

SFB 649 Discussion Paper 2008-052

# Bayesian Demographic Modeling and Forecasting: An Application to U.S. Mortality

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

# Bayesian Demographic Modeling and Forecasting: An Application to U.S. Mortality<sup>1</sup>

Wolfgang H. Reichmuth

Humboldt-Universität zu Berlin

Samad Sarferaz

Humboldt-Universität zu Berlin

*Preliminary Draft*

## Abstract

We present a new way to model age-specific demographic variables with the example of age-specific mortality in the U.S., 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 hitherto 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 to 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

---

<sup>1</sup>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 as they address the most fundamental attributes of human life. Respective 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 brings forth 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 high importance in areas like social security and public health. In the private sector such advancements of knowledge can have a substantial monetary value as they improve the calculation of life insurances or pension schemes for the insurance industry. Population forecasts that can be derived from demographic rates give another example of interest beyond pure science due to their implications for investment decisions in the public and private sector. 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 with 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 with 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 along 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 Methods (MCMC), to estimate the model with an efficient one-step procedure. By the choice of priors this Bayesian estimation approach also reveals clearly the assumptions made. Most notably, it also not only yields point estimates, but distributional statements for the results in the most natural 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 GDP and unemployment as important macroeconomic variables. Due 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 popular Lee-Carter

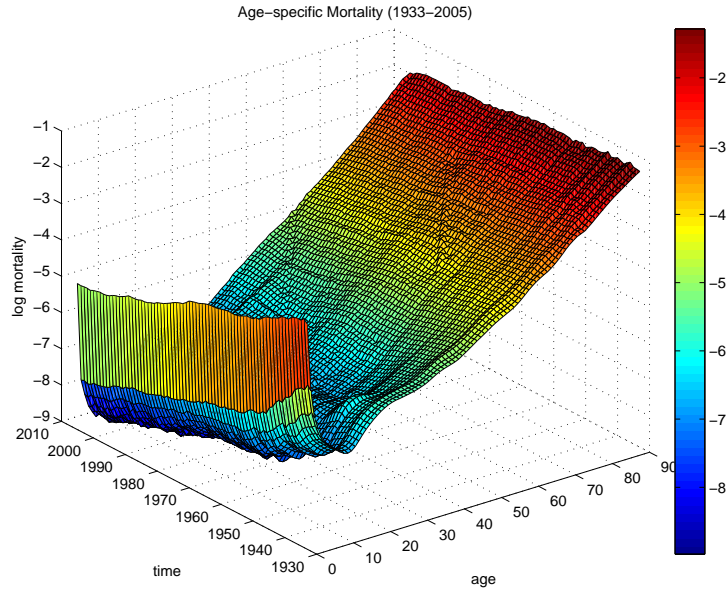


Figure 1: Mortality surface of logarithmized age-specific total (female and male combined) mortality in the U.S. 1933–2005.

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 like in the Lee-Carter model. In-sample forecasts yield 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 to 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.<sup>2</sup>

The presented approach can be applied to model, forecast and analyze all kinds of age-specific variables. Mortality just forms a prominent example due to its high 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 its value on its own, forecasts of 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.

<sup>2</sup>Cf. REICHMUTH AND SARFERAZ (2008).

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, before 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 concludes.

## 2 Literature on Modeling and Forecasting Mortality

We start with a short overview of some developments in modeling and forecasting age-specific mortality.<sup>3 4</sup> Models that map age to age-specific mortality take advantage of the obvious, strong regularities in the age pattern of mortality. 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 data to be handled.

### 2.1 Parametric Modeling of Age-specific Mortality

Systematic patterns in mortality have been known since the development of 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 this time, HELIGMAN AND POLLARD (1980) suggest a sum of three terms representing different components of mortality

$$m(x) = A^{(x+B)^C} + D e^{-E(\ln x - \ln F)^2} + GH^x / (1 + GH^x)$$

with eight time-dependent parameters  $A_t, \dots, H_t$ . 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 by univariate time series methods as ARIMA processes which may lead to inconsistencies in the long run.

---

<sup>3</sup>Of course, we can only briefly sketch some major issues. BOOTH (2006) gives a comprehensive survey on demographic forecasting.

<sup>4</sup>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, i.e. 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) have developed a two-dimensional set of four regional patterns each with 24 different *mortality levels* identified by 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

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

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.  $b_x$  is the first principal component and is estimated by singular value decomposition. The subsequent estimation of the mortality index  $k_t$  as ARIMA process results in a simple random walk with drift. However, the outcome of forecasting age-specific mortality by this method with one time-dependent parameter is similar to if each age-specific time series was extrapolated along its own historic time trend, potentially leading to an implausible age pattern in the long run.<sup>5</sup> This disadvantage is especially severe if the Lee-Carter approach is applied to single-cause mortality, for which it was not indeed assigned.<sup>6</sup> Nevertheless, the Lee-Carter method and its several enhancements have become the standard for mortality forecasts and have been used for the newly emerged stochastic population forecasts since LEE AND TULJAPURKAR (1994) and LEE (1998), too.

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 problems and extensions of it. Quantitative comparisons of several recent models are given by CAIRNS ET AL. (2007) and CAIRNS ET AL. (2008). But they only apply data for the age classes 60–89, i.e., 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) smooth along the age dimension by restricting the impact of several  $k_t$  on particular age classes with a matrix containing splines.<sup>7</sup> Smoothing with a roughness penalty in the estimation of both the Lee-Carter and a Poisson log-bilinear model is done by DELWARDE ET AL. (2007).

---

<sup>5</sup>This critique goes back to MCNOWN (1992) and ALHO (1992).

<sup>6</sup>GIROSI AND KING (2008, pp. 38–42) discuss this point and give examples.

<sup>7</sup>In a different approach of a generalized linear model with Poisson errors, CURRIE ET AL. (2004) smooth 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 expert's 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 like 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 vector autoregression (VAR). We assume an autoregression (AR) process for the coefficients, which link the explanatory variables with the age-specific demographic variables, to ensure a smooth development along the age dimension. For the estimation of this state space model we use Bayesian methods, providing an appropriate Markov Chain Monte Carlo (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 following equation

$$d_{x,t} = \bar{d}_x + \beta_x z_t + \epsilon_{x,t}^d \quad (1)$$

with the arithmetic mean  $\bar{d}_x = \frac{1}{T} \sum_{t=1}^T d_{x,t}$  and explanatory variables  $z_t \equiv [\kappa_t \ Y_t]'$ , where  $\kappa_t$  is a  $K \times 1$  vector of unobservables and  $Y_t$  is a  $N \times 1$  vector of observed covariates. The corresponding coefficient vector  $\beta_x \equiv [\beta_x^\kappa \ \beta_x^Y]$  is  $1 \times M$ , where  $\beta_x^\kappa$  is a  $1 \times K$  vector and  $\beta_x^Y$  is a  $1 \times N$  vector with  $M = K + N$ . We assume for  $z_t$  and  $\beta_x$  to follow vector autoregressive processes

$$z_t = c + \phi_1 z_{t-1} + \phi_2 z_{t-2} + \dots + \phi_p z_{t-p} + \epsilon_t^z, \quad (2)$$

$$\beta_x = \alpha_1 \beta_{x-1} + \alpha_2 \beta_{x-2} + \dots + \alpha_q \beta_{x-q} + \epsilon_x^\beta, \quad (3)$$

where  $c$  is a  $M \times 1$  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  $\epsilon_{x,t}^d \sim i.i.d. \mathcal{N}(0, \sigma_d^2)$  for the disturbances in Equation (1),  $\epsilon_t^z \sim i.i.d. \mathcal{N}(0, \Sigma_z)$  for the disturbances in Equation (2) and  $\epsilon_x^\beta \sim i.i.d. \mathcal{N}(0, \Sigma_\beta)$  for the disturbances in Equation (3), where the covariance matrix  $\Sigma_\beta$  is a diagonal matrix. Thus each component of  $\beta_x$  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  $\phi_1, \phi_2, \dots, \phi_q$ , where we specify the prior on  $\phi_1$  very tightly around one and the prior on  $\phi_2, \dots, \phi_q$  very tightly around zero. 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  $\phi_1, \phi_2, \dots, \phi_q$  and we obtain, approximately, the following model

$$d_{x,t} = \bar{d}_x + \beta_x^\kappa \kappa_t + \epsilon_{x,t}^d \quad (4)$$

with an AR-process for the mortality index  $\kappa_t$

$$\kappa_t = c + \kappa_{t-1} + \epsilon_t^\kappa, \quad (5)$$

which is the Lee-Carter Model set up in state space representation as it is described in PEDROZA (2006).

### 3.3 Augmenting the Simple Model with Covariates

The inclusion of covariates may noticeably improve the forecasts of demographic models.<sup>8</sup> 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  $z_t = [\kappa_t \ Y_t]'$  instead of the simpler special case  $z_t = \kappa_t$  according to Lee-Carter. The informational gain of this inclusion depends of course on the specifications of the demographic variable and of 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 besides 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 in 50 years from now. Nevertheless, we are very confident that this mortality is quite similar to the mortality 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 a smooth development along the age dimension, because the coefficients  $\beta_x, \dots, \beta_{x-q}$  are connected by autoregressive processes for each component

---

<sup>8</sup>This issue is discussed extensively in GIROSI AND KING (2008).



of the  $\beta$ 's. 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<sup>9</sup>

$$\beta_{\tilde{x}} = \tilde{\alpha}_{-\frac{q}{2}} \beta_{\tilde{x}-\frac{q}{2}} + \cdots + \tilde{\alpha}_{-1} \beta_{\tilde{x}-1} + \tilde{\alpha}_1 \beta_{\tilde{x}+1} + \cdots + \tilde{\alpha}_{\frac{q}{2}} \beta_{\tilde{x}+\frac{q}{2}} + \tilde{\epsilon}_{\tilde{x}}^{\beta}. \quad (6)$$

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 U.S., e.g., in the case of single-cause mortality or of data from non-industrialized countries in the past and present.

### 3.5 Cohort Effects

The general model described above 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

$$d_{x,t} = \bar{d}_x + \beta_x z_t + \beta_x^{\gamma} \gamma_{t-x} + \epsilon_{x,t}^d. \quad (7)$$

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

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

where  $\epsilon_t^{\gamma}$  is not serially correlated and independent of  $\epsilon_{x,t}^d$ ,  $\epsilon_t^z$  and  $\epsilon_x^{\beta}$  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

$$d_{x,t} = \bar{d}_x + (\beta_x^{\kappa} \gamma) \left( \frac{\kappa_t}{\gamma} \right) + \beta_x^Y Y_t + \epsilon_{x,t}^d. \quad (9)$$

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

$$d_{x,t} = \bar{d}_x + (\beta_x P') (P z_t) + \epsilon_{x,t}^d,$$

---

<sup>9</sup>Set  $\alpha_0 \equiv -1$ ,  $\tilde{\alpha}_i \equiv -\frac{\alpha_{(q/2)-i}}{\alpha_{q/2}}$  for  $i \in \{-\frac{q}{2}, \dots, \frac{q}{2}\}$ ,  $\tilde{x} \equiv x - \frac{q}{2}$  and  $\tilde{\epsilon}_{\tilde{x}}^{\beta} \equiv -\frac{\epsilon_x^{\beta}}{\alpha_{q/2}}$ .

where

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

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

## 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

$$\begin{aligned} d_x^H &\equiv [d_{x,T+1} \dots d_{x,T+H}], \\ d_x^T &\equiv [d_{x,1} \dots d_{x,T}], \\ z &\equiv [z_1 \ z_2 \ \dots \ z_T], \\ \beta &\equiv [\beta_0 \ \beta_1 \ \dots \ \beta_A]', \\ \Psi &\equiv \{(c, \phi_1, \phi_2, \dots, \phi_p, \Sigma_z), (\alpha_1, \alpha_2, \dots, \alpha_q, \Sigma_\beta), (\sigma_d^2)\}. \end{aligned}$$

Thus, the posterior predictive density can be expressed as

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

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

$$z_{T+i} = c + \phi_1 z_{T+i-1} + \phi_2 z_{T+i-2} + \dots + \phi_p z_{T+i-p} + \epsilon_{T+i}^z. \quad (10)$$

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

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

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

---

<sup>10</sup>This is similar to the dynamic factor literature. See, amongst others, GEWEKE AND ZHOU (1996) and BERNANKE ET AL. (2005).

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

We consider the following model

$$Z^* = X^* \Phi^* + \epsilon^*, \quad (11)$$

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  $\lambda_1$  is called the *overall tightness of beliefs* around the random walk prior and  $\hat{\sigma}_1, \hat{\sigma}_2, \dots, \hat{\sigma}_M$  are the empirical standard deviations taken from the first  $p$  observations. Increasing values for  $\lambda_1$  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  $\Sigma_z$  as given the dummy observations imply the following conjugate prior for our VAR-parameters

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

The prior for the AR-parameters in Equation (3) is similar to the one specified for the VAR parameters with  $\lambda_2$  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 Markov Chain Monte Carlo 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 following conditional distributions  $\mathcal{P}(\Psi | z, \beta)$ ,  $\mathcal{P}(z | \Psi, \beta)$  and  $\mathcal{P}(\beta | \Psi, z)$  and draw iteratively from them. Taken initialized values for  $z^{(0)}$  and  $\beta^{(0)}$  as given, we sample in the  $i$ -th iteration  $\Psi^{(i)}$  from  $\mathcal{P}(\Psi | z^{(i-1)}, \beta^{(i-1)})$ ,  $z^{(i)}$  from  $\mathcal{P}(z | \Psi^{(i)}, \beta^{(i-1)})$  and  $\beta^{(i)}$  from  $\mathcal{P}(\beta | \Psi^{(i)}, z^{(i)})$  successively. Under weak conditions and for  $i \rightarrow \infty$  the Gibbs sampler converges and we obtain samples from the desired joint distribution  $\mathcal{P}(\Psi, z, \beta)$ .<sup>11</sup> For a more detailed description of the estimation procedure we refer to Appendix A.

## 7 Data

We apply our model to age-specific total (combining female and male) mortality data from the U.S. with 91 individual age classes from 0 to 90 as shown in Figure 1 as specification of the demographic variable  $d_{x,t}$ .<sup>12</sup> These time series provided by the Human Mortality Database span the period 1933–2005 of which we use the post-WW II period.<sup>13</sup> We add macroeconomic time series of real gross domestic product (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, the unemployment rate is measured as number of unemployed in percentage of the civilian labor force.<sup>14</sup>

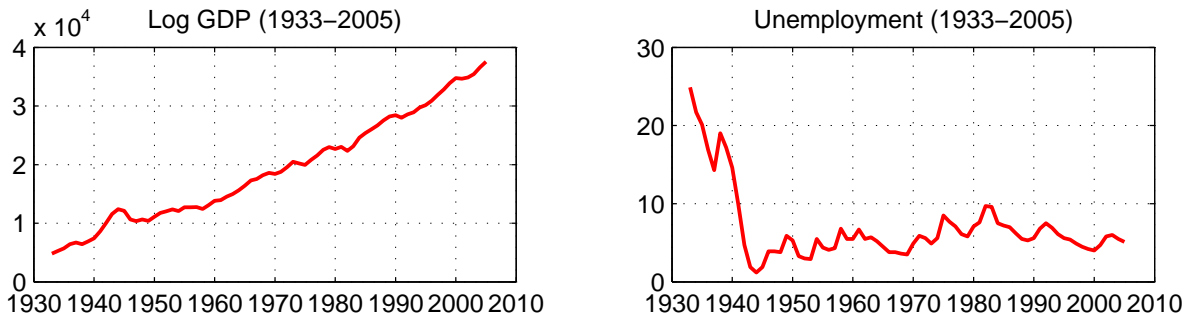


Figure 2: Logarithmized GDP and unemployment rate for the U.S. 1946–2005.

<sup>11</sup>Cf. GEMAN AND GEMAN (1984).

<sup>12</sup>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.

<sup>13</sup>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 U.S., population estimates for 1940–1969 are adjusted to exclude the Armed Forces overseas and to correct the inclusion of Alaska and Hawaii. Moreover, due 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).

<sup>14</sup>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 c.f. U.S. CENSUS BUREAU (2007).

## 8 Results

We apply our model to mortality data from the U.S. in the period 1946–2005 and gradually vary the model specification. With the objective of comparability with the results of Lee-Carter, we first assume  $\kappa_t$  to consist of only one unobserved time series, which may be called mortality index, and abstain from using covariates. Afterwards, the macroeconomic time series are included as covariates.

### 8.1 Preliminaries

For the results we used a lag length of  $p = 4$  for the  $z$ 's and  $q = 4$  for the  $\beta$ 's. The prior specifications, which we describe in Section 5, are  $\lambda_1 = 5$  for the VAR parameters of  $z$ <sup>15</sup> and a flat prior  $\lambda_2 = 0$  for the AR parameters of  $\beta$ . 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 the time period and of the age span under consideration. To check whether our results depend on the initial  $\beta$  parameters we conduct the following exercise. We leave out mortality of the youngest age classes and estimate our model with  $\beta_s, \dots, \beta_A$ , where  $s > 0$ . We obtain very similar results to the full model  $\beta_s, \dots, \beta_A$ , suggesting that the choice of initial values for the  $\beta$ '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 the forecasts on circumstances relatively close to 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 avoid that the starting values influence our results we discard the first half of the chain as 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  $\kappa$  and no covariates. Figure 3 shows the estimated  $\kappa$  and the corresponding coefficient matrix  $\beta$  which reveals how close the mortality of particular age classes is associated with the developing of the latent variable  $\kappa$ . The age classes 0–15 are higher-than-average exposed to  $\kappa$ . However, all age classes are positively related to the latent variable.

In Figure 4 we show different in-sample forecasts for  $\kappa$  over a fifteen year horizon from 1991 onwards, that can be compared with the 'realized' developing (red line), which means the median of the estimated  $\kappa$  for the entire period.

---

<sup>15</sup>Which is also used by, amongst others, SIMS AND ZHA (1998).

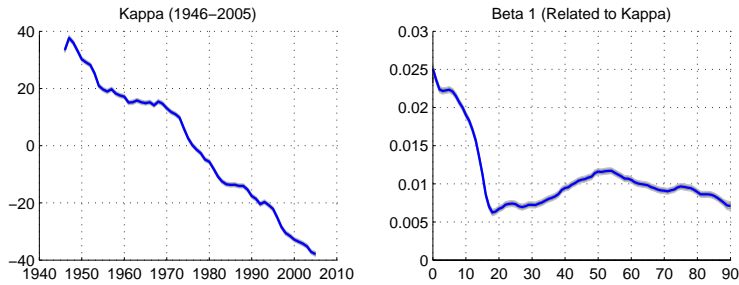


Figure 3: Estimated  $\kappa$  and  $\beta$ . The small gray shaded area around the blue median represents 90% of the posterior probability mass regarding both parameter and random term uncertainty.

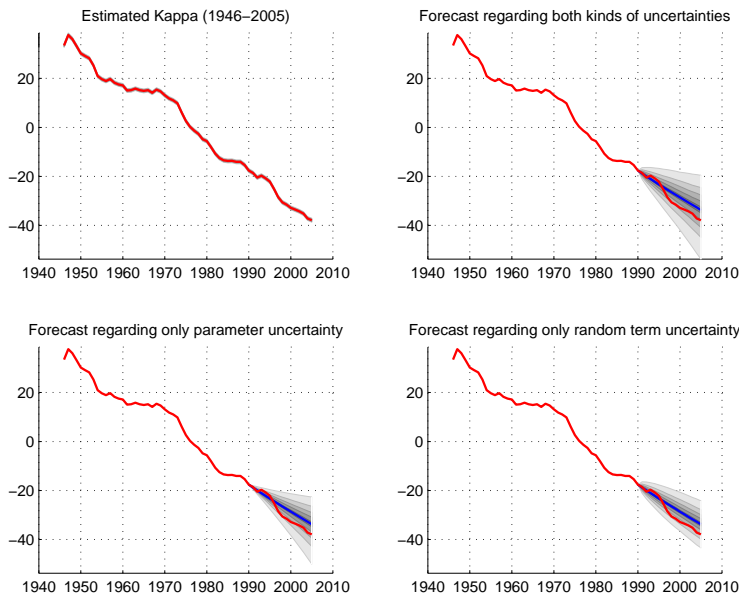


Figure 4: Panel with in-sample forecasts of  $\kappa$  with respect to different sources of uncertainty for the period 1991–2005. The red line displays always the median estimation of  $\kappa$  based on the observations for the whole period 1946–2005. The blue line displays the median forecast of  $\kappa$  based only on the information up to 1990. The entire gray shaded area represents 90%, 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  $\epsilon$ , 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 close but widen faster than they do linearly. The forecast 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.<sup>16</sup> This result demonstrates to which extent 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 precautionary 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  $\kappa$  in the observation period (starting point for the forecast) are also non-deterministic.

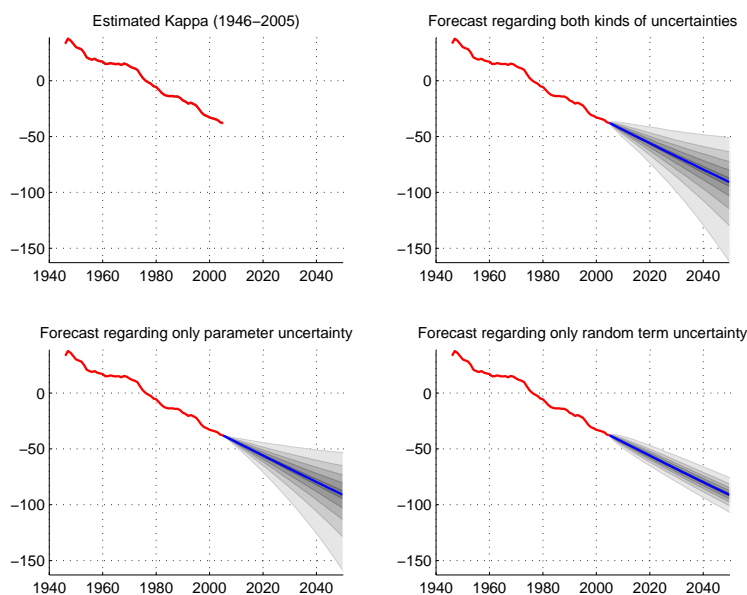


Figure 5: Panel with long-run forecasts of  $\kappa$  with respect to different sources of uncertainty for the period 2006–2050. The red line displays the median estimation of  $\kappa$  based on the observations in the period 1946–2005. The blue line displays the median forecast of  $\kappa$  based on this information. The entire gray shaded area represents 90%, 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.

<sup>16</sup>LEE AND CARTER (1992) mention a dissenting relationship in their Appendix B.

### 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  $\kappa_2$  to the specification with the two covariates. Figure 6 shows the estimated coefficients  $\beta$  related to  $\kappa_1$  and  $\kappa_2$ , GDP and unemployment revealing to what extent 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).

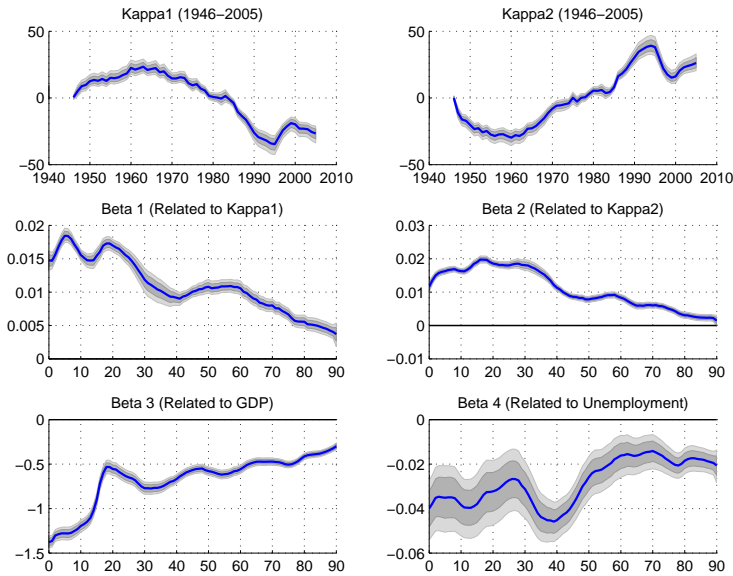


Figure 6: Estimated  $\kappa$ 's and  $\beta$ 's for the model specification with two latent variables and GDP and unemployment as covariates. The entire gray shaded area around the blue median represents 90% and the dark gray shaded area represents 68% of the posterior probability mass regarding both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

Figure 7 shows the median of 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 row of Figure 8 including macro variables as covariates improves the forecasts for the age-classes 0–12 and 60–90. Furthermore, adding  $\kappa_2$  to the two covariates leads to even better forecasts, which is shown in the third row of Figure 8. However, for the age-classes 15–35 the forecasts deteriorate. This leads us to the conclusion that the covariates have to be chosen very carefully, as they might help predicting particular age-classes and at the same time even worsen the forecasts of others.



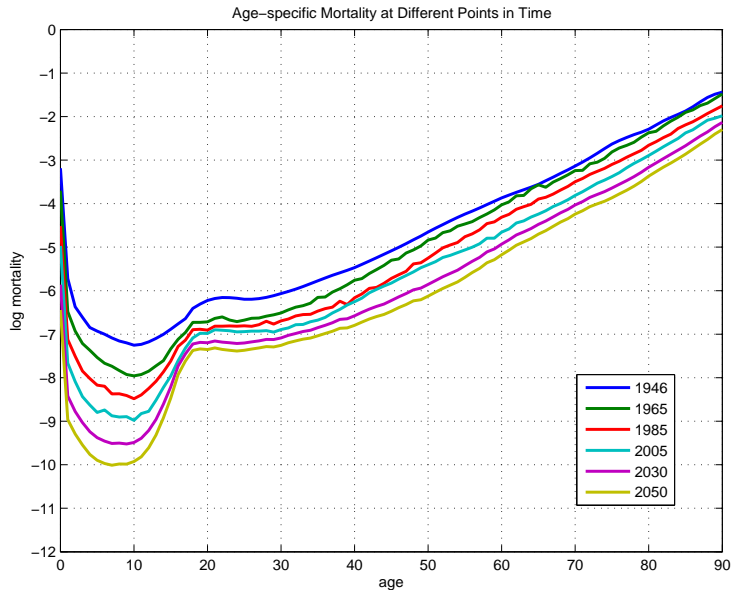


Figure 7: Observations and forecasts of age-specific mortality  $m_{x,t}$  at different points in time. The lines for the years 2030 and 2050 display the median forecasts regarding both parameter and random term uncertainty.

The figures discussed in this section demonstrate the smooth transition along the age dimension as it is described in Section 3.4. Admittedly, the difference to the Lee-Carter results is not so obvious due 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 insurances. In both cases major deviations have different consequences to smaller ones. This means that not only mean and variance, but also higher moments (skewness and kurtosis) of the distribution of predicted mortality matter. 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 quite low probability for realizations close to the median, which further challenges traditional forecast methods with misleadingly tight error bands.

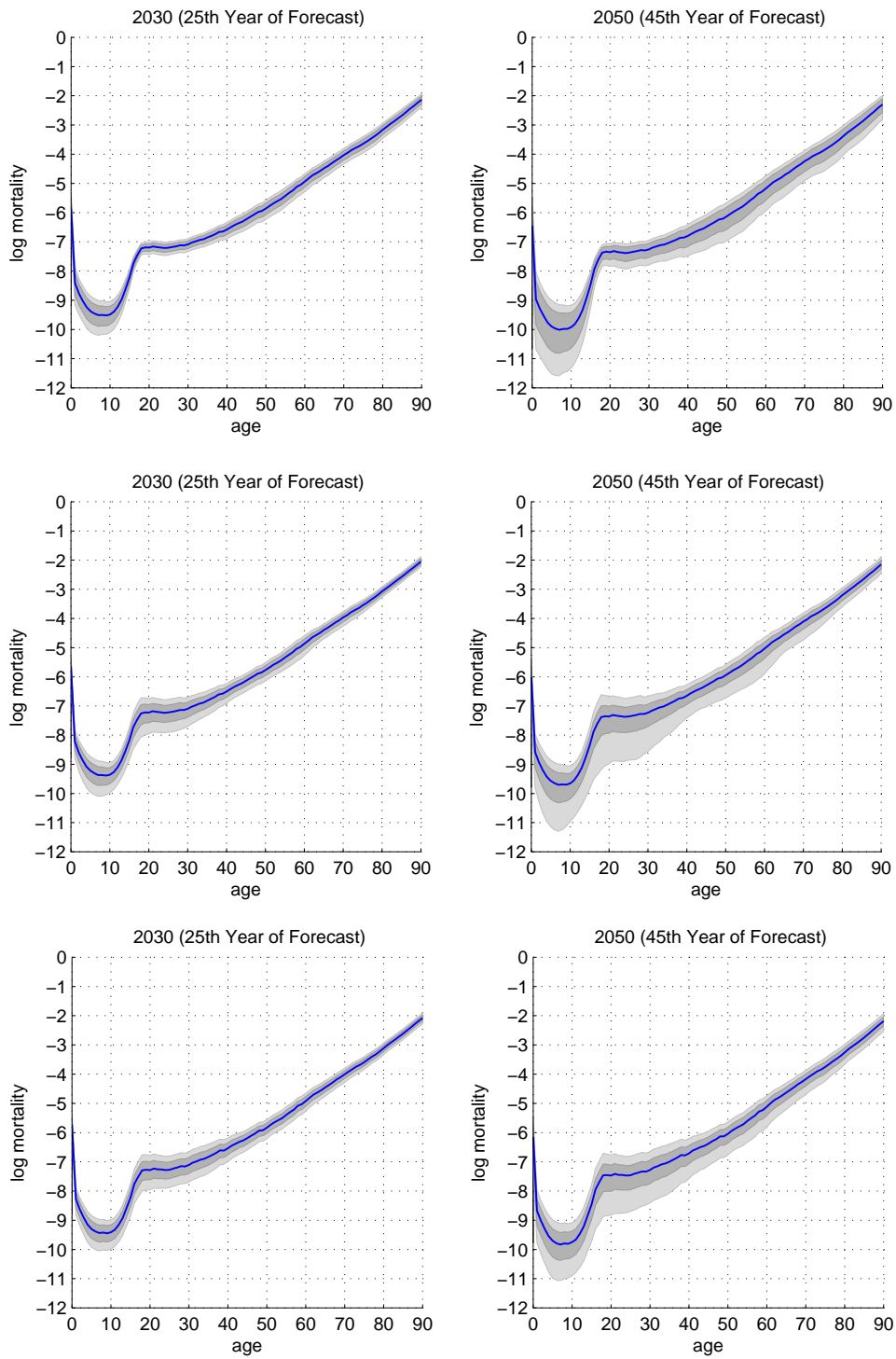


Figure 8: Panel with forecasts of age-specific mortality  $m_{x,t}$  25 and 45 years ahead for different model specifications. The first row shows the  $K = 1, N = 0$ , the second row the  $K = 1, N = 2$  and the third row the  $K = 2, N = 2$  specification. The entire gray shaded area around the blue median represents 90% 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 or 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  $\kappa$  and no covariates to forecast mortality for all age classes up to 110+.<sup>17</sup> 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 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 non-linearly from 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 displays the hypothetical birth-time probabilities  $l_{x,t}$  of surviving up to exact age  $x$  if a female would be 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 to the northeast. First, reductions of child mortality mainly shift the curve upwards, 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 displays the corresponding birth-time probabilities  $d_{x,t}$  of dying at age  $x$ . Of course, the values rise over most of the life time and peak somewhere in the old age before they fall again.<sup>18</sup> Remarkably, these probabilities do not only shift to the right, but also increasingly concentrate on a smaller age range. With respect to the survival curve, this corresponds to a transformation towards a long relatively flat initial course followed by a steep fall, which is known as rectangularization.

Finally, in Figure 11 we present time series of life expectancies at different ages for the whole observation plus forecast period 1946–2050. Life expectancy means always 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 rising 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

<sup>17</sup>The inclusion of very high ages is necessary for the best possible calculation of remaining life expectancies.

<sup>18</sup>In today's industrialized countries child mortality is no longer a major threat.

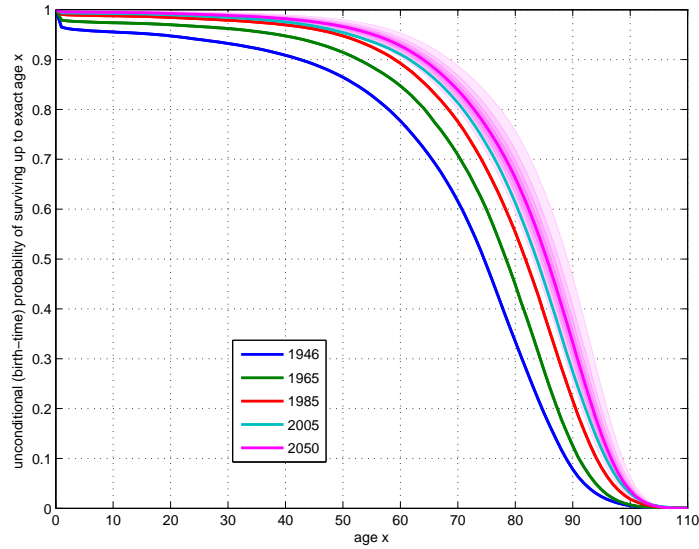


Figure 9: Probabilities  $l_{x,t}$  of surviving up to exact age  $x$  for females based on period life tables for different points in time. The figures for the years 1946–2005 are calculated from observations. The thick magenta line displays the median forecast of  $l_{x,2050}$ . The entire magenta shaded area represents 90%, 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.

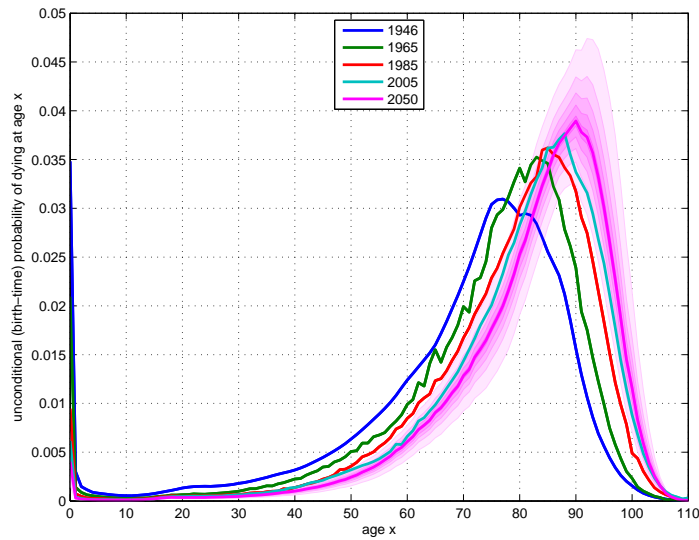


Figure 10: Probabilities  $d_{x,t}$  of dying at age  $x$  for females based on period life tables for different points in time. The figures for the years 1946–2005 are calculated from observations. The thick magenta line displays the median forecast of  $d_{x,2050}$ . The entire magenta shaded area represents 90%, 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.

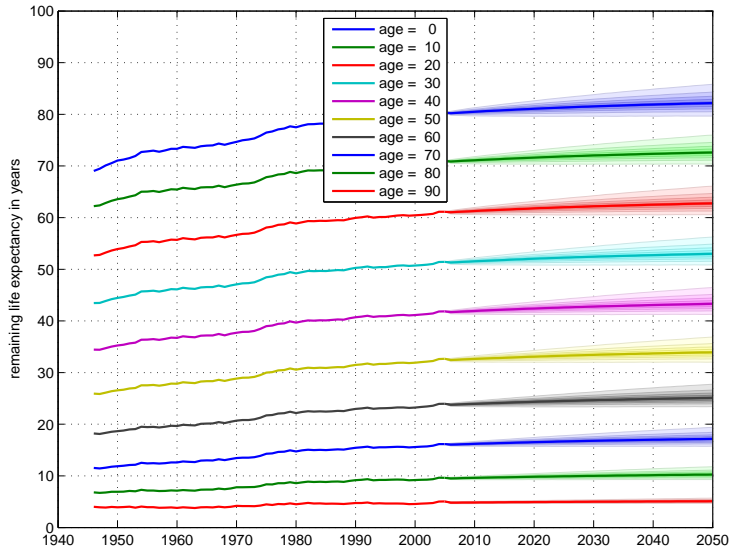


Figure 11: Remaining life expectancies  $e_{x,t}$  for females of different age classes based on period life tables. The thick lines display figures calculated from observations in the period 1946–2005 and median forecasts for  $e_{x,t}$  in the period 2006–2050. For each age class the entire shaded area represents 90% 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.

gain in female life expectancy based on period life tables between 2005 and 2050 is about 4.5 years for a new born and 2.8 years for a 60-year-old. Once again, the error bands of the forecasts can be further reduced by including covariates.

## 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 embodiment of 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 Markov Chain Monte Carlo algorithm, which enables us to estimate the parameters and the 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 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, i.e., the forecasts exhibit a smooth development along the age dimension with sufficiently tight error bands. Comparing different specifications it turns out that covariates can indeed help to 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 at the danger that hitherto existing forecasting methods for age-specific mortality, ignoring certain sources of uncertainty, 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 & 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. Forthcoming. Preliminary version provided at the URL: <http://gking.harvard.edu/files/smooth> [version 02/04/2008].
- 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: Giving some Account of the Present Undertakings, Studies and Labours of the Ingenious; In many Considerable Parts of the World*, 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). Quit your job, stay alive? Business cycles and mortality. *SFB 649 Discussion Paper*, 2008-xxx. Forthcoming.
- 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 DC. 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  $\Psi$  we condition on values for  $z$  and  $\beta$ . 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

$$Z = X\Phi + \epsilon^z, \quad (13)$$

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

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

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

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

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  $\Phi$  taken values for  $\Sigma_z^{-1}$  as given

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

and the marginal density for  $\Sigma_z^{-1}$

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

Expression (16) is a Normal density and (17) a Wishart density. Thus, the likelihood function can be described as a product of a Normal density for  $\Phi$  conditional on  $\Sigma_z$  and an inverted Wishart density for  $\Sigma_z$

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

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  $\Phi$

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

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

$$\beta^i = G^i \alpha^i + \epsilon^{\beta^i} \quad \text{for } i = 1, \dots, M, \quad (20)$$

where  $\beta^i \equiv [\beta_q^i \ \beta_{q+1}^i \ \dots \ \beta_A^i]'$  is a  $(A - q + 1) \times 1$  vector,  $\alpha^i \equiv [\alpha_1^i \ \alpha_2^i \ \dots \ \alpha_q^i]'$  is a  $q \times 1$  vector,  $\epsilon^{\beta^i} \equiv [\epsilon_q^{\beta^i} \ \epsilon_{q+1}^{\beta^i} \ \dots \ \epsilon_A^{\beta^i}]'$ , which is  $(A - q + 1) \times 1$  vector and

$$G^i \equiv \begin{bmatrix} \beta_{q-1}^i & \beta_{q-2}^i & \dots & \beta_0^i \\ \beta_q^i & \beta_{q-1}^i & \dots & \beta_1^i \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{A-1}^i & \beta_{A-2}^i & \dots & \beta_{A-q}^i \end{bmatrix}$$

is a  $(A - q + 1) \times q$  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

$$\mathcal{P}(\alpha^i, \sigma_\beta^i) = \mathcal{F}(\alpha^i | \sigma_\beta^i) \mathcal{F}(\sigma_\beta^i). \quad (21)$$

The posterior for  $\alpha^i$  conditional on the variance  $\sigma_\beta^i$  is

$$\alpha^i | \sigma_\beta^i \sim \mathcal{N} \left( \hat{\alpha}^i, \sigma_\beta^i (G^{i'} G^i)^{-1} \right), \quad (22)$$

where  $\hat{\alpha}^i$  is the OLS estimate and the marginal posterior for  $\sigma_\beta^i$  is the following inverted gamma distribution

$$\sigma_\beta^i \sim \mathcal{IG} \left( \frac{\hat{s}}{2}, \frac{(A - q)}{2} \right), \quad (23)$$

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 for 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 following Inverted Gamma distribution

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

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

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

To calculate the latent  $z$  we condition on values for  $\Psi$  and  $\beta$ . 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

$$Z_t = \tilde{\Phi} Z_{t-1} + \tilde{\epsilon}_t^z, \quad (25)$$

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,  $\tilde{\epsilon}_t^z \equiv [\epsilon_t^z \ 0 \ \dots \ 0]'$ , which is a  $(Mp+1) \times 1$  vector and

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

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

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

$$\bar{D}_t = \beta z_t + \epsilon_t^d \quad (26)$$

with

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

which is a  $(A+N) \times 1$  matrix with  $D_t \equiv [d_{0,t} \ d_{1,t} \ \dots \ d_{A,t}]'$ ,  $\bar{D} \equiv [\bar{d}_0 \ \bar{d}_1 \ \dots \ \bar{d}_A]'$ , 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 a  $(A+N) \times M$  matrix with  $\beta^\kappa \equiv [(\beta_0^\kappa)' \ (\beta_1^\kappa)' \ \dots \ (\beta_A^\kappa)']'$ , which is a  $A \times K$  matrix and  $\beta^Y \equiv [(\beta_0^Y)' \ (\beta_1^Y)' \ \dots \ (\beta_A^Y)']'$ , which is a  $A \times N$  matrix.

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

$$\bar{D}_t = H Z_t + \epsilon_t^d, \quad (27)$$

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

To calculate  $z$  we apply the algorithm suggested by CARTER AND KOHN (1994) and FRÜHWIRTH-SCHNATTER (1994).<sup>19</sup> We draw with this procedure  $z$  from its joint distribution

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

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

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

with

$$z_{T|T} = E(z_T|D), \quad (30)$$

$$P_{T|T} = Cov(z_T|D) \quad (31)$$

and

$$z_{t|t, z_{t+1}} = E(z_t|z_{t+1}, D), \quad (32)$$

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

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

$$z_{t|t-1} = \tilde{\Phi} z_{t-1|t-1}, \quad (34)$$

$$P_{t|t-1} = \tilde{\Phi} P_{t-1|t-1} \tilde{\Phi}' + Q, \quad (35)$$

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 a  $(Mp+1) \times (Mp+1)$  matrix. Accordingly the forecast error is

$$\nu_t = \bar{D}_t - H z_{t|t-1} \quad (36)$$

with the corresponding variance

$$\Omega = H P_{t|t-1} H' + R, \quad (37)$$

---

<sup>19</sup>Cf. also KIM AND NELSON (1999).

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

$$K_t = P_{t|t-1} H' \Omega^{-1}. \quad (38)$$

Thus, the updating equations are:

$$z_{t|t} = z_{t|t-1} + K_t \nu_t, \quad (39)$$

$$P_{t|t} = P_{t|t-1} + K_t H P_{t|t-1}. \quad (40)$$

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

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

and

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

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

To calculate  $\beta$  we take values for  $\Psi$  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  $x = 0, 1, \dots, A$  and not in  $t = 1, 2, \dots, T$  as in section A.2.

Our state equation can be expressed as

$$\tilde{\beta}_x = \tilde{\alpha} \tilde{\beta}_{x-1} + \tilde{\epsilon}_x^\beta, \quad (43)$$

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

$$\tilde{\alpha} = \begin{bmatrix} \alpha_1 & \dots & \alpha_q \\ I_{M(p-1) \times M(p-1)} & & 0_{M(p-1) \times (M+1)} \end{bmatrix},$$

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

$$\tilde{D}_x - \bar{d}_x = W \tilde{\beta}_x + \epsilon_x^d, \quad (44)$$

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

$$Q \equiv \begin{bmatrix} \Sigma_\beta & 0_{M \times M(q-1)} \\ 0_{M(q-1) \times M(q-1)} & 0_{M(q-1) \times M} \end{bmatrix}$$

we can apply the procedure described in section A.2 to calculate  $\beta$ .

## 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 results for the probability  $l_{x,t}$  of surviving up to 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 being still alive at exact age  $x$ . The respective calculations are standard.<sup>20</sup>

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  $x$ -th 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):

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

and

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

Consistent values for  $\alpha_{0,t}^{total}$  would require information about 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 total figures of both sexes combined, we adopt a simple approximation roughly reflecting 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 perceptibly influence the results. The highest recorded age class  $\tilde{x}$  is open, i.e., not restricted to one 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

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

---

<sup>20</sup>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 numbers  $l_{x,t}$  and  $d_{x,t}$  of survivors and deaths out of 100000 live-births.

The person-years lived at age  $x$  and from age  $x$  onwards are

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

and

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

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 basic variables of several age classes are synthetic measures for this period. They mix up 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 are adequate, which are calculated from data of a single cohort 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 higher 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.



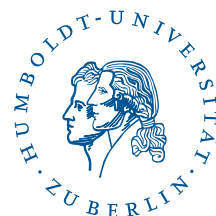
## SFB 649 Discussion Paper Series 2008

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, January 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.

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

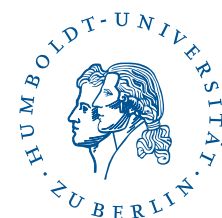
This research was supported by the Deutsche  
Forschungsgemeinschaft through the SFB 649 "Economic Risk".



- 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.
- 042 "Gruppenvergleiche bei hypothetischen Konstrukten – Die Prüfung der Übereinstimmung von Messmodellen mit der Strukturgleichungsmethodik" 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.
- 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.

**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".



- 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".

