Insurance: Mathematics and Economics 45 (2009) 410–423

Contents lists available at ScienceDirect

Insurance: Mathematics and Economics

journal homepage: www.elsevier.com/locate/ime

Dynamic mortality factor model with conditional heteroskedasticity

Quansheng Gao[a,*,1], Chengjun Hu[b,1]

[a] Department of Mathematics and Physics, Wuhan Polytechnic University, Wuhan, 430023, PR China

[b] Department of Anatomy and Embryology, Wuhan University School of Medicine, Wuhan, 430071, PR China

ARTICLE INFO

Article history:
Received January 2009
Received in revised form
June 2009
Accepted 2 September 2009

Keywords:
Lee–Carter model
Generalized dynamic factor model
Multivariate generalized autoregressive conditionally heteroskedastic model
Mortality forecasting
Kalman filter

ABSTRACT

In most methods for modeling mortality rates, the idiosyncratic shocks are assumed to be homoskedastic. This study investigates the conditional heteroskedasticity of mortality in terms of statistical time series. We start from testing the conditional heteroskedasticity of the period effect in the naïve Lee–Carter model for some mortality data. Then we introduce the Generalized Dynamic Factor method and the multivariate BEKK GARCH model to describe mortality dynamics and the conditional heteroskedasticity of mortality. After specifying the number of static factors and dynamic factors by several variants of information criterion, we compare our model with other two models, namely, the Lee–Carter model and the state space model. Based on several error-based measures of performance, our results indicate that if the number of static factors and dynamic factors is properly determined, the method proposed dominates other methods. Finally, we use our method combined with Kalman filter to forecast the mortality rates of Iceland and period life expectancies of Denmark, Finland, Italy and Netherlands.

1. Introduction

A remarkable increasing interest has been observed in the past few years in methods that forecast mortality rates. Stochastic models of mortality are a significant topic of current research in actuarial science and demography. The modeling approach proposed by Lee and Carter (1992) is amongst the most widely used mortality trend fitting and projection tools. The Lee and Carter model (hereafter LC model) is a parsimonious demographic model combined with statistical time-series methods. The model assumes that the dynamics of the logarithm of central death rates over time are driven by a single time varying period effect. The mortality projection relies on the extrapolation of the period effect under an appropriate statistical time-series model.

The original LC model entails two equations. The first equation decomposes a time series of age-specific vital rates into three sets of parameters: age pattern parameters, period effect parameters and parameters that represent age-specific reactions to the period-specific effect. The second equation is a model of the time path of the period effect. The LC model adopts a combination of singular value decomposition (SVD) estimation for the first equation and time-series methods for modeling the evolution of the period effect (Wolf, 2004).

Over the last decade, a host of extensions have been proposed to improve the LC model. Renshaw and Haberman (2003) propose adding additional time varying factors. Renshaw and Haberman (2006) show that a cohort effect is required in order to fit gender-based 1961–2003 UK data. Hyndman and Ullah (2007) develop a multi-factor modeling approach using functional principal components to fit demographic data. See Sherris and Wills (2008) for recent discussions.

The LC model in its original form has some shortcomings although it is relative simple, easily applied and provides fairly accurate mortality projections. The simplicity of the LC model means that the error terms in two equations are assumed to be white noise with zero mean and small constant variance over all ages and all times, i.e., the model error terms are homoskedastic. The assumption of constant volatility, however, is always unrealistic: the observed logarithm of central death rates is much more variable and the volatility is time varying (Lee and Miller, 2001).

In this paper, we focus on the mortality heterogeneity in terms of statistical time series. Not only would the mortality rates fluctuate but also the patterns of mortality may change over time due to many reasons. Weiland et al. (2006) observe changes in mortality for varying ages at different times. The cohort effect such as changes in human behavior has contributed significantly to the volatility in mortality. A sustained fall in smoking prevalence may lead to lower gains in mortality improvement from the effects of smoking behavior in future years (Olshansky et al., 2005). Deaths related to AIDS, drug, alcohol abuse and violence make the future course of mortality rates at young ages considerably uncertain. Increasing obesity in children and young adults has continued

* Corresponding author. Tel.: +86 27 65087648.
E-mail addresses: gqsh0101@yahoo.com.cn (Q. Gao), chengjunh2003@yahoo.com.cn (C. Hu).

[1] The two authors contributed equally to this work.

0167-6687/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.insmatheco.2009.09.001

to lead to a regime of decreasing life expectancy (Flegal et al., 2004). Whilst medical advances and other factors will continue to lead to further mortality decline, infectious diseases, for example tuberculosis, SARS, hepatitis C and HIV, could work in the opposite direction. It is always supposed that there is a smooth underlying force of mortality. However, Kerkby and Currie (2008) find that observed mortality is subjected to more than stochastic deviation from the smooth surface; for example, flu epidemics, hot summers or cold winters can disproportionately effect the mortality of certain age groups in particular years. They call such an effect a period shock. In sum, when one chooses an appropriate model for forecasting future mortality trends, one must foresee whether the model would reflect the heterogeneity (Gallop, 2007).

The warnings about heteroskedasticity, however, have usually been applied only to cross sectional models, not to time-series models. Schrager (2006) observes that the logarithm of the observed force of mortality is much more variable at older ages than at younger ages because of the much smaller absolute number of deaths at older ages. They model heterogeneity by adding a heterogeneous component to mortality intensity. Bauer et al. (2008) specify adequate volatility structures for forward models to project mortality. So far knowledge of the volatility of mortality rate has been still insufficient and we know little about mortality heterogeneity in time series.

Our approach, inspired from method proposed by Alessi et al. (2006), which combines the Generalized Dynamic Factor Model (GDFM) and the multivariate Generalized Autoregressive Conditionally Heteroskedastic (GARCH) model, purposes to avoid the drawback of time-series heterogeneity inherent to the original mortality forecasting methodology. Our model is a multivariate forecasting method and can also be seen as a two-layer model. The external layer involves the observed mortality processes. This process is assumed to follow a measurement equation that contains two parts: the common component and the idiosyncratic component. The common component is a combination of state vector of the system. It is specified in such a way that it captures the long and short run relationships among mortality processes of different ages. The internal layer is called the state equation where state vectors are similar to period effects in the LC model. State vectors, which are assumed to be a random walk process in the LC model, are assumed to follow a VAR(1) process.

Our specifications are similar to the state space framework that comprises of two latent components for each variable. The advantages of utilizing the state space approach are two fold. First, the state space model can capture most of the common properties (correlation) among separate age groups. Inter-series relationships can also be disaggregated to the latent component level. Second, the state space model allows for both a dynamic representation of the common components and non-orthogonal idiosyncratic components. These advantages enable us to address the question of whether an increase in volatility in one age induces additional volatility in another. These advantages also provide us a greater degree of insight that may be useful for mortality forecasting.

Thus modeling time-series heterogeneity of mortality can be specified from two sides. On one hand, error terms in measurement equation are assumed to follow ARCH or GARCH models. These models are widespread tools for dealing with time-series heteroskedasticity. On the other hand, we use the BEKK model of Engle and Kroner (1995) or Dynamic Conditional Correlation (DCC) model of Engle (2002) for error terms in state equation. The BEKK model is a parameterization of conditional variance that guarantees positivity of the conditional covariance matrix and reduces the number of parameters to be estimated. The DCC model adapts GARCH models specifically for the estimation of time varying correlations. The BEKK model focuses on the dynamic of the conditional covariance matrix, whereas the DCC model focuses on the dynamic of the conditional variances and the conditional correlation matrix.

The rest of the paper will be organized as follows. Section 2 reviews the LC model and provides some detailed tests for conditional heteroskedasticity of the period effect in the original LC model. Section 3 presents our dynamic mortality factor model with conditional heteroskedasticity and briefly describes the estimation procedure and forecasting methods with Kalman filter. Section 4 details some hypothesis tests and specifications such as the number of static factors and dynamic factors and tests for conditional heteroskedasticity. Section 5 justifies out-of-sample performance of the proposed method. Sections 6 and 7 provide some forecast results for Iceland and some more countries. Concluding remarks can be found in Section 8.

2. Testing the conditional heteroskedasticity of the period effect in the original LC model

In this section, we first outline the LC model and the way it is usually estimated. Then we will test the conditional heteroskedasticity of the period effect. In the next section we will introduce our alternative approach.

2.1. Overview of the LC model

Because this paper focuses on demography applications, we use demography notations. Let $m_{x,t}$ denote the logarithm of central death rates for age x in year t with $x \in \{1, 2, \dots, N\}$ and $t \in \{1, 2, \dots, T\}$. Let $m_t = (m_{1,t}, m_{1,t}, \dots, m_{N,t})'$ be an N-dimensional vector process, where the notation M' indicates the transpose of matrix M.

In the original LC model, it is assumed that each mortality rate series m_t can be written as the sum of two mutually orthogonal unobservable components, the common component κ_t and the idiosyncratic component ε_t with $\varepsilon_t = (\varepsilon_{1,t}, \varepsilon_{2,t}, \dots, \varepsilon_{N,t})'$. The common component is driven by age effect $\theta^{(x)}$ that describes the relationship between ages and mortality rates and period effect f_t that captures the impact that time-specific events have on the number of deaths experienced in a population, including effects such as general health status of the population, availability of health services, and critical weather conditions (Olivieri, 2007). We neglect here the cohort effect factor that captures the influence of year of birth on mortality improvement rates. Thus the LC model can be formulated as:

$$m_{x,t} = f^{(x)} + \theta^{(x)} f_t + \varepsilon_{x,t} \quad (x = 1, 2, \dots, N,\ t = 1, 2, \dots, T) \quad (1)$$

or

$$m_t = f + \theta f_t + \varepsilon_t = \kappa_t + \varepsilon_t \quad (2)$$

where $f = (f^{(1)}, f^2, \dots, f^{(N)})'$, $\theta = (\theta^{(1)}, \theta^2, \dots, \theta^{(N)})'$ and $\kappa_t = f + \theta f_t \cdot f_t$ is assumed to be governed by:

$$f_t = d + f_{t-1} + u_t. \quad (3)$$

Thus f_t is a random walk with drift. In the original LC model, error term vector ε_t and the error term u_t are supposed to be white noise, satisfying the distributional assumption:

$$\begin{pmatrix} \varepsilon_t \\ u_t \end{pmatrix} / \mathbb{F}_{t-1} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \Sigma_\varepsilon & 0 \\ 0 & \sigma_u^2 \end{pmatrix} \right) \quad (4)$$

where $\mathbb{F}_{t-1}$ contains all the information available at time t, Σ_ε is the unknown covariance matrix of ε_t and σ_u^2 is the unknown variance of u_t.