average (ARIMA) processes, which may lead to inconsistencies in the long run.

[3] Of course, we can only briefly sketch some major issues. Booth (2006) gives a comprehensive survey of demographic forecasting.

[4] For the sake of simplicity, except for the final life table calculations, we use the term *age-specific mortality* for both the probability ${}_1q_x = (l_x - l_{x+1}) \,/\, l_x$ of dying at age x, which is related to the population at risk, that is, the number l_x of survivors to age x, and the death rate ${}_1m_x = (l_x - l_{x+1}) \,/\, {}_1L_x$ at age x, which is related to the person-years ${}_1L_x$ lived at age x ($l_{x+1} \leq {}_1L_x \leq l_x$).

2.2 Lee–Carter and Non-parametric Modeling of Age-Specific Mortality

Non-parametric approaches to modeling age-specific mortality span from early model life tables to the nowadays well-established method of Lee and Carter (1992). After the first set of model life tables released by the United Nations (1955), Coale et al. (1966) developed a two-dimensional set of four regional patterns, each with 24 different mortality levels identified by the life expectancy of children. Brass (1971) presents a relational model that maps a tabulated standard age pattern of mortality with two parameters to actual mortality.

Lee and Carter (1992) apply principal component analysis and propose a model

$$\ln\left(m_{x,t}\right) = a_x + b_x k_t + \varepsilon_{x,t}$$

with mortality $m_{x,t}$ at age x and time t, fixed age effect a_x equal to the average observed log death rate, and an age-specific impact b_x of a time-specific general mortality index k_t. This single parameter k_t maps the average age pattern of mortality deviation from a_x to the actual pattern and b_x is the first principal component and is estimated by singular value decomposition. The subsequent estimation of the mortality index k_t as an ARIMA process results in a simple random walk with drift. The outcome, however, of forecasting age-specific mortality by this method with one time-dependent parameter is similar to that if each age-specific time series were extrapolated along its own historic time trend, potentially leading to an implausible age pattern in the long run.[5] This disadvantage is especially severe if the Lee–Carter approach is applied to single-cause mortality, for which it was not indeed assigned.[6] Nevertheless, the Lee–Carter method and its several enhancements have become the standard for mortality forecasts and have also been used for the newly emerged stochastic population forecasts since Lee and Tuljapurkar (1994) and Lee (1998).

There is broad literature introducing models more or less similar to the Lee–Carter approach. Lee (2000) reviews the original model as well as some of its problems and extensions. Quantitative comparisons of several recent models are given by Cairns et al. (2007, 2008), but they only apply data for the age classes 60–89, that is, model a relatively even part of the full pattern of age-specific mortality, which is of course of special interest for the insurance industry. Renshaw and Haberman (2006) include an additional cohort effect estimated in a two-step procedure. To overcome potential roughness De Jong and Tickle (2006) apply smoothing along the age dimension by restricting the impact of several k_t on particular age classes with a spline matrix.[7] Delwarde et al. (2007) apply smoothing with a roughness penalty for both the Lee–Carter and a Poisson log-bilinear model.

[5] This critique goes back to McNown (1992) and Alho (1992).

[6] Girosi and King (2008, pp. 38–42) discuss this point and give examples.

[7] In a different approach of a generalized linear model with Poisson errors, Currie et al. (2004) apply smoothing along both the age and time dimensions with splines and handle future values to be forecasted as missing values which are estimated simultaneously.

Pedroza (2006) applies Bayesian methodology to mortality forecasting and adopts it to a state space reformulation of the Lee–Carter model. Girosi and King (2008) also generalize the Lee–Carter method to an analysis with several principal components instead of considering only the first one. Nevertheless, they advocate a completely different approach and run Bayesian regressions on socio-economic time series as explanatory covariates for mortality. Their main purpose is to establish a formalized way to incorporate additional information about regularities along a cross-section dimension of mortality, which may comprise age, sex, country, or cause of death, and generate priors to express experts' assessments of these similarities.

3 A Bayesian State Space Model

The dynamics of age-specific demographic variables can be captured by models based on a latent common component as in Lee and Carter (1992). We follow this line of research and extend these models by including additional macro variables as covariates and relating them with the latent variable by a VAR. We assume an autoregressive (AR) process for the coefficients, which link the explanatory variables with the age-specific demographic variables, to ensure smoothness along the age dimension. For the estimation of this state space model we use Bayesian methods, providing an appropriate MCMC algorithm. Although in this paper we apply our model to mortality, we present it in a more general way for any age-specific demographic variable.

3.1 General Model

Given an observed demographic variable $d_{x,t}$ with age classes $x = 0, \dots, A$ and time periods $t = 1, \dots, T$, we can formulate the equation

$$d_{x,t} \;=\; \overline{d}_x + \beta_x z_t + \epsilon^d_{x,t} \tag{1}$$

with arithmetic mean $\overline{d}_x = \frac{1}{T}\sum_{t=1}^{T} d_{x,t}$ and explanatory variables $z_t \equiv [\kappa_t \;\; Y_t]'$, where κ_t is a $K \times 1$ vector of unobservables and Y_t is an $N \times 1$ vector of observed covariates. The corresponding coefficient vector $\beta_x \equiv [\beta^\kappa_x \;\; \beta^Y_x]$ is $1 \times M$, where β^κ_x is a $1 \times K$ vector and β^Y_x is a $1 \times N$ vector with $M = K + N$. We assume z_t and β_x follow vector autoregressive processes,

$$z_t \;=\; c + \phi_1 z_{t-1} + \phi_2 z_{t-2} + \cdots + \phi_p z_{t-p} + \epsilon^z_t\,, \tag{2}$$

$$\beta_x \;=\; \alpha_1 \beta_{x-1} + \alpha_2 \beta_{x-2} + \cdots + \alpha_q \beta_{x-q} + \epsilon^\beta_x\,, \tag{3}$$

where c is an $M \times 1$ vector of constants, $\phi_1, \dots, \phi_p$ are $M \times M$ matrices, and $\alpha_1, \dots, \alpha_q$ are $M \times M$ diagonal matrices. We assume $\epsilon^d_{x,t} \sim i.i.d.\ \mathcal{N}(0, \sigma^2_d)$ for the disturbances in Equation (1), $\epsilon^z_t \sim i.i.d.\ \mathcal{N}(0, \Sigma_z)$ for the disturbances in Equation (2), and $\epsilon^\beta_x \sim i.i.d.\ \mathcal{N}(0, \Sigma_\beta)$ for the disturbances in Equation (3), where the covariance matrix Σ_β is a diagonal matrix. Thus each component of β_x in fact follows an autoregressive process on its own. All disturbances are assumed to be independent of each other.

3.2 Special Case Lee–Carter

To give a more intuitive introduction to our model, we will show in the following that the Lee–Carter model can be seen as a special case of our model. We begin by assuming that $z_t \equiv \kappa_t$, dropping Equation (3), and specifying an extremely strong prior on $\phi_1, \phi_2, \ldots, \phi_q$, where we specify the prior on ϕ_1 very tightly around 1 and the prior on $\phi_2, \ldots, \phi_q$ very tightly around 0. Of course, this can be applied by subsequently strengthening the power of the priors. For the extreme case, when the priors are very dominant, information emerging from the data will be completely ignored for the VAR parameters $\phi_1, \phi_2, \ldots, \phi_q$ and we obtain, approximately, the model

$$d_{x,t} \;=\; \overline{d}_x + \beta_x^{\kappa}\kappa_t + \epsilon_{x,t}^{d} \tag{4}$$

with an AR process for the mortality index κ_t,

$$\kappa_t = c + \kappa_{t-1} + \epsilon_t^{\kappa}\;, \tag{5}$$

which is the Lee–Carter model in state space representation as described in Pedroza (2006).

3.3 Augmenting the Simple Model with Covariates

The inclusion of covariates may noticeably improve the forecasts of demographic models.[8] Respective time series provide additional information, which is ignored otherwise, if these covariates exhibit a possibly small but systematic impact on the demographic variable. Hence, in principle, the co-evolution of the demographic variable and its covariates should be modeled together. In our case, this means choosing $N > 0$, resulting in the full model with $z_t = [\kappa_t \quad Y_t]'$ instead of the simpler special case where $z_t = \kappa_t$, according to the Lee–Carter model. The informational gain of this inclusion depends of course on the specifications of the demographic variable and appropriate covariates and has to be weighted against the increased number of parameters to be estimated. By the vector autoregression in Equation (2), our model enables the requested utilization of covariates in an appropriate way. Nevertheless, this is only a further alternative to the parsimonious version without covariates, which already exhibits good forecasting features.

3.4 Smoothing Along the Age Dimension

When trying to predict future mortality, we have to consider the knowledge about its systematic pattern. To exemplify this point, we might have no idea in the first place about the level of mortality of a 40-year-old 50 years from now; nevertheless, we are very confident that this mortality is quite similar to that of a 41-year-old. Hence any forecast missing this basic feature with diverging developments of adjacent age classes should be mistrusted. As already discussed in Section 2.2, the Lee–Carter model cannot prevent potential implausible age patterns in out-of-sample forecasts. Our model mitigates this problem. Equation (3) guarantees smoothness along the age dimension because the coefficients $\beta_x, \ldots, \beta_{x-q}$ are

[8] This issue is discussed extensively in Girosi and King (2008).

connected by autoregressive processes. For $\frac{q}{2} \in \mathbb{N}$ and $\alpha_{q/2} \neq 0$, Equation (3) can easily be reformulated to get a symmetric representation of smoothing between adjacent age classes:[9]

$$\beta_{\tilde{x}} = \tilde{\alpha}_{-\frac{q}{2}}\,\beta_{\tilde{x}-\frac{q}{2}} + \cdots + \tilde{\alpha}_{-1}\,\beta_{\tilde{x}-1} + \tilde{\alpha}_{1}\,\beta_{\tilde{x}+1} + \cdots + \tilde{\alpha}_{\frac{q}{2}}\,\beta_{\tilde{x}+\frac{q}{2}} + \tilde{\epsilon}^{\beta}_{\tilde{x}}\,. \tag{6}$$

Assuring a plausible age pattern without jumps might be even more important when looking at more volatile data than in our example of current all-cause mortality from the United States, for example, as for the case of single-cause mortality or for data from non-industrialized countries in the past and present.

3.5 Cohort Effects

The general model described previously can theoretically be extended to also capture cohort effects. We just have to extend Equation (1) with an additional variable corresponding to the cohort dimension, which can be expressed as

$$d_{x,t} \;=\; \overline{d}_x + \beta_x z_t + \beta^{\gamma}_x \gamma_{t-x} + \epsilon^{d}_{x,t}\,. \tag{7}$$

With $N = 0$ Equation (7) is similar to the model described in Renshaw and Haberman (2006). One deviation from their model is that we assume the following law of motion:

$$\gamma_{t-x} = \varphi_1 \gamma_{(t-x)-1} + \varphi_2 \gamma_{(t-x)-2} + \cdots + \varphi_r \gamma_{(t-x)-r} + \epsilon^{\gamma}_t\,, \tag{8}$$

where ϵ^{γ}_t is not serially correlated and independent of $\epsilon^{d}_{x,t}$, ϵ^{z}_t, and ϵ^{β}_x at all leads and lags. The other deviation to Renshaw and Haberman (2006) is that they estimate Equation (7) in a two-step procedure, whereas we would be able to estimate the extended model in a more efficient one-step procedure by introducing an additional step to the Gibbs sampler described in Section 6.

3.6 Indeterminacies

In the estimation procedure we have to deal with three kinds of potential indeterminacies, namely, sign, scale, and rotational indeterminacies. The former two can be illustrated with the following example. Presume we multiply Equation (1) by $1 = \frac{\gamma}{\gamma}, \gamma \neq 0$; then we obtain

$$d_{x,t} \;=\; \overline{d}_x + (\beta^{\kappa}_x \gamma)\left(\frac{\kappa_t}{\gamma}\right) + \beta^{Y}_x Y_t + \epsilon^{d}_{x,t}\,. \tag{9}$$

Of course, this equation implies the same data-generating process as Equation (1), even though we have $\widetilde{\beta}^{\kappa}_x \equiv \beta^{\kappa}_x \gamma$ and $\widetilde{\kappa}_t \equiv \kappa_t/\gamma$ with different scale or sign than before. To solve these indeterminacies we need additional constraints. Following the Lee–Carter model, we impose $\sum_{t=0}^{T} \kappa^{k}_t = 0$ and $\sum_{x=0}^{A} \beta^{k}_x = 1$ for all $k \in \{1, \ldots, K\}$. In the case of $K > 1$ an additional rotational indeterminacy occurs, because appropriate rotations yield

$$d_{x,t} = \overline{d}_x + \left(\beta_x P'\right)\left(P z_t\right) + \epsilon^{d}_{x,t}\,,$$

[9] Set $\alpha_0 \equiv -1$, $\tilde{\alpha}_i \equiv -\frac{\alpha_{(q/2)-i}}{\alpha_{q/2}}$ for $i \in \{-\frac{q}{2}, \ldots, \frac{q}{2}\}$, $\tilde{x} \equiv x - \frac{q}{2}$, and $\tilde{\epsilon}^{\beta}_{\tilde{x}} \equiv -\frac{\epsilon^{\beta}_x}{\alpha_{q/2}}$.

where

$$P = \begin{bmatrix} A & B \\ 0 & I \end{bmatrix}$$

is an orthogonal matrix with $\tilde{\tilde{\beta}}_x \equiv \beta_x P'$ and $\tilde{\tilde{z}}_t \equiv P z_t$, implying the same data-generating process as Equation (1). Sufficient conditions for unique identification are to set the lower $K \times K$ block of β_x^κ to a diagonal matrix and the lower $K \times N$ block of β_x^Y to zero.[10]

4 Predictive Densities

In order to derive analytically distributional statements on the probabilities of outcomes we describe the posterior predictive densities corresponding to the future path of the demographic variables up to horizon H. In this context we find it useful to define

$$d_x^H \equiv [d_{x,T+1} \ \ldots \ d_{x,T+H}] \,,$$

$$d_x^T \equiv [d_{x,1} \ \ldots \ d_{x,T}] \,,$$

$$z \equiv [z_1 \ z_2 \ \ldots \ z_T] \,,$$

$$\beta \equiv [\beta_0 \ \beta_1 \ \ldots \ \beta_A]' \,,$$

$$\Psi \equiv \left\{(c, \phi_1, \phi_2, \ldots, \phi_p, \Sigma_z), (\alpha_1, \alpha_2, \ldots, \alpha_q, \Sigma_\beta), (\sigma_d^2)\right\} \,.$$

Thus the posterior predictive density can be expressed as

$$p\left(d_x^H | d_x^T\right) = \int\int\int p\left(d_x^H | z, \beta, \Psi, d_x^T\right) \ p\left(z, \beta, \Psi, | d_x^T\right) \ dz \ d\beta \ d\Psi \,.$$

In order to obtain values for the future path of the observations we draw ϵ_{T+i}^z from $\mathcal{N}(0, \Sigma_z)$ for $i = 1, \ldots, H$ and iterate on

$$z_{T+i} = c + \phi_1 z_{T+i-1} + \phi_2 z_{T+i-2} + \cdots + \phi_p z_{T+i-p} + \epsilon_{T+i}^z \,. \tag{10}$$

Following this, we use the values from Equation (10), draw $\epsilon_{x,T+i}^d$ from $\mathcal{N}(0, \sigma_d^2)$, and iterate on

$$d_{x,T+i} \ = \ \overline{d}_x + \beta_x z_{T+i} + \epsilon_{x,T+i}^d$$

to get draws from the joint posterior distribution of d_x^H.

5 Priors

We introduce priors on the VAR parameters via dummy observations by simulating an artificial dataset with certain assumed properties and add it to our actual dataset. This goes back to the *mixed estimation* procedure suggested by Theil and Goldberger (1961) and was recently applied by Sims and Zha (1998) and Del Negro and Schorfheide (2004). We generate dummy observations, implying that the series produced include a

[10]This is similar to the dynamic factor literature. See, among others, Geweke and Zhou (1996) and Bernanke et al. (2005).

random walk process. We do this by centering the probability mass for the first lagged coefficient around 1 and for all subsequent lags around 0, while we subsequently decrease the uncertainty that the coefficients are zero for more distant lags.

We consider the following model:

$$Z^* = X^*\Phi^* + \epsilon^* \,, \tag{11}$$

where

$$Z^* \equiv \begin{bmatrix} \lambda_1\hat{\sigma} \\ 0_{M(p-1)\times M} \end{bmatrix}$$

and

$$X^* \equiv \begin{bmatrix} \lambda_1\hat{\sigma} & 0 & \cdots & 0 \\ 0 & 2\lambda_1\hat{\sigma} & 0 & \vdots \\ \vdots & 0 & \ddots & 0 \\ 0 & \cdots & 0 & p\lambda_1\hat{\sigma} \end{bmatrix},$$

with

$$\hat{\sigma} \equiv \begin{bmatrix} \hat{\sigma}_1 & 0 & \cdots & 0 \\ 0 & \hat{\sigma}_2 & 0 & \vdots \\ \vdots & 0 & \ddots & 0 \\ 0 & \cdots & 0 & \hat{\sigma}_M \end{bmatrix},$$

where λ_1 is called the *overall tightness of beliefs* around the random walk prior and $\hat{\sigma}_1, \hat{\sigma}_2, \ldots, \hat{\sigma}_M$ are the empirical standard deviations taken from the first p observations. Increasing values for λ_1 imply that we are more certain concerning our prior and hence the prior gets more weight in comparison to information emerging from the dataset via the likelihood function. Taken values for Σ_z as given, the dummy observations imply the following conjugate prior for our VAR parameters:

$$\Phi^*|\Sigma_z \sim \mathcal{N}\left(vec(\hat{\Phi^*}), \Sigma_z \otimes (X^{*\prime}X^*)^{-1}\right) . \tag{12}$$

The prior for the AR parameters in Equation (3) is similar to the one specified for the VAR parameters with λ_2 as the overall tightness of beliefs of the prior. For the variance of the disturbance in Equation (1) we assume an inverted gamma distribution $\mathcal{IG}(\frac{\tau_1}{2}, \frac{\tau_2}{2})$.

6 Estimation

We estimate our model using MCMC methods; more precisely, we apply the Gibbs sampler. This method enables us to draw from the joint distribution $\mathcal{P}(\Psi, z, \beta)$ by subdividing it into the conditional distributions $\mathcal{P}(\Psi \mid z, \beta)$, $\mathcal{P}(z \mid \Psi, \beta)$, and $\mathcal{P}(\beta \mid \Psi, z)$ and draw iteratively from them. Taking initialized values for $z^{(0)}$ and $\beta^{(0)}$ as given, we sample in the ith iteration $\Psi^{(i)}$ from $\mathcal{P}(\Psi \mid z^{(i-1)}, \beta^{(i-1)})$, $z^{(i)}$ from $\mathcal{P}(z \mid \Psi^{(i)}, \beta^{(i-1)})$, and $\beta^{(i)}$ from $\mathcal{P}(\beta \mid \Psi^{(i)}, z^{(i)})$ successively. Under weak conditions and for $i \to \infty$ the Gibbs sampler converges and we obtain samples from the desired joint distribution $\mathcal{P}(\Psi, z, \beta)$.[11] For a more detailed description of the estimation procedure see Appendix A.

[11] Cf. Geman and Geman (1984).

7 Data

We apply our model to age-specific total (combining female and male) mortality data from the United States with 91 individual age classes from 0 to 90 as shown in Figure 1 as specification of the demographic variable $d_{x,t}$.[12] These time series provided by the Human Mortality Database span the period 1933–2005, of which we use the post-World War II period.[13] We add macroeconomic time series of real GDP per capita and of unemployment, which are displayed in Figure 2. The data for real GDP per capita are expressed in logarithms of chained 2000 Dollars, and the unemployment rate is measured as the number of unemployed as a percentage of the civilian labor force.[14]

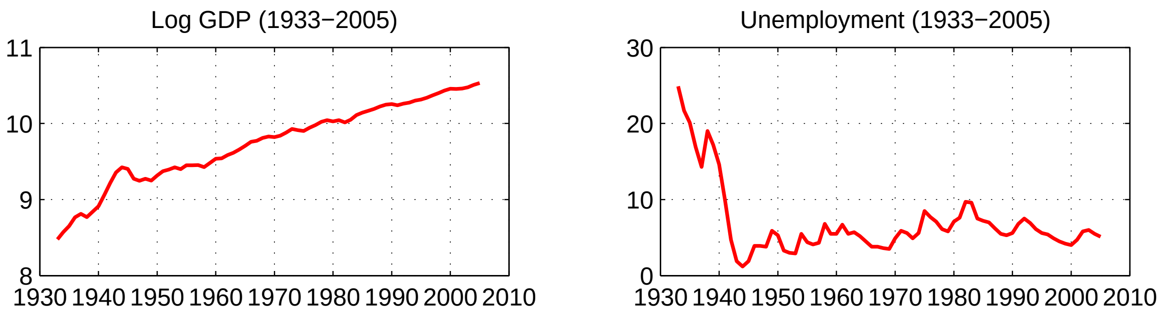

Figure 2: Logarithmized GDP and unemployment rate for the United States 1946–2005.

8 Results

We apply our model to mortality data from the United States in the period 1946–2005 and gradually vary the model specification. With the objective of comparing with the Lee–Carter results, we first assume κ_t to consist of only one unobserved time series, which may be called mortality index, and abstain from using covariates. Afterward, the macroeconomic time series are included as covariates.

[12] Unlike Lee and Carter (1992), where each age class comprises 5 years, we refrain from age grouping and keep the detailed information of single age classes.

[13] C.f. Human Mortality Database (2008). In the Human Mortality Database raw data are corrected for obvious mistakes and, for the calculation of life tables, death rates for the age classes 80 and above are smoothed by fitting a logistic function according to Thatcher et al. (1998) if the number of observations becomes too small. Wilmoth et al. (2007) supply a detailed method protocol. In the case of the United States, population estimates for 1940–1969 are adjusted to exclude the Armed Forces overseas and to correct for the inclusion of Alaska and Hawaii. Moreover, owing to the lack of data for the age classes 75 and above in the period 1933–1939, the extinct cohort method is applied as supposed by Kannisto (1994).

[14] Although the pre-1947 unemployment figures refer to persons aged 14 and above, whereas the post-1947 figures refer to persons aged 16 and above, this minor change causes no jump in 1947, when both definitions yield the same number. With respect to GDP and the unemployment rate, see the U.S. Census Bureau (2007).

8.1 Preliminaries

For the results we used a lag length of $p = 4$ for the z's and $q = 4$ for the β's. The prior specifications, which we describe in Section 5, are $\lambda_1 = 5$ for the VAR parameters of z and a flat prior $\lambda_2 = 0$ for the AR parameters of β.[15] For the variance of the disturbances in Equation (1) we choose $\tau_1 = 0.01$ and $\tau_2 = 3$.

The estimation results may be affected by the choice of time period and age span under consideration. To check whether our results depend on the initial β parameters we conduct the following exercise. We leave out mortality of the youngest age classes and estimate our model with $\beta_s, \ldots, \beta_A$, where $s > 0$. We obtain very similar results to the full model $\beta_s, \ldots, \beta_A$, suggesting that the choice of initial values for the β's does not bias our results. With respect to the time period we mainly focus on the postwar era 1946–2005 to base the analysis and forecasts on circumstances relatively close to the present and to avoid the influence of very high unemployment after the Great Depression and possible distortions from World War II. Nevertheless, we also test for specifications that span the entire period 1933–2004 and get very similar results for the forecasts.

To ensure that our Gibbs sampler converges we restart the algorithm several times, each time using different starting values drawn from an overdispersed distribution. The results for all these different chains are very similar. Our sampler already reaches convergence after a few thousand draws. Furthermore, to keep the starting values from influencing our results we discard the first half of the chain as the burn-in phase.

8.2 One Kappa but No Covariates ($K = 1$, $N = 0$)

First we present the simplest version, with only one latent variable κ and no covariates. Figure 3 shows the estimated κ and the corresponding coefficient matrix β, which reveals how close the mortality of particular age classes is associated with development of the latent variable κ. The age classes 0–15 are higher than average exposed to κ; however, all age classes are positively related to the latent variable.

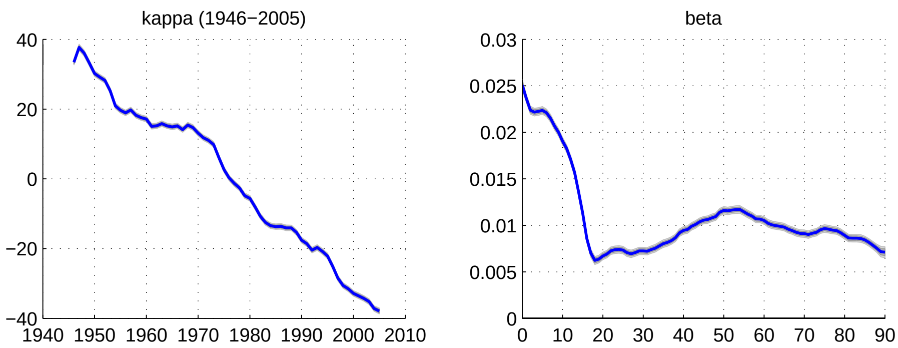

Figure 3: Estimated κ and β. The small gray shaded area around the blue median represents 90% of the posterior probability mass regarding both parameter and random term uncertainty.

[15] $\lambda_1 = 5$ is also used by, among others, Sims and Zha (1998).

In Figure 4 we show different in-sample forecasts for κ over a 15-year horizon from 1991 onward, that can be compared with the "realized" developing (red line), which means the median of the estimated κ for the entire period.

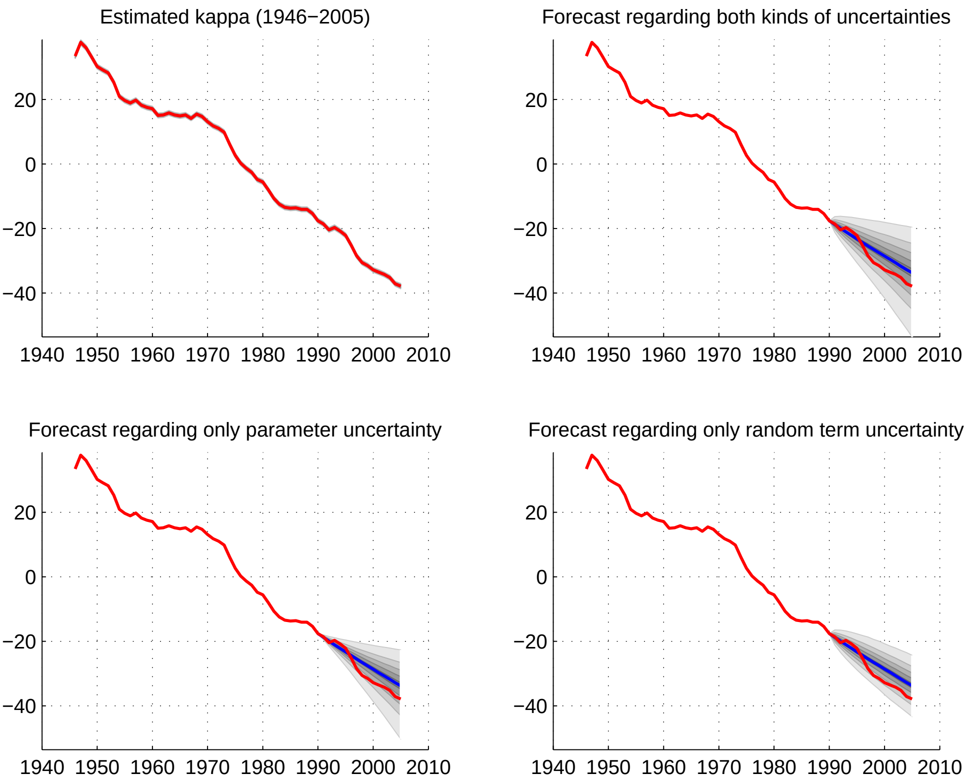

Figure 4: Panel with in-sample forecasts of κ with respect to different sources of uncertainty for the period 1991–2005. The red line always displays the median estimation of κ based on the observations for the whole period 1946–2005. The blue line displays the median forecast of κ based only on the information up to 1990. The entire gray shaded area represents 90% of the posterior probability mass and each of the different gray shaded bands represents 10% of the posterior probability mass. Note that the innermost band is largely covered by the blue line.

Additionally, we show in Figure 5 out-of-sample forecasts for a longer horizon up to the year 2050. These forecasts are of course subject to different kinds of uncertainty. In each case, we give an overview of forecasts, where either only the uncertainty due to the random terms ϵ, only the uncertainty due to the estimation of the parameters of the model, or both kinds of uncertainty are considered. The resulting distributional features of the forecasts are illustrated by the probability mass around the medium forecast. In all cases, accounting only for the random term uncertainty results in quite close forecasts which have the form of a parabola and widen only a little over time. In contrast to this, the forecasts accounting only for parameter uncertainty start very narrow but widen faster than they do linearly. The forecasts with respect to both sources of uncertainty are of course the widest. In this case, the overall accuracy of the forecast is dominated by the effect of the random term in the short run and by the effect of the parameter estimation in the long run.[16] This result demonstrates the extent to which presentations of forecasts can be misleading by giving rise to an illusion of sureness if important sources of uncertainty are ignored. Moreover, even the most precautious versions of our plots give only lower bounds for the real forecast uncertainty, which can be even larger, because the specification of the model (model choice) and the estimation of κ in the observation period (starting point for the forecast) are also nondeterministic.

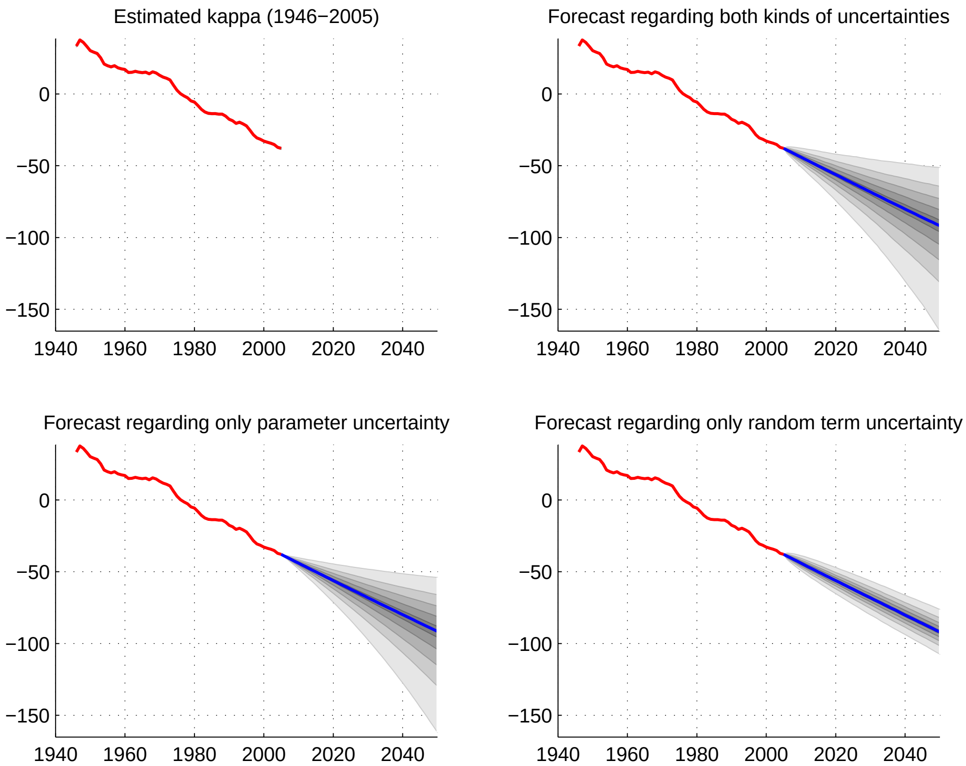

Figure 5: Panel with long-run forecasts of κ with respect to different sources of uncertainty for the period 2006–2050. The red line displays the median estimation of κ based on the observations in the period 1946–2005. The blue line displays the median forecast of κ based on this information. The entire gray shaded area represents 90% of the posterior probability mass and each of the different gray shaded bands represents 10% of the posterior probability mass. Note that the innermost band is largely covered by the blue line.

8.3 Improving Forecasts, with Covariates ($K = 1$, $N = 2$ and $K = 2$, $N = 2$)

In order to improve our predictions we extend our model by including logarithmized real GDP per capita and the unemployment rate as covariates and, in a further step, by adding a second latent variable κ_2 to the specification with the two covariates. Figure 6 shows

[16] Lee and Carter (1992) mention a dissenting relationship in their Appendix B.

the estimated coefficients β related to κ_1 and κ_2, GDP, and unemployment, revealing the extent to which age-specific mortality is affected by the latent variables and covariates. Of course, this paves the way for structural analysis of the systematic interactions of mortality and covariates using impulse responses analyses, which is presented in detail in Reichmuth and Sarferaz (2008).

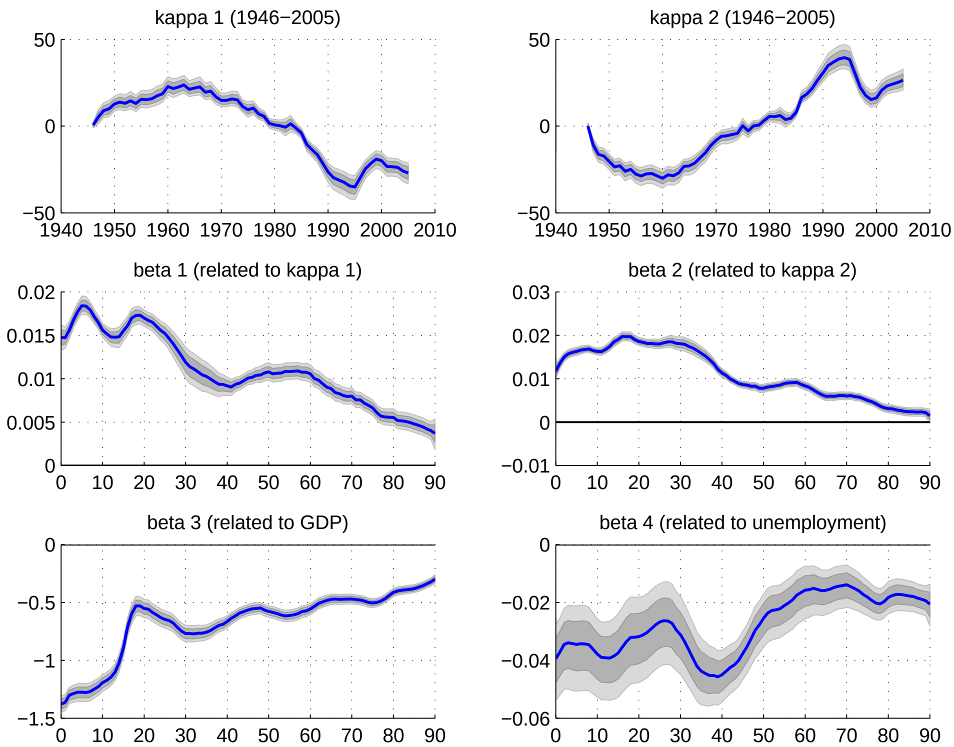

Figure 6: Estimated κ's and β's for the model specification with two latent variables and GDP and unemployment as covariates. The entire gray shaded area around the blue median represents 90% of the posterior probability mass and the dark gray shaded area represents 68% of the posterior probability mass regarding both parameter and random term uncertainty.

For the simplest specification without covariates, Figure 7 shows the median out-of-sample forecasts of age-specific mortality about the middle and at the end of the forecast period in comparison to actual observations. As can be seen, the overall level of mortality declines steadily but the shape stays more or less the same. Figure 8 shows different out-of-sample forecasts for the longer horizon until 2050, where the error bands widen by time. As can be seen in the first and second rows of Figure 8, including macro variables as covariates improves the forecasts for the higher age classes, whereas the forecasts for the age classes 15–40 deteriorate. This leads us to the conclusion that covariates have to be chosen very carefully in general, as they might help predict particular age classes but at the same time worsen the forecasts of others. The third row of Figure 8 shows that adding κ_2 to the specification with two covariates improves the forecasts again. For the age classes below 15 or above 40, they are the best of all specifications.

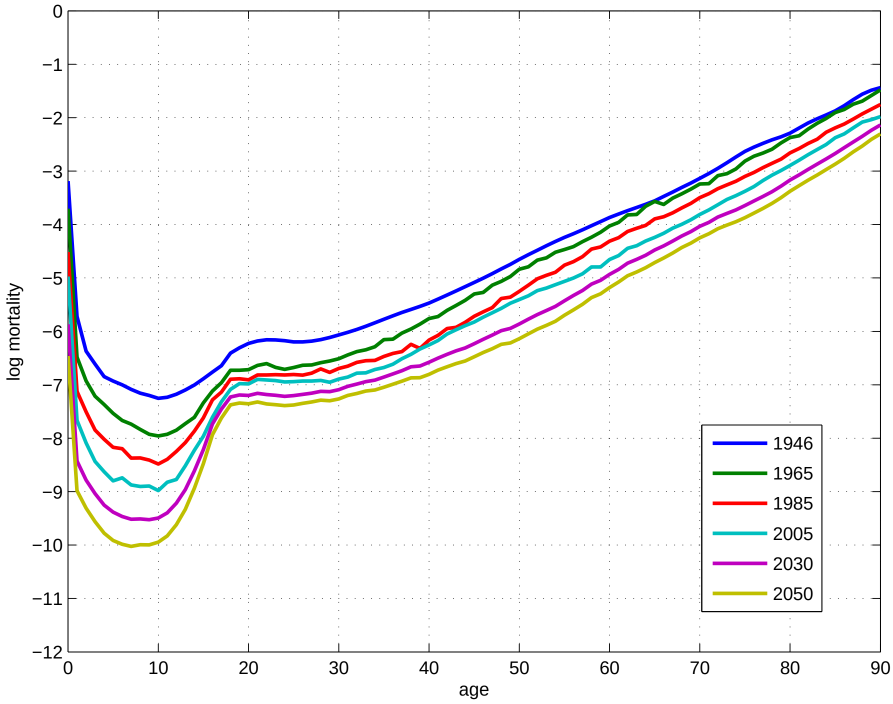

Figure 7: Observations and forecasts of age-specific mortality $m_{x,t}$ at different points in time. The lines for the years 2030 and 2050 display the median forecasts regarding both parameter and random term uncertainty.

The figures discussed in this section demonstrate the smooth transition along the age dimension as described in Section 3.4. Admittedly, the difference to the Lee–Carter results is not so obvious owing to their previous age grouping, but note that we prevent divergence for single age classes in the long-run independent of the choice of all-cause mortality.

The forecast errors presented in this paper can be interpreted differently, depending on the particular research interest of the reader. For example, overestimating future mortality may jeopardize pension schemes, whereas underestimating is a danger for life insurance calculations. In both cases major misjudgments have more severe consequences for the stakeholders than smaller ones. This means that not only mean and variance but also higher moments (skewness and kurtosis) of the distribution of predicted mortality matter for the risk assessment. Our Bayesian presentation of the forecast results with a detailed allocation of probability masses provides the information needed.

Moreover, the relatively wide dispersion of our forecasts assigns only a rather low probability for realizations close to the median, which further challenges traditional forecast methods with misleadingly tight error bands.

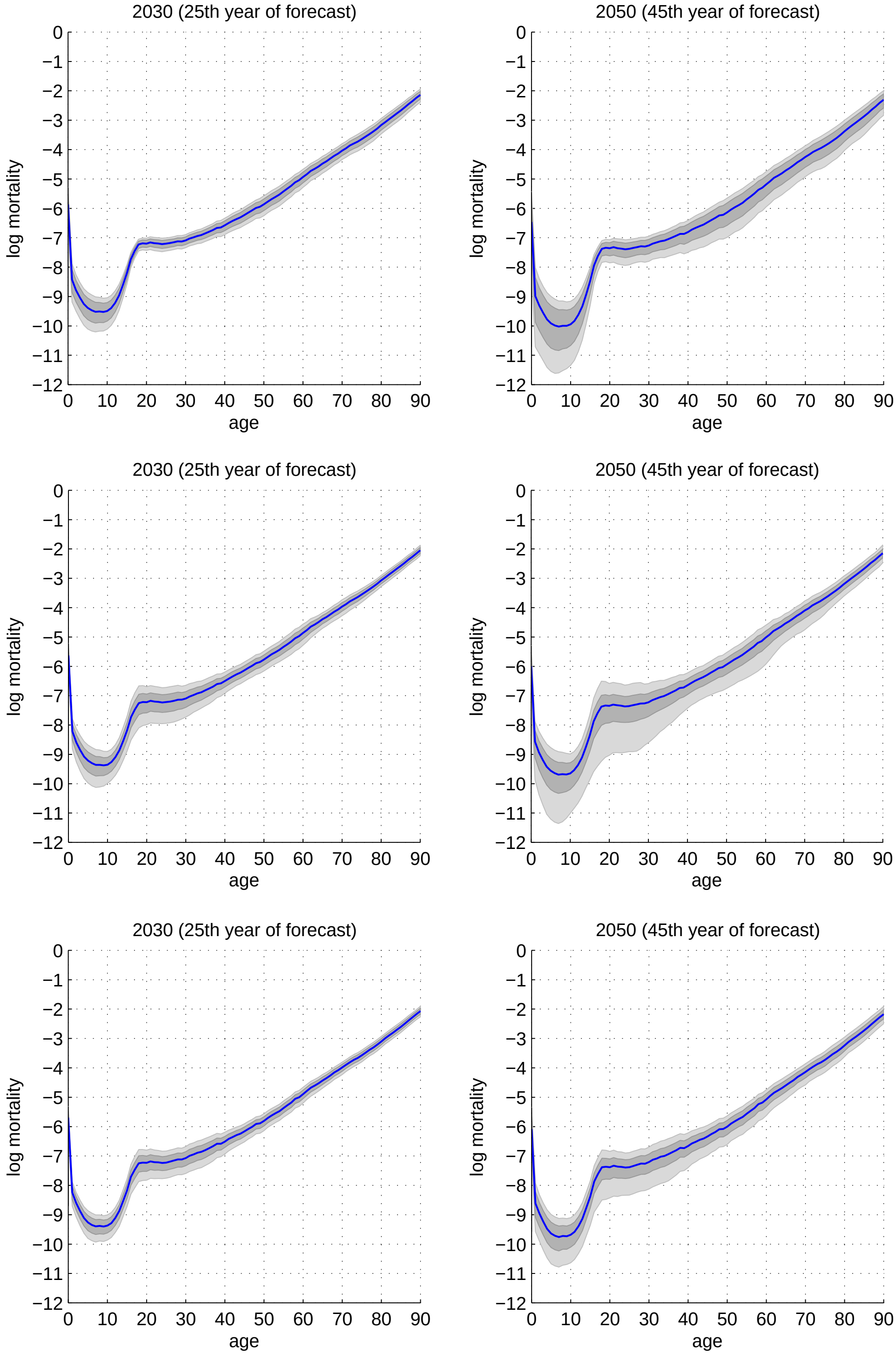

Figure 8: Panel with forecasts of age-specific mortality $m_{x,t}$ 25 and 45 years ahead for different model specifications. The first row shows the specification for $K = 1$, $N = 0$, the second row for $K = 1$, $N = 2$, and the third row for $K = 2$, $N = 2$. The entire gray shaded area around the blue median represents 90% of the posterior probability mass and the dark gray shaded area represents 68% of the posterior probability mass regarding both parameter and random term uncertainty.

8.4 Life Tables

Life tables deliver some intuitively interpretable variables such as surviving probabilities and life expectancies, which can be calculated from a complete set of age-specific mortalities. For this purpose, we use the simplest specification of our model with one latent variable κ and no covariates to forecast mortality for all age classes up to 110+.[17] We do so for female and male mortality separately, because the resulting life tables are quite different and would not be represented adequately by a version for "total" mortality. Finally, we compute the respective period life tables up to the year 2050 and present the results for females. The detailed calculations are given in Appendix B. Note that the life table variables depend nonlinearly on a whole set of mortalities at different ages. Thus, to get proper percentiles for the forecasts of these variables, we do not use percentiles of age-specific mortality directly but compute the life tables from the particular mortalities for the second half of 30,000 independent draws separately. Once again, the error bands with respect to both parameter and error term uncertainty are the widest.

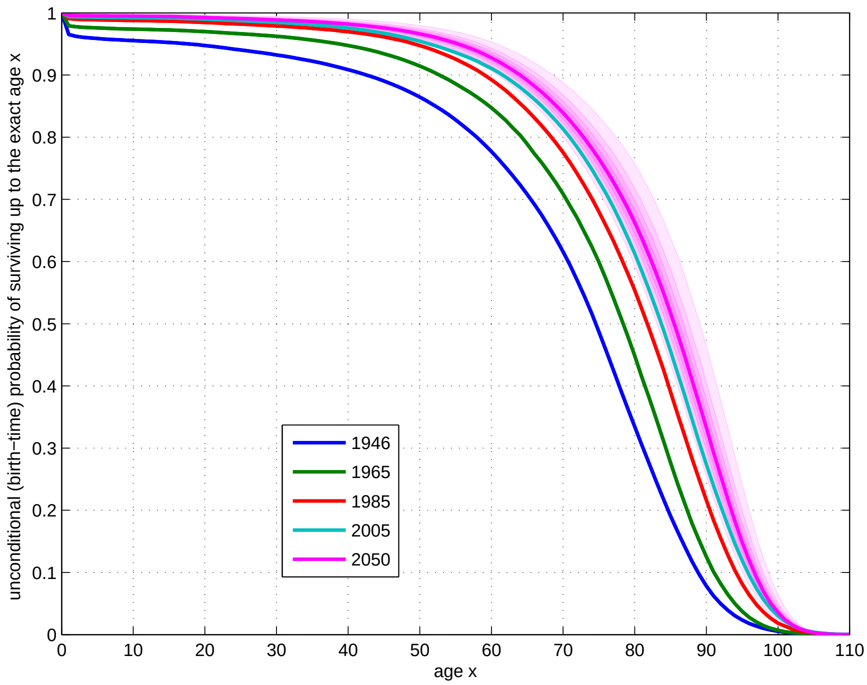

Figure 9: Probabilities $l_{x,t}$ of surviving up to the exact age x for females based on period life tables for different points in time. The figures for the years 1946–2005 are calculated from observations. The thick magenta line displays the median forecast of $l_{x,2050}$. The entire magenta shaded area represents 90% of the posterior probability mass and each of the different magenta shaded bands represents 10% of the posterior probability mass regarding both parameter and random term uncertainty. Note that the innermost band is largely covered by the thick line for the median.

[17] The inclusion of very high ages is necessary for the best possible calculation of remaining life expectancies.

Figure 9 displays the hypothetical birth-time probabilities $l_{x,t}$ of surviving up to the exact age x if a female were subject to the age-specific mortalities of one particular period over her whole life cycle. During the observation period 1946–2005 the curves consistently move upward and to the right. First, reductions of child mortality mainly shift the curve upward, whereas later on reductions of old-age mortality shift it to the right. The forecast for 2050 shows that this trend will probably continue, though the error bands show the relatively high uncertainty about the future survival curve. However, the forecast accuracy of the life table variables, which depend in particular on old-age mortality, can also be improved by the inclusion of covariates.

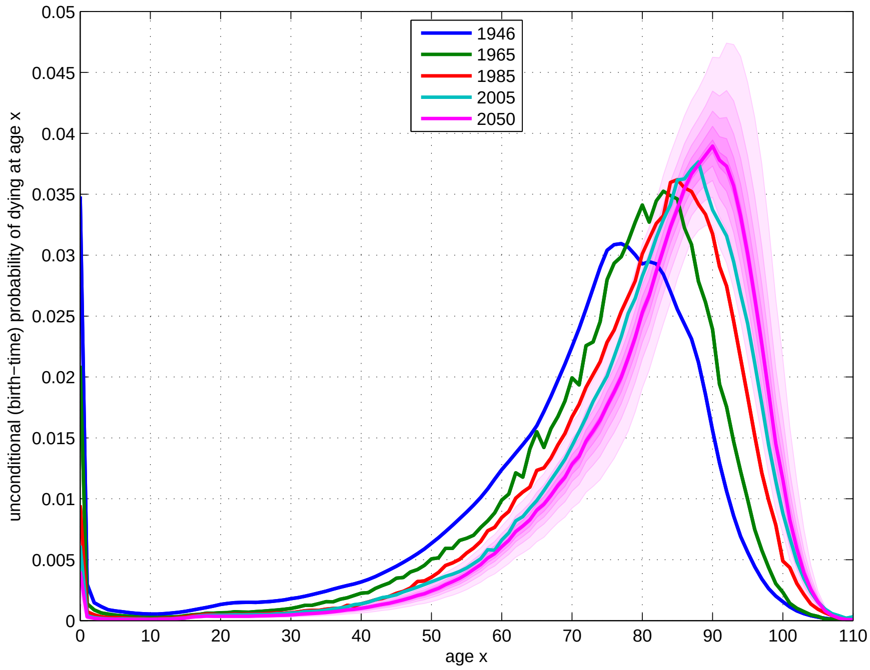

Figure 10: Probabilities $d_{x,t}$ of dying at age x for females based on period life tables for different points in time. The figures for the years 1946–2005 are calculated from observations. The thick magenta line displays the median forecast of $d_{x,2050}$. The entire magenta shaded area represents 90% of the posterior probability mass and each of the different magenta shaded bands represents 10% of the posterior probability mass regarding both parameter and random term uncertainty. Note that the innermost band is largely covered by the thick line for the median.

Figure 10 displays the corresponding birth-time probabilities $d_{x,t}$ of dying at age x. Of course, the values rise over most of the lifetime and peak somewhere in old age before they fall again.[18] Remarkably, these probabilities not only shift to the right but also concentrate increasingly on a smaller age range. With respect to the survival curve, this corresponds to a transformation toward a long relatively flat initial course followed by a steep fall, which is known as rectangularization.

[18]In today's industrialized countries child mortality is no longer a major threat.

Finally, in Figure 11 we present time series of life expectancies at different ages for the whole observation plus the forecast period 1946–2050. Life expectancy always means the remaining life expectancy for those who have already achieved a particular age. In our application, the life expectancies of older people are always lower than those of younger people, because there is no phase of life with such a high mortality that survivors of this phase would have a higher remaining life expectancy than younger people prior to this phase. The life expectancies for all age classes increase quite evenly over time. The rise for the younger people is the strongest, because they benefit from the mortality reduction at all age classes lying ahead of them. Our forecasts clearly show that the trend of increasing life expectancies at all age classes will continue with high probability. For example, the median forecast of the gain in female life expectancy based on period life tables between 2005 and 2050 is about 4.5 years for a newborn and 2.8 years for a 60-year-old. Once again, the error bands of the forecasts can be further reduced by including covariates.

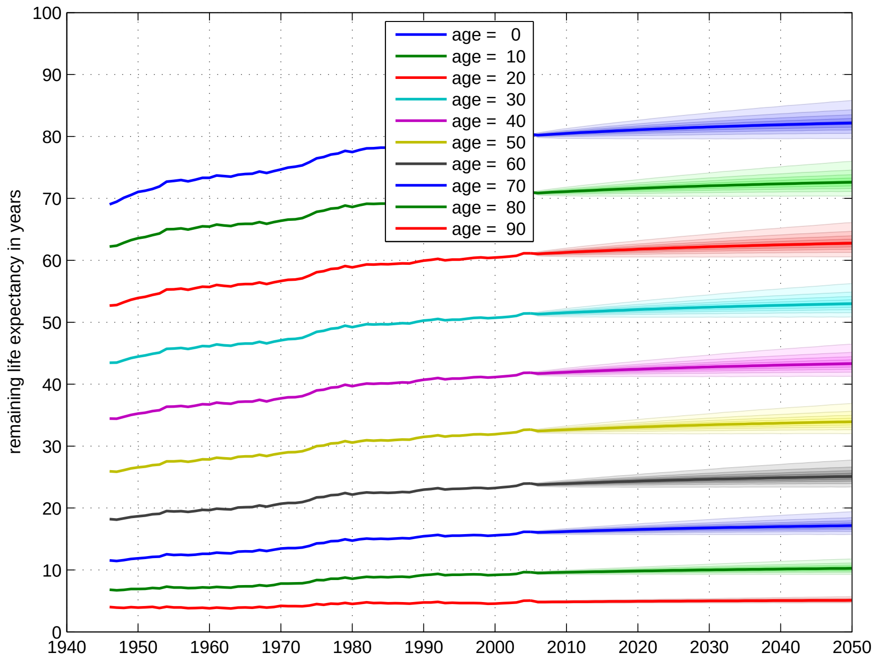

Figure 11: Remaining life expectancies $e_{x,t}$ for females of different age classes based on period life tables. The thick lines display figures calculated from observations in the period 1946–2005 and median forecasts for $e_{x,t}$ in the period 2006–2050. For each age class the entire shaded area represents 90% of the posterior probability mass and the different shaded bands represent 10% of the posterior probability mass regarding both parameter and random term uncertainty. Note that some of the bands are largely covered by the thick lines for the medians.

9 Conclusion

In this paper we present an alternative approach to modeling age-specific mortality. We build on the model introduced in Lee and Carter (1992) and extend it in several dimensions. We incorporate covariates and model their dynamics jointly with the latent variable underlying mortality of all age classes by a VAR process. Furthermore, we resolve the shortcomings in the age dimension from which previous models suffered by connecting adjacent age groups through an AR process. Our new modeling approach thus allows for consistent forecasts of age-specific mortality and the other variables.

We develop an appropriate MCMC algorithm, which enables us to estimate the parameters and latent variables jointly in an efficient one-step procedure. With our Bayesian approach we formalize priors for the parameters and thus include information into our model in a formal way. Additionally, we are able to assess uncertainty intuitively by constructing error bands for our forecasts.

We apply our model to U.S. mortality for 1946–2005 and test its forecast ability by means of in-sample and out-of-sample forecasts up to the year 2050. Our model performs well, that is, the forecasts exhibit smoothness along the age dimension with sufficiently tight error bands. Comparing different specifications, it turns out that covariates can indeed help improve the forecasts for particular age classes. Moreover, we demonstrate that uncertainty stemming from the error term is more important in the short run, whereas parameter uncertainty is very important for long-run forecasts. This points to the danger that existing forecasting methods for age-specific mortality, which ignore certain sources of uncertainty, may yield misleadingly sure predictions.

The link we provide between age-specific mortality and covariates can be exploited in a more structural way than is pursued in this present paper. An analysis of this relationship is conducted in Reichmuth and Sarferaz (2008).

References

ALHO, J. M. (1992). Modeling and forecasting U.S. mortality: Comment. *Journal of the American Statistical Association*, 87(419), 673–674.

BERNANKE, B. S., J. BOIVIN AND P. ELIASZ (2005). Measuring the effects of monetary policy: A factor-augmented vector autoregressive (FAVAR) approach. *Quarterly Journal of Economics*, 120(1), 387–422.

BOOTH, H. (2006). Demographic forecasting: 1980 to 2005 in review. *International Journal of Forecasting*, 22, 547–581.

BRASS, W. (1971). On the scale of mortality. In: BRASS, W. (ed.). *Biological Aspects of Demography.* Taylor and Francis, London, 69–110.

CAIRNS, A. J. G., D. BLAKE, K. DOWD, G. D. COUGHLAN, D. EPSTEIN AND M. K. ALLAH (2008). Mortality density forecasts: An analysis of six stochastic mortality models. *Pensions Institute Discussion Paper*, PI-0801.

CAIRNS, A. J. G., D. BLAKE, K. DOWD, G. D. COUGHLAN, D. EPSTEIN, A. ONG AND I. BALEVICH (2007). A quantitative comparison of stochastic mortality models using data from England and Wales and the United States. *Pensions Institute Discussion Paper*, PI-0701.

CARTER, C. K. AND R. KOHN (1994). On Gibbs sampling for state space models. *Biometrika*, 81(3), 541–553.

COALE, A. J. AND P. G. DEMENY (1983). *Regional Model Life Tables and Stable Populations.* Academic Press, New York.

COALE, A. J., P. G. DEMENY AND B. VAUGHAN (1966). *Regional Model Life Tables and Stable Populations.* Princeton University Press, Princeton.

CURRIE, I. D., M. DURBAN AND P. H. C. EILERS (2004). Smoothing and forecasting mortality rates. *Statistical Modelling*, 4(1), 279–298.

DE JONG, P. AND L. TICKLE (2006). Extending Lee–Carter mortality forecasting. *Mathematical Population Studies*, 13, 1–18.

DEL NEGRO, M. AND F. SCHORFHEIDE (2004). Priors from general equilibrium models for VARs. *International Economic Review*, 45(2), 643–673.

DELWARDE, A., M. DENUIT AND P. H. C. EILERS (2007). Smoothing the Lee–Carter and Poisson log-bilinear models for mortality forecasting: A penalized log-likelihood approach. *Statistical Modelling*, 7(1), 29–48.

FRÜHWIRTH-SCHNATTER, S. (1994). Data augmentation and dynamic linear models. *Journal of Time Series Analysis*, 15, 183–202.

Geman, D. and S. Geman (1984). Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-6, 721–741.

Geweke, J. and G. Zhou (1996). Measuring the pricing error of the arbitrage pricing theory. *Review of Financial Studies*, 9(2), 557–587.

Girosi, F. and G. King (2008). *Demographic Forecasting.* Princeton University Press, Princeton.

Gompertz, B. (1825). On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies. *Philosophical Transactions of the Royal Society of London*, 115, 513–585.

Graunt, J. (1662). *Natural and Political Observations Mentioned in a following Index, and made upon the Bills of Mortality.* John Martyn and James Allestry, London.

Halley, E. (1693). An estimate of the degrees of the mortality of mankind, drawn from curious tables of the births and funerals at the city of Breslaw; with an attempt to ascertain the price of annuities upon lives. *Philosophical Transactions*, 17, 596–610, 654–656.

Heligman, L. and J. H. Pollard (1980). The age pattern of mortality. *Journal of the Institute of Actuaries*, pp. 49–80.

Human Mortality Database (2008). *Human Mortality Database.* University of California, Berkeley, and Max Planck Institute for Demographic Research, Rostock. URL: `http://mortality.org` [data download 17/04/2008].

Kannisto, V. (1994). *Development of Oldest-old Mortality, 1950–1990: Evidence from 28 Developed Countries*, volume 1 of *Odense Monographs on Population Aging.* Odense University Press, Odense, Denmark.

Kim, C.-J. and C. R. Nelson (1999). *State-Space Models with Regime-Switching: Classical and Gibbs-Sampling Approaches with Applications.* MIT Press, Cambridge, Massachusetts.

Lee, R. D. (1998). Probabilistic approaches to population forecasting. In: Lutz, W., J. W. Vaupel and D. A. Ahlburg (eds.). *Frontiers of Population Forecasting*, Supplement to *Population and Development Review*, 24, 156–190.

——— (2000). The Lee–Carter method for forecasting mortality, with various extensions and applications. *North American Actuarial Journal*, 4(1), 80–93.

Lee, R. D. and L. R. Carter (1992). Modeling and forecasting U. S. mortality. *Journal of the American Statistical Association*, 87(419), 659–671.

Lee, R. D. and S. Tuljapurkar (1994). Stochastic population forecasts for the United States: Beyond high, medium, and low. *Journal of the American Statistical Association*, 89(428), 1175–1189.

McNown, R. (1992). Modeling and forecasting U.S. mortality: Comment. *Journal of the American Statistical Association*, 87(419), 671–672.

McNown, R. and A. Rogers (1989). Forecasting mortality. A parameterized time series approach. *Demography*, 26(4), 645–660.

Pedroza, C. (2006). A Bayesian forecasting model: Predicting U.S. male mortality. *Biostatistics*, 7(4), 530–550.

Preston, S. H., P. Heuveline and M. Guillot (2005). *Demography. Measuring and Modeling Population Processes.* Blackwell Publishing, Malden, Massachusetts.

Reichmuth, W. H. and S. Sarferaz (2008). The influence of the business cycle on mortality. *SFB 649 Discussion Paper*, 2008-059.

Renshaw, A. E. and S. Haberman (2006). A cohort-based extension to the Lee–Carter model for mortality reduction factors. *Insurance: Mathematics and Economics*, 38, 556–570.

Sims, C. A. and T. Zha (1998). Bayesian methods for dynamic multivariate models. *International Economic Review*, 39(4), 949–968. Symposium on Forecasting and Empirical Methods in Macroeconomics and Finance.

Thatcher, A. R., V. Kannisto and J. W. Vaupel (1998). *The Force of Mortality at Ages 80 to 120*, volume 5 of *Odense Monographs on Population Aging.* Odense University Press, Odense, Denmark.

Theil, H. and A. S. Goldberger (1961). On pure and mixed statistical estimation in economics. *International Economic Review*, 2(1), 65–78.

United Nations (1955). *Age and Sex Patterns of Mortality: Model Life Tables for Underdeveloped Countries.* United Nations publication E.55.XIII.9. United Nations, New York.

U.S. Census Bureau (2007). *The 2007 Statistical Abstract, The National Data Book.* U.S. Census Bureau, Washington, D.C. URL: `http://www.census.gov/compendia/statabl` [data download 27/08/2007].

Wilmoth, J. R., K. Andreev, D. Jdanov and D. A. Glei (2007). *Methods Protocol for the Human Mortality Database.* University of California, Berkeley, and Max Planck Institute for Demographic Research, Rostock. URL: `http://mortality.org` [version 31/05/2007].

A Gibbs Sampler

A.1 Sampling from $\mathcal{P}(\Psi \mid z, \beta)$

To calculate the parameters summarized in Ψ we condition on values for z and β. However, for notational convenience we will not state this explicitly throughout the section.

VAR Parameters

We derive the posterior for the VAR parameters by using the prior specified in Section 5 and by combining them with the likelihood function described in this section. To make the description of the estimation procedure more convenient we rewrite Equation (2) as

$$Z = X\Phi + \epsilon^z \,, \tag{13}$$

where $Z \equiv [z_{p+1} \; z_1 \; \dots \; z_T]'$ is a $T-p \times M$ matrix, $\Phi \equiv [\phi_1 \; \phi_2 \; \dots \; \phi_p \; c]'$ is a $Mp+1 \times M$ matrix, and

$$X \equiv \begin{bmatrix} z'_p & z'_{p-1} & \cdots & z'_1 & 1 \\ z'_{p+1} & z'_p & \cdots & z'_2 & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ z'_{T-1} & z'_{T-2} & \cdots & z'_{T-p} & 1 \end{bmatrix}$$

is a $T-p \times Mp+1$ matrix including lagged Z's. Thus its likelihood function conditional on the first p observation can be expressed as

$$\mathcal{L}(\Phi, \Sigma_z) \propto |\Sigma_z|^{-\frac{T-p}{2}} \exp\left\{ tr \left\{ -\frac{1}{2}\Sigma_z^{-1}(Z - X\Phi)'(Z - X\Phi) \right\} \right\} , \tag{14}$$

where tr is the trace operator. The likelihood function can be decomposed into

$$\mathcal{L}(\Phi, \Sigma_z) \propto |\Sigma_z|^{-\frac{T-p}{2}} \exp\left\{ tr \left\{ -\frac{1}{2}\Sigma_z^{-1} \left(\hat{S} + \frac{1}{2}(\Phi - \hat{\Phi})'X'X(\Phi - \hat{\Phi}) \right) \right\} \right\} , \tag{15}$$

where $\hat{S} \equiv (Z - X\hat{\Phi})'(Z - X\hat{\Phi})$ is the squared sample error matrix, with $\hat{\Phi} \equiv (X'X)^{-1}X'Z$. Furthermore we subdivide it into the conditional density for Φ, taking values for Σ_z^{-1} as given,

$$\mathcal{F}(\Phi|\Sigma_z) \propto |\Sigma_z|^{-\frac{M}{2}} \exp\left\{ -\frac{1}{2} \left(vec(\Phi) - vec(\hat{\Phi}) \right)' \left(\Sigma_z^{-1} \otimes X'X \right) \left(vec(\Phi) - vec(\hat{\Phi}) \right) \right\} \tag{16}$$

and the marginal density for Σ_z^{-1}

$$\mathcal{F}(\Sigma_z) \propto |\Sigma_z|^{-\frac{T-M-p}{2}} \exp\left\{ tr \left\{ -\frac{1}{2}\Sigma_z^{-1}\hat{S} \right\} \right\} . \tag{17}$$

Expression (16) is a normal density and Equation (17) a Wishart density. Thus the likelihood function can be described as a product of a normal density for Φ conditional on Σ_z and an inverted Wishart density for Σ_z,

$$\mathcal{L}(\Phi, \Sigma_z) \propto \mathcal{N}\left(vec(\hat{\Phi}), \Sigma_z \otimes X'X^{-1} \right) \mathcal{IW}\left(\hat{S}, TA - pM \right) , \tag{18}$$

where for the inverted Wishart density $\hat{S}$ serves as the scale matrix and $TA - pM$ as the degrees of freedom. Combining the likelihood function with the conjugate prior described in Section 5, we obtain the following normal posterior for Φ,

$$\Phi|\Sigma_z \sim \mathcal{N}\left(vec(\overline{\Phi}), \Sigma_z \otimes \overline{X}'\overline{X}^{-1}\right) , \tag{19}$$

where $\overline{\Phi} \equiv \overline{X}'\overline{X}^{-1}(X^{*\prime}Y^* + X'Y)$ with $\overline{X}'\overline{X} \equiv (X^{*\prime}X^* + X'X)$ and, as we assume an improper prior on Σ_z, the posterior is proportional to the second term described in Equation(18).

AR Parameters

As the error terms in equation (3) are independent of each other, we can estimate the AR parameters equation by equation. We rewrite Equation (3) as

$$\beta^i = G^i\alpha^i + \epsilon^{\beta^i} \quad \text{for} \quad i = 1, \ldots, M , \tag{20}$$

where $\beta^i \equiv [\beta^i_q \; \beta^i_{q+1} \; \ldots \; \beta^i_A]'$ is an $(A - q + 1) \times 1$ vector, $\alpha^i \equiv [\alpha^i_1 \; \alpha^i_2 \; \ldots \; \alpha^i_q]'$ is a $q \times 1$ vector, $\epsilon^{\beta^i} \equiv [\epsilon^{\beta^i}_q \; \epsilon^{\beta^i}_{q+1} \; \ldots \; \epsilon^{\beta^i}_A]'$, which is $(A - q + 1) \times 1$ vector, and

$$G^i \equiv \begin{bmatrix} \beta^i_{q-1} & \beta^i_{q-2} & \cdots & \beta^i_0 \\ \beta^i_q & \beta^i_{q-1} & \cdots & \beta^i_1 \\ \vdots & \vdots & \ddots & \vdots \\ \beta^i_{A-1} & \beta^i_{A-2} & \cdots & \beta^i_{A-q} \end{bmatrix}$$

is an $(A - q + 1) \times q$ matrix. Because we assume a flat prior for the AR parameters, the posterior of the AR parameters is proportional to the likelihood function. We can apply a similar decomposition as in Section A.1 and obtain the following normal inverted gamma posterior

$$\mathcal{P}\left(\alpha^i, \sigma^i_\beta\right) = \mathcal{F}\left(\alpha^i|\sigma^i_\beta\right)\mathcal{F}\left(\sigma^i_\beta\right) . \tag{21}$$

The posterior for α^i conditional on the variance σ^i_β is

$$\alpha^i|\sigma^i_\beta \sim \mathcal{N}\left(\hat{\alpha}^i, \sigma^i_\beta(G^{i\prime}G^i)^{-1}\right) , \tag{22}$$

where $\hat{\alpha}^i$ is the ordinary least squares (OLS) estimate and the marginal posterior for σ^i_β is the inverted gamma distribution

$$\sigma^i_\beta \sim \mathcal{IG}\left(\frac{\hat{s}}{2}, \frac{(A-q)}{2}\right) , \tag{23}$$

where $\hat{s} = (\beta^i - G^i\alpha^i)'(\beta^i - G^i\alpha^i)$ is used as the scale parameter and $A - q$ as the degrees of freedom.

Variance

We assume the variances of the disturbances in Equation (1) to be the same for the dimensions $x = 0, 1, \dots, A$ and $t = 1, 2, \dots, T$. Hence the posterior can be expressed as the inverted gamma distribution

$$\sigma_d^2 \sim \mathcal{IG}\left(\frac{TA + \tau_1}{2}, \frac{\hat{s}_d + \tau_2}{2}\right), \tag{24}$$

where $\hat{s}_d = \sum_{t=1}^{T} \sum_{x=0}^{A} \left(d_{x,t} - \overline{d}_x - \beta_x z_t\right)^2$.

A.2 Sampling from $\mathcal{P}(z \mid \Psi, \beta)$

To calculate the latent z we condition on values for Ψ and β. However, for notational convenience we will not state this explicitly throughout the section. As z contains latent variables, we set up a state space system, which we will describe in the following.

We rewrite Equation (2) into its canonical form and use it as our state equation

$$Z_t = \widetilde{\Phi} Z_{t-1} + \widetilde{\epsilon}_t^z, \tag{25}$$

where $Z_t \equiv [z_t \; z_{t-1} \; \dots \; z_{t-p+1} 1]'$ is $(Mp+1) \times 1$, which is the state vector, $\widetilde{\epsilon}_t^z \equiv [\epsilon_t^z \; 0 \; \dots \; 0]'$, which is a an $(Mp + 1) \times 1$ vector, and

$$\widetilde{\Phi} \equiv \begin{bmatrix} \phi_1 & \dots & \phi_p & c \\ & I_{M(p-1)\times M(p-1)} & & 0_{M(p-1)\times(M+1)} \\ 0 & \dots & 0 & 1 \end{bmatrix},$$

which is an $(Mp + 1) \times (Mp + 1)$ matrix, where I is the identity matrix.

To derive our observation equation we first rewrite Equation (1) as

$$\overline{D}_t = \beta z_t + \epsilon_t^d, \tag{26}$$

with

$$\overline{D}_t \equiv \begin{bmatrix} D_t & -\overline{D} \\ & Y_t \end{bmatrix},$$

which is an $(A + N) \times 1$ matrix, with $D_t \equiv [d_{0,t} \; d_{1,t} \; \dots d_{A,t}]'$, $\overline{D} \equiv [\overline{d}_0 \; \overline{d}_1 \; \dots \overline{d}_A]'$, where both are $A \times 1$ vectors, $\epsilon_t^d = [\epsilon_{0,t}^d \; \epsilon_{1,t}^d \; \dots \; \epsilon_{A,t}^d \; 0_{1\times N}]'$ is a $(A + N) \times 1$, and

$$\beta \equiv \begin{bmatrix} \beta^\kappa & \beta^Y \\ 0_{N\times K} & I_{N\times N} \end{bmatrix},$$

which is an $(A + N) \times M$ matrix, with $\beta^\kappa \equiv [(\beta_0^\kappa)' \; (\beta_1^\kappa)' \dots \; (\beta_A^\kappa)']'$, which is an $A \times K$ matrix, and $\beta^Y \equiv \left[(\beta_0^Y)' \; (\beta_1^Y)' \dots \; (\beta_A^Y)'\right]'$, which is an $A \times N$ matrix.

We rewrite Equation (26) to match the state equation and finally obtain our observation equation

$$\overline{D}_t = HZ_t + \epsilon_t^d \,, \tag{27}$$

where $H \equiv [\beta \; 0_{A+N\times M(p-1)+1}]$ is an $(A+N)\times(Mp+1)$ matrix.

To calculate z we apply the algorithm suggested by Carter and Kohn (1994) and Frühwirth-Schnatter (1994).[19] With this procedure we draw z from its joint distribution

$$\mathcal{P}(z|D) = \mathcal{P}\left(z_T|\overline{D}_T\right)\prod_{t=1}^{T-1}\mathcal{P}\left(z_t|z_{t+1}, D^t\right)\,, \tag{28}$$

where $D = [\overline{D}_1 \; \overline{D}_2 \; \ldots \; \overline{D}_T]$ and $D^t = [\overline{D}_1 \; \overline{D}_2 \; \ldots \; \overline{D}_t]$. Because the disturbances in Equations (25) and (27) are Gaussian, Equation (28) can be rewritten as

$$\mathcal{P}(z|D) = \mathcal{N}(z_{T|T}, P_{T|T})\prod_{t=1}^{T-1}\mathcal{N}(z_{t|t,z_{t+1}}, P_{t|t,z_{t+1}})\,, \tag{29}$$

with

$$z_{T|T} = E(z_T|D)\,, \tag{30}$$
$$P_{T|T} = Cov(z_T|D)\,, \tag{31}$$

and

$$z_{t|t,z_{t+1}} = E(z_t|z_{t+1}, D)\,, \tag{32}$$
$$P_{t|t,z_{t+1}} = Cov(z_t|z_{t+1}, D)\,. \tag{33}$$

We obtain $z_{T|T}$ and $P_{T|T}$ from the last step of the Kalman filter iteration and use them as the conditional mean and covariance matrix for the multivariate normal distribution $\mathcal{N}(z_{T|T}, P_{T|T})$ in order to draw z_T. In the following we will describe the Kalman filter procedure.

We begin with the prediction steps

$$z_{t|t-1} = \widetilde{\Phi} z_{t-1|t-1}\,, \tag{34}$$
$$P_{t|t-1} = \widetilde{\Phi} P_{t-1|t-1}\widetilde{\Phi} + Q\,, \tag{35}$$

where

$$Q \equiv \begin{bmatrix} \Sigma_z & 0_{M\times M(p-1)+1} \\ 0_{M(p-1)+1\times M(p-1)+1} & 0_{M(p-1)+1\times M} \end{bmatrix}\,,$$

which is an $(Mp+1)\times(Mp+1)$ matrix. Accordingly, the forecast error is

$$\nu_t = \overline{D}_t - Hz_{t|t-1}\,, \tag{36}$$

[19] Cf. also Kim and Nelson (1999).

with the corresponding variance

$$\Omega = HP_{t|t-1}H' + R\,, \tag{37}$$

where $R \equiv \sigma_d^2 I_N$. The Kalman gain can be expressed as

$$K_t = P_{t|t-1}H'\Omega^{-1}\,. \tag{38}$$

Thus the updating equations are

$$\begin{aligned} z_{t|t} &= z_{t|t-1} + K_t\nu_t\,, && (39)\\ P_{t|t} &= P_{t|t-1} + K_t H P_{t|t-1}\,. && (40) \end{aligned}$$

To obtain draws for $z_1, z_2, \ldots, z_{T-1}$ we sample from $N\left(z_{t|t,z_{t+1}}, P_{t|t,z_{t+1}}\right)$, using a backward-moving updating scheme, incorporating at time t information about z_t contained in period $t+1$. More precisely, we move backward and generate z_t for $t = T-1, \ldots, 1$ at each step while using information from the Kalman filter and z_{t+1} from the previous step. The updating equations are

$$z_{t|t,z_{t+1}} = z_{t|t} + P_{t|t}\Phi' P_{t+1|t}^{-1}(z_{t+1} - z_{t+1|t}) \tag{41}$$

and

$$P_{t|t,F_{t+1}} = P_{t|t} - P_{t|t}\Phi' P_{t+1|t}^{-1}\Phi P_{t|t}\,. \tag{42}$$

A.3 Sampling from $\mathcal{P}(\beta \mid \Psi, z)$

To calculate β we take values for Ψ and z as given. The procedure applied here is very similar to the one described in Section A.2. Hence we will just give a brief overview of the estimation procedure. However, there is one important difference, namely, that now we move in the age dimension $x = 0, 1, \ldots, A$ and not in $t = 1, 2, \ldots, T$ as in Section A.2.
Our state equation can be expressed as

$$\widetilde{\beta}_x = \widetilde{\alpha}\widetilde{\beta}_{x-1} + \widetilde{\epsilon}_x^{\beta}\,, \tag{43}$$

where $\widetilde{\beta}_x = [\beta_{x-1}\ \beta_{x-2}\ \ldots\ \beta_{x-q+1}]'$ is $Mq \times 1$, which is denoted as the state vector, $\widetilde{\epsilon}_x^{\beta} = [\epsilon_x^{\beta}\ 0\ \ldots\ 0]'$ is $Mq \times 1$, and

$$\widetilde{\alpha} = \left[\begin{array}{ccc} \alpha_1 & \ldots & \alpha_q \\ I_{M(p-1)\times M(p-1)} & & 0_{M(p-1)\times(M+1)} \end{array}\right],$$

which is an $Mq \times Mq$ matrix. Hence our observation equation can be expressed as

$$\widetilde{D}_x - \overline{d}_x = W\widetilde{\beta}_x + \epsilon_x^d\,, \tag{44}$$

where $\widetilde{D}_x \equiv [d_{x,1}\ d_{x,2}\ \ldots, d_{x,T}]'$ is a $T \times 1$ vector, $\epsilon_x^d \equiv [\epsilon_{x,1}^d\ \epsilon_{x,1}^d \ldots\ \epsilon_{x,1}^d]$ is a $T \times 1$ vector, and $W \equiv [z'\ 0_{T,M(q-1)}]$ is a $T \times Mq$ matrix. For $x = 0, 1, \ldots, A$ instead of $t = 1, 2, \ldots T$, $\widetilde{\Phi} \equiv \widetilde{\alpha}$, $H \equiv W$, $R \equiv \sigma_d^2 I_T$, and

$$Q \equiv \left[\begin{array}{cc} \Sigma_\beta & 0_{M\times M(q-1)} \\ 0_{M(q-1)\times M(q-1)} & 0_{M(q-1)\times M} \end{array}\right],$$

we can apply the procedure described in Section A.2 to calculate β.

B Life Table Calculations

We use both observed and estimated age-specific death rates $m_{x,t}$ to calculate period life tables by single years of age and time and present the results for the probability $l_{x,t}$ of surviving up to the exact age x and the probability $d_{x,t}$ of dying at age x. Both variables represent birth time probabilities for all born living. Thus they are unconditional. In contrast to this, the remaining life expectancy $e_{x,t}$ is conditional on still being alive at exact age x. The respective calculations are standard.[20]

The conditional probability of dying before arriving at exact age $x + 1$ if still alive at exact age x is

$$q_{x,t} \equiv \frac{m_{x,t}}{1 + (1 - \alpha_{x,t})m_{x,t}} \ .$$

The factor $\alpha_{x,t}$ reflects the average fraction of a year that people dying at age x still live after their xth birthday. For infants, with their high mortality in the first weeks, we apply, according to Preston et al. (2005, pp. 47–48) and Wilmoth et al. (2007, p. 38), sex-specific values originally proposed by Coale and Demeny (1983):

$$\alpha_{0,t}^{male} \equiv \begin{cases} 0.045 + 2.684 m_{0,t}^{male} & , m_{0,t}^{male} < 0.107 \\ 0.330 & , m_{0,t}^{male} \geq 0.107 \end{cases}$$

and

$$\alpha_{0,t}^{female} \equiv \begin{cases} 0.053 + 2.800 m_{0,t}^{female} & , m_{0,t}^{female} < 0.107 \\ 0.350 & , m_{0,t}^{female} \geq 0.107 \end{cases}$$

Consistent values for $\alpha_{0,t}^{total}$ would require information about the total numbers of deaths for both sexes to weight the respective values for $m_{0,t}^{male}$ and $m_{0,t}^{female}$. Instead of that, when using the total figures of both sexes combined, we adopt a simple approximation roughly reflecting the higher infant mortality and higher birth rates of males

$$\alpha_{0,t}^{total} \equiv 0.56\alpha_{0,t}^{male} + 0.44\alpha_{0,t}^{female} \ ,$$

which does not perceivably influence the results. The highest recorded age class $\tilde{x}$ is open, that is, not restricted to 1 year. We set $\alpha_{\tilde{x},t} \equiv \frac{1}{m_{\tilde{x},t}}$ resulting in $q_{\tilde{x},t} = 1$. For all other age classes $0 < x < \tilde{x}$ we assume a uniform distribution of cases of death and apply

$$\alpha_{x,t} \equiv 0.5 \ .$$

The conditional probability of surviving up to exact age $x+1$ if still alive at exact age x is

$$p_{x,t} \equiv 1 - q_{x,t} \ .$$

[20]Cf. Preston et al. (2005, pp. 38–54) or Wilmoth et al. (2007, pp. 35–39). Unlike the life table calculations of the Human Mortality Database, we do not smooth observed death rates $m_{x,t}$ for the higher age classes at the beginning of the calculations.

For all born living the unconditional probability of surviving up to exact age x is

$$l_{x,t} \equiv l_{0,t} \prod_{i=0}^{x-1} p_{i,t} = l_{x-1,t} p_{x-1,t}$$

and the unconditional probability of dying at age x is

$$d_{x,t} \equiv l_{0,t} \prod_{i=0}^{x-1} p_{i,t} q_{x,t} = l_{x,t} q_{x,t} \ .$$

We normalize $l_{0,t} \equiv 1$ to get values for $l_{x,t}$ and $d_{x,t}$ interpretable as probabilities for the life table population. The alternative choice of $l_{0,t} \equiv 100000$ would result in the numbers $l_{x,t}$ and $d_{x,t}$ of survivors and deaths out of $100,000$ live births.

The person-years lived at age x and from age x onward are

$$L_{x,t} \equiv l_{x,t} - (1 - \alpha_{x,t}) d_{x,t}$$

and

$$T_{x,t} \equiv \sum_{i=x}^{\tilde{x}} L_{i,t} \ .$$

Finally, we get the conditional remaining life expectancy if still alive at exact age x

$$e_{x,t} \equiv \frac{T_{x,t}}{l_{x,t}} \ .$$

Note that all variables in a period life table refer to the same point in time t and reflect its time-specific conditions. Variables such as $l_{x,t}$, $d_{x,t}$, and $e_{x,t}$ that are aggregated from the basic variables of several age classes are synthetic measures for this period. They mix up the values of the different age classes belonging to different cohorts because they correspond to a cross section of the Lexis diagram. Hence the aggregated variables of a period life table do not describe the conditions for the members of any real age cohort, who pass through many different periods but are always subject to the mortality of their very own cohort. To analyze these conditions along the life cycle, cohort life tables, which are calculated from data of a single cohort, are adequate and correspond to diagonal sections of the Lexis diagram. Unfortunately, they can only be accurately calculated retrospectively. Of course, short-run fluctuations that last only a few periods but affect many age classes have a greater effect on period life tables than on cohort life tables. The latter exhibit, in general, less volatility, because time-specific anomalies are not wrongly extrapolated but on the contrary often counterbalanced later on.

SFB 649 Discussion Paper Series 2008

For a complete list of Discussion Papers published by the SFB 649, please visit http://sfb649.wiwi.hu-berlin.de.

001 "Testing Monotonicity of Pricing Kernels" by Yuri Golubev, Wolfgang Härdle and Roman Timonfeev, January 2008.
002 "Adaptive pointwise estimation in time-inhomogeneous time-series models" by Pavel Cizek, Wolfgang Härdle and Vladimir Spokoiny, January 2008.
003 "The Bayesian Additive Classification Tree Applied to Credit Risk Modelling" by Junni L. Zhang and Wolfgang Härdle, January 2008.
004 "Independent Component Analysis Via Copula Techniques" by Ray-Bing Chen, Meihui Guo, Wolfgang Härdle and Shih-Feng Huang, January 2008.
005 "The Default Risk of Firms Examined with Smooth Support Vector Machines" by Wolfgang Härdle, Yuh-Jye Lee, Dorothea Schäfer and Yi-Ren Yeh, January 2008.
006 "Value-at-Risk and Expected Shortfall when there is long range dependence" by Wolfgang Härdle and Julius Mungo, Januray 2008.
007 "A Consistent Nonparametric Test for Causality in Quantile" by Kiho Jeong and Wolfgang Härdle, January 2008.
008 "Do Legal Standards Affect Ethical Concerns of Consumers?" by Dirk Engelmann and Dorothea Kübler, January 2008.
009 "Recursive Portfolio Selection with Decision Trees" by Anton Andriyashin, Wolfgang Härdle and Roman Timofeev, January 2008.
010 "Do Public Banks have a Competitive Advantage?" by Astrid Matthey, January 2008.
011 "Don't aim too high: the potential costs of high aspirations" by Astrid Matthey and Nadja Dwenger, January 2008.
012 "Visualizing exploratory factor analysis models" by Sigbert Klinke and Cornelia Wagner, January 2008.
013 "House Prices and Replacement Cost: A Micro-Level Analysis" by Rainer Schulz and Axel Werwatz, January 2008.
014 "Support Vector Regression Based GARCH Model with Application to Forecasting Volatility of Financial Returns" by Shiyi Chen, Kiho Jeong and Wolfgang Härdle, January 2008.
015 "Structural Constant Conditional Correlation" by Enzo Weber, January 2008.
016 "Estimating Investment Equations in Imperfect Capital Markets" by Silke Hüttel, Oliver Mußhoff, Martin Odening and Nataliya Zinych, January 2008.
017 "Adaptive Forecasting of the EURIBOR Swap Term Structure" by Oliver Blaskowitz and Helmut Herwatz, January 2008.
018 "Solving, Estimating and Selecting Nonlinear Dynamic Models without the Curse of Dimensionality" by Viktor Winschel and Markus Krätzig, February 2008.
019 "The Accuracy of Long-term Real Estate Valuations" by Rainer Schulz, Markus Staiber, Martin Wersing and Axel Werwatz, February 2008.
020 "The Impact of International Outsourcing on Labour Market Dynamics in Germany" by Ronald Bachmann and Sebastian Braun, February 2008.
021 "Preferences for Collective versus Individualised Wage Setting" by Tito Boeri and Michael C. Burda, February 2008.

This research was supported by the Deutsche Forschungsgemeinschaft through the SFB 649 "Economic Risk".

022 "Lumpy Labor Adjustment as a Propagation Mechanism of Business Cycles" by Fang Yao, February 2008.
023 "Family Management, Family Ownership and Downsizing: Evidence from S&P 500 Firms" by Jörn Hendrich Block, February 2008.
024 "Skill Specific Unemployment with Imperfect Substitution of Skills" by Runli Xie, March 2008.
025 "Price Adjustment to News with Uncertain Precision" by Nikolaus Hautsch, Dieter Hess and Christoph Müller, March 2008.
026 "Information and Beliefs in a Repeated Normal-form Game" by Dietmar Fehr, Dorothea Kübler and David Danz, March 2008.
027 "The Stochastic Fluctuation of the Quantile Regression Curve" by Wolfgang Härdle and Song Song, March 2008.
028 "Are stewardship and valuation usefulness compatible or alternative objectives of financial accounting?" by Joachim Gassen, March 2008.
029 "Genetic Codes of Mergers, Post Merger Technology Evolution and Why Mergers Fail" by Alexander Cuntz, April 2008.
030 "Using R, LaTeX and Wiki for an Arabic e-learning platform" by Taleb Ahmad, Wolfgang Härdle, Sigbert Klinke and Shafeeqah Al Awadhi, April 2008.
031 "Beyond the business cycle – factors driving aggregate mortality rates" by Katja Hanewald, April 2008.
032 "Against All Odds? National Sentiment and Wagering on European Football" by Sebastian Braun and Michael Kvasnicka, April 2008.
033 "Are CEOs in Family Firms Paid Like Bureaucrats? Evidence from Bayesian and Frequentist Analyses" by Jörn Hendrich Block, April 2008.
034 "JBendge: An Object-Oriented System for Solving, Estimating and Selecting Nonlinear Dynamic Models" by Viktor Winschel and Markus Krätzig, April 2008.
035 "Stock Picking via Nonsymmetrically Pruned Binary Decision Trees" by Anton Andriyashin, May 2008.
036 "Expected Inflation, Expected Stock Returns, and Money Illusion: What can we learn from Survey Expectations?" by Maik Schmeling and Andreas Schrimpf, May 2008.
037 "The Impact of Individual Investment Behavior for Retirement Welfare: Evidence from the United States and Germany" by Thomas Post, Helmut Gründl, Joan T. Schmit and Anja Zimmer, May 2008.
038 "Dynamic Semiparametric Factor Models in Risk Neutral Density Estimation" by Enzo Giacomini, Wolfgang Härdle and Volker Krätschmer, May 2008.
039 "Can Education Save Europe From High Unemployment?" by Nicole Walter and Runli Xie, June 2008.
040 "Solow Residuals without Capital Stocks" by Michael C. Burda and Battista Severgnini, August 2008.
041 "Unionization, Stochastic Dominance, and Compression of the Wage Distribution: Evidence from Germany" by Michael C. Burda, Bernd Fitzenberger, Alexander Lembcke and Thorsten Vogel, March 2008
042 "Gruppenvergleiche bei hypothetischen Konstrukten – Die Prüfung der Übereinstimmung von Messmodellen mit der Strukturgleichungs-methodik" by Dirk Temme and Lutz Hildebrandt, June 2008.
043 "Modeling Dependencies in Finance using Copulae" by Wolfgang Härdle, Ostap Okhrin and Yarema Okhrin, June 2008.
044 "Numerics of Implied Binomial Trees" by Wolfgang Härdle and Alena Mysickova, June 2008.

SFB 649, Spandauer Straße 1, D-10178 Berlin
http://sfb649.wiwi.hu-berlin.de

This research was supported by the Deutsche Forschungsgemeinschaft through the SFB 649 "Economic Risk".

045 "Measuring and Modeling Risk Using High-Frequency Data" by Wolfgang Härdle, Nikolaus Hautsch and Uta Pigorsch, June 2008.

046 "Links between sustainability-related innovation and sustainability management" by Marcus Wagner, June 2008.

047 "Modelling High-Frequency Volatility and Liquidity Using Multiplicative Error Models" by Nikolaus Hautsch and Vahidin Jeleskovic, July 2008.

048 "Macro Wine in Financial Skins: The Oil-FX Interdependence" by Enzo Weber, July 2008.

049 "Simultaneous Stochastic Volatility Transmission Across American Equity Markets" by Enzo Weber, July 2008.

050 "A semiparametric factor model for electricity forward curve dynamics" by Szymon Borak and Rafał Weron, July 2008.

051 "Recurrent Support Vector Regreson for a Nonlinear ARMA Model with Applications to Forecasting Financial Returns" by Shiyi Chen, Kiho Jeong and Wolfgang K. Härdle, July 2008.

052 "Bayesian Demographic Modeling and Forecasting: An Application to U.S. Mortality" by Wolfgang Reichmuth and Samad Sarferaz, July 2008.

SFB 649, Spandauer Straße 1, D-10178 Berlin
http://sfb649.wiwi.hu-berlin.de

This research was supported by the Deutsche Forschungsgemeinschaft through the SFB 649 "Economic Risk".