



AARHUS
UNIVERSITY
BUSINESS AND SOCIAL SCIENCES
DEPARTMENT OF ECONOMICS AND BUSINESS



CREATES

Center for Research in Econometric Analysis of Time Series

Deterministic and stochastic trends in the Lee-Carter mortality model

Laurent Callot, Niels Haldrup and Malene Kallestrup Lamb

CREATES Research Paper 2014-44

Deterministic and stochastic trends in the Lee-Carter mortality model

LAURENT CALLOT^{abcd} NIELS HALDRUP^{cd} MALENE KALLESTRUP LAMB^{cd}

November 19, 2014

ABSTRACT. The Lee and Carter (1992) model assumes that the deterministic and stochastic time series dynamics loads with identical weights when describing the development of age specific mortality rates. Effectively this means that the main characteristics of the model simplifies to a random walk model with age specific drift components. But restricting the adjustment mechanism of the stochastic and linear trend components to be identical may be a too strong simplification. In fact, the presence of a stochastic trend component may itself result from a bias induced by properly fitting the linear trend that characterizes mortality data. We find empirical evidence that this feature of the Lee-Carter model overly restricts the system dynamics and we suggest to separate the deterministic and stochastic time series components at the benefit of improved fit and forecasting performance. In fact, we find that the classical Lee-Carter model will otherwise over estimate the reduction of mortality for the younger age groups and will under estimate the reduction of mortality for the older age groups. In practice, our recommendation means that the Lee-Carter model instead of a one-factor model should be formulated as a two (or several)-factor model where one factor is deterministic and the other factors are stochastic. This feature generalizes to the range of models that extend the Lee-Carter model in various directions.

KEYWORDS: Mortality modelling, factor models, principal components, stochastic and deterministic trends.

JEL-CLASSIFICATIONS: C2, C23, J1, J11.

1. INTRODUCTION

One of the most commonly used models to forecast age-specific mortality rates is based on the Lee and Carter (1992) model. The model has attracted a lot of attention and has become a benchmark for mortality modelling and life table predictions although the model has also been subject to criticism. Even though the model was mainly intended to describe the statistical variation in all-cause mortality in the United States and similar developed countries, the model is now widely used to predict all-cause and cause specific mortality for a large range of developed and less

^aFree University Amsterdam, ^bTinbergen Institute, ^cAarhus University, ^dCREATES.

Corresponding author: Niels Haldrup, Department of Economics and Business, CREATES, Aarhus University, nhaldrup@creates.au.dk. This research is supported by CREATES - Center for Research in Econometric Analysis of Time Series (DNRF78), funded by the Danish National Research Foundation. The program code and data is accessible at <https://github.com/lcallot/LeeCarter>.

developed countries around the world, see e.g. Lee (2000), Booth et al. (2002), Renshaw and Haberman (2003, 2010), Girosi and King (2006) amongst many others.

The basic Lee-Carter model is rather simple. It describes the age specific (log) mortality in terms of age-specific intercepts and a single time factor (known as the mortality index) with age-specific loadings. In most applications the mortality index is modelled as a random walk with drift which is a fundamental assumption when using the model for projections into the future. The model parameters, i.e. the age-specific intercepts and loadings and the time factor, can be estimated rather easily by use of singular value decomposition of the matrix of historical mortality rates over age-groups and time. It is well documented that over long histories of time log (all cause) mortality has evolved linearly which is clearly the most dominant dynamic feature of the data. The assumption that the time-factor of the Lee-Carter model is governed by a random walk with drift will clearly capture this feature, but because the model is a one factor model it also means that the loadings of the time trend (or drift) as well as any stochastic components will be the same even though the order of the linear trend will dominate the stochastic trend. This may appear to be an inappropriate restriction of the model dynamics. In fact, because an empirical regularity of log mortality measured over time is a dominant linear trend, the recommendation of Lee and Carter to extract a single factor via singular value decomposition effectively implies that the linear time trend will determine that factor and thus the model simplifies to a random walk with age specific drift terms and with a covariance structure that potentially is biased. The common practice of modelling the mortality time index as a random walk with drift can be the result of modelling the time index as a single factor whereby the stochastic component will be biased to have a unit root. A similar reasoning can also be found in Girosi and King (2007).

In this paper we argue that an improved model fit can be gained by separating the deterministic and stochastic dynamics of the model and by allowing different loadings depending upon the source of the dynamics. Still, the model encompasses the basic Lee-Carter model as a special case. We argue that a two (or several) factor model with one deterministic factor and the other factor(s) being stochastic is preferable when in fact the mortality series exhibit a strong linear trend. For a number of countries we demonstrate empirically that the improved fit can be significant and that the loadings of the deterministic and stochastic factors are indeed rather different. We find that the classical Lee-Carter model will tend to over estimate the reduction of mortality for the younger age groups and will under estimate the reduction of mortality for the older age groups. Generally, the transient dynamics of the log mortality will be rather different than the long run trend and hence will have implications in particular for short and medium term forecasts.

Even though the analysis of the present paper uses the classical Lee-Carter model as the benchmark model our results generalize to the range of models with strongly negatively trending mortality that extends the Lee-Carter model such as Lee and Miller (2001), Booth et al. (2002), Renshaw and Haberman (2003, 2010) and Plat (2009).

The plan of the paper is the following. In section 2 some features of the classical Lee-Carter model are presented and in section 3 we suggest a modification of the model that facilitates separation of the deterministic and stochastic dynamics. Subsequently, empirical illustrations to gender specific (all cause) mortality data for U.S., Japan, and France are provided to demonstrate the advances of the modified model.

2. THE CLASSICAL LEE-CARTER MODEL

Assume that we have observed age-specific death rates m_{xt} for a set of calendar years $t = t_1, t_2, \dots, t_n$ and ages $x = x_1, x_2, \dots, x_m$. Subsequently, we will refer to m_{xt} as the age-specific mortality rates which are assumed to be constant within each interval of age and time, but is allowed to vary from one interval to the next. The classical Lee-Carter (CLC) model assumes a log bi-linear model for mortality rates along the age dimension in terms of the parameters α_x and β_x , and along the time dimension by the time factor k_t :

$$\ln m_{xt} = \alpha_x + \beta_x k_t + \varepsilon_{xt} \quad (1)$$

The parameters α_x represent age-specific constants describing the general pattern of mortality averaged over time. k_t , known as the mortality index, summarizes the development in the level of mortality over time and thus will capture the general time trend of the death rates. The parameters β_x measure the loadings to the particular age groups when the mortality index changes. The error term ε_{xt} has mean zero and variance σ_ε^2 and reflects the age-specific historical fluctuations not captured by the model. Often ε_{xt} is assumed to be normally distributed. Notice that a total of $m \times n$ observations m_{xt} are available for estimation and the model thus needs $2m + n + 1$ parameters to be estimated, i.e. α_x, β_x, k_t and σ_ε^2 . As seen, the one-factor model (1) is a special case of a principal components model with $r = 1$ principal component.

The model parameters are not identified, but a simple identification scheme can be chosen. It is common to impose the constraints $\sum_t k_t = 0$ and $\sum_x \beta_x = 1$ whereby it follows that

$$\hat{\alpha}_x = \overline{\ln m_x} = n^{-1} \sum_t \ln m_{xt} \quad (2)$$

that is, the empirical average over time of the age profile for age group x .

Subsequently, $\hat{\beta}_x$ and \hat{k}_t can be calculated by singular-value decomposition (SVD) of the $m \times n$ matrix of demeaned log mortality rates $\overline{M} = \{\ln m_{xt} - \overline{\ln m_x}\}$ subject to the chosen identification scheme. For $\overline{M} = ULV'$ the estimate of β_x is given by the first column of U corresponding to the largest singular value and scaled to have unit variance. k_t is subsequently calculated as $k_t = \beta' \overline{M}_t$, where \overline{M}_t is the row mean of \overline{M} .

After estimation of the model parameters and extraction of the mortality index \hat{k}_t , the next step is to model a process for \hat{k}_t which can be used to generate h -steps ahead forecasts $E(k_{t+h}|t)$ that serve as input to the forecasts $E(\ln m_{x,t+h}|t)$. Typically the model for \hat{k}_t is based on an ARIMA time series specification and in most cases it

is found appropriate to model \widehat{k}_t as a random walk with drift, e.g. $\Delta\widehat{k}_t = \mu + v_t$. A consequence of this is that the n log mortality rates each cointegrate pairwise and are driven by the same stochastic trend, see Lazar and Denuit (2009). For a given sample, the level of the mobility index can be described as

$$k_t = k_1 + \mu(t - 1) + \sum_{j=2}^t v_j, \quad \text{for } t = 2, \dots, n \quad (3)$$

where the drift parameter μ can be estimated as

$$\widehat{\mu} = \frac{1}{n-1}(\widehat{k}_n - \widehat{k}_1) \quad (4)$$

From these assumptions, the mortality index is seen to be governed by a linear trend as well as a stochastic trend (random walk) component, $\sum_{j=2}^t v_j$. By mean-adjusting the trend, log mortality can thus be described as

$$\ln m_{xt} = \alpha_x + \beta_x \mu(t - \bar{t}) + \beta_x \left\{ \sum_{j=2}^t v_j + \frac{\mu}{2}(n-1) + k_1 \right\} + \varepsilon_{xt} \quad (5)$$

where $\bar{t} = n^{-1} \sum_t t = \frac{1}{2}(n+1)$ and the identifying restrictions imply that $\frac{1}{n} \sum_{t=1}^n (\sum_{j=2}^t v_j) + \frac{\mu}{2}(n-1) + k_1 = 0$. Hence the random walk with drift assumption means that log mortality for each age group is affected through identical loadings β_x associated with the linear trend term and the (level corrected) stochastic trend component.

3. SEPARATING DETERMINISTIC AND STOCHASTIC DYNAMICS OF THE LEE-CARTER MODEL

In terms of long range forecasting it is clear that the drift (or trend) component of the mortality index will dominate the stochastic trend component so the fact that the loadings are restricted to be the same for the two components will have little impact for the long horizon point forecasts. However, a more flexible specification of the CLC model seems natural in light of the foregoing discussion. Consider a modified Lee-Carter model which allows the loadings of the deterministic and the stochastic trend to be different, that is

$$\ln m_{xt} = \widetilde{\alpha}_x + \widetilde{\gamma}_x(t - \bar{t}) + \widetilde{\beta}_x \widetilde{k}_t + \widetilde{\varepsilon}_{xt} \quad (6)$$

where \widetilde{k}_t is a time series process which typically will be modelled as a random walk *without* drift or as a stationary process, e.g. an AR(p) process. We will denote this model specification the detrended Lee-Carter (DLC) model. Notice, that if $\widetilde{\gamma}_x = \widetilde{\beta}_x$ then the CLC model (1) with a random walk plus trend specification of k_t and the DLC model (6) will coincide. However, to the extent that the loadings are different it is expected that the DLC will be superior in terms of model fit as well as in terms of forecasting due to the increased flexibility of the model. Also, the (stochastic)

time series properties of \tilde{k}_t may be rather different from those of k_t unless the CLC model is the right specification. In fact, if the model (6) with $\tilde{\gamma}_x \neq \tilde{\beta}_x$ is the correct model specification, then it is likely that estimation of the stochastic component of k_t based on the model (1) will be biased due to the presence of a linear trend component that will dominate k_t overall. Furthermore, if the Lee-Carter model (1) is estimated and the true model is (6) with $\tilde{\gamma}_x \neq \tilde{\beta}_x$ then the covariance structure is obviously distorted and will depend upon the estimate β_x leading to a biased two-stage Lee-Carter estimator, see also Girosi and King (2007). The cost of the more flexible DLC model specification is that m additional parameters need to be estimated but this is necessary to avoid the bias that the Lee-Carter model would otherwise imply when $\tilde{\gamma}_x \neq \tilde{\beta}_x$.

The estimation of the modified model can be conducted by prior detrending and an identification scheme along the lines of the classical Lee-Carter model can be easily conducted. If we impose the identifying restrictions of Lee and Carter, $\sum_t \tilde{k}_t = 0$ and $\sum_x \tilde{\beta}_x = 1$ it is seen that

$$\hat{\alpha}_x = \hat{\alpha}_x = n^{-1} \sum_t \ln m_{xt} \quad (7)$$

In fact, both $\tilde{\alpha}_x$ and $\tilde{\gamma}_x$ can be straightforwardly estimated by least squares detrending via the regression

$$\ln m_{xt} = \tilde{\alpha}_x + \tilde{\gamma}_x(t - \bar{t}) + \omega_t \quad (8)$$

and subsequently extracting $\tilde{\beta}_x$ and \tilde{k}_t by singular value decomposition of the matrix of *detrended* mortality rates $\overline{M}^* = \left\{ \ln m_{xt} - \tilde{\alpha}_x - \tilde{\gamma}_x(t - \bar{t}) \right\}$. It is well known that even when ω_t has a unit root the least squares estimate $\hat{\tilde{\gamma}}_x$ will be consistent.

4. EMPIRICAL ILLUSTRATION

In this section we compare some of the empirical features of the CLC and DLC models. The mortality series considered are all-cause mortality rates for men and women for France, Japan, and the USA. The sample period is 1950-2010. The source of the data is The Human Mortality Database¹.

Figure 1 displays the log mortality for selected age groups, $x = (0, 1, 50, 60, 80, 90)$. Notice that the scale of the graphs is different due to the various age groups being differently exposed to death. As seen, all series exhibit a strong decline in mortality over time that is almost linear but clearly with a different slope depending upon the age group considered. In Figure 2 the different trend slopes as a function of age group x is shown for the six mortality series. The trend coefficient $\hat{\tilde{\gamma}}_x$ is estimated from the detrending regression (8). Despite the different shape across countries and genders the graph also demonstrate the general similarity across age groups.

¹Human Mortality Database. University of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany). Available at www.mortality.org or www.humanmortality.de (data downloaded September 9, 2014).

To compare the trend estimates from the DLC and CLC models figure 3 displays a cross plot of the trend estimates. To facilitate the comparison the trend slope of the CLC is calculated as the estimate $\beta_x \mu$ where μ is estimated as in (4) and β_x is estimated from the CLC singular value decomposition. The two estimates are rather similar as one would expect since the dominant feature of the mortality series is the steadily declining trend. The estimates mainly differ for middle aged and people in their 20s but otherwise are almost identical.

Figures 1-5 about here

The fact that the trend slope estimates generally are similar for the two models is an indication that the one-factor CLC model is driven by the linear trend which potentially may bias the stochastic part of the mortality index k_t . To examine this further, the CLC and DLC estimates β_x and $\tilde{\beta}_x$ are compared in Figure 4. In Figure 5 k_t of the CLC model corrected for the trend, $k_t - k_1 - \hat{\mu}(t - 1)$, is compared with \tilde{k}_t estimated from the DLC model (6). By construction, the shape of the β_x curves for the CLC is driven by the trend slope. As seen, the separation of the dynamics underlying the DLC model gives a very different picture. If the CLC model was correctly specified, then the curves for the CLC and DLC would be identical, but obviously this is far from being the case. Discrepancies are seen to be present especially for the younger ages, but a general pattern is that the constraints imposed by the CLC relative to the detrended model leads to over estimating the reduction of mortality for children and young people and under estimate the decrease of mortality for older age groups.

Turning to the dynamics of the (detrended) mortality indices, the shape seems to be rather similar and indeed exhibits a fairly high degree of persistence. Notwithstanding, visual inspection of Figure 5 indicates that the CLC detrended k_t series has larger persistence than the DLC \tilde{k}_t series. This is also confirmed by estimating the AR(1) parameter of each of the two trend terms. Table 1 shows that the AR parameter and hence the persistence of the trend series is always smaller for the DLC model compared with the CLC model. The sample size is not big enough to reject a unit root in either series, but after all the estimates indicate that the CLC estimate of the stochastic component is biased towards persistency.

To compare the model fit of the various model specifications Table 2 displays the (pseudo) coefficient of determination R^2 for a range of models. The results are reported for men and women as well as for the total death rates. The coefficient of determination is calculated across all time periods and age groups, that is

$$R^2 = \frac{\sum_x \sum_t \hat{v}_{xt}^2}{\sum_x \sum_t (\log m_{xt} - \overline{\log m_x})^2} \quad (9)$$

where \hat{v}_{xt}^2 is the squared estimation error from the relevant model specification. The numerator is the total sum of squares of log mortality in deviations from the mean.

Tables 1-2 about here

The first three columns report R^2 for the CLC, DLC and the detrending regression (8). It is remarkable that the linear trend model alone accounts for 90-95 pct of the total variation in the data. The CLC model contributes with additionally 2-4 percentage points in explanatory power. Notwithstanding, by separating the deterministic and stochastic trend components as done in the DLC model additionally 2-3 percentage points can be gained in terms of fit.

Since the fit is already very high due to the presence of a trend, we also calculated an alternative measure of fit where log mortality was measured in deviations from the deterministic trend fitted by either the CLC or DLC model, i.e. the numerator in (9) used detrended log mortality series. The fourth column of Table 2 shows the fit of the CLC measured in deviations from the mean and trend, and the last column displays the fit for the DLC model. The improved fit of the DLC model, after correcting for the linear trend, is remarkable. For most cases the fit of the DLC model is more than 50% larger than for the CLC model, and for Japan (total) and France (total) the fit is even larger. These results stress the importance of treating deterministic and stochastic components differently when modelling mortality data that is strongly trending.

5. CONCLUSION

Many age-specific log mortality series exhibit a very strong negative time trend. We demonstrate that the dynamic model features of the Lee-Carter mortality model are distorted by the presence of a deterministic trend component. However, within the Lee-Carter class of models a simple solution exists where essentially the detrended rather than the demeaned mortality data is used in the analysis. We are not claiming that this is the right modelling approach for all purposes, but it shows that separation of the deterministic and stochastic components is generally important and may lead to improved fit for mortality model building and hence will potentially lead to improved forecasting performance.

REFERENCES

- [1] Booth, H, Maindonald J, Smith L., 2002. Applying Lee-Carter under conditions of variable mortality decline. *Population Studies* **56**, 325-336.
- [2] Girosi, F., and King G., 2006. *Demographic Forecasting*. Cambridge University Press: Cambridge, 2006.
- [3] Girosi, F., and King G., 2007. Understanding the Lee-Carter Mortality Forecasting Method, Mimeo, Harvard University.
- [4] Lazar, D., and Denuit M. M., 2009. A multivariate time series approach to projected life tables. *Applied Stochastic Models in Business and Industry* **25**, 806-823.
- [5] Lee, R. D., and Carter L.R., 1992. Modelling and forecasting U.S. mortality. *Journal of the American Statistical Association* **87**, 659-671.
- [6] Lee, R. D., 2000, The Lee-Carter method for forecasting mortality, with various extensions and applications. *North American Actuarial Journal* **4**, 80-91.
- [7] Lee, R. D., and Miller T, 2001. Evaluating the performance of the Lee-Carter model for forecasting mortality. *Demography* **38**, 537-549.
- [8] Plat, R., 2009. On stochastic mortality modelling. *Insurance: Mathematics and Economics* **45**, 393-404.
- [9] Renshaw, A. E. and S. Haberman, 2003. Lee-Carter mortality forecasting with age specific enhancement. *Insurance: Mathematics and Economics* **33**, 255-272.
- [10] Renshaw, A. E. and S. Haberman, 2010. Lee-Carter mortality forecasting: a parallel generalized linear modelling approach for England and Wales mortality projections. *Applied Statistics* **52**, 119-137.

6. TABLES AND FIGURES

Table 1. AR(1) parameter estimates for k_t corrected for mean (CLC) and for \tilde{k}_t (DLC).

Country	Gender	CLC	DLC
USA	female	0.993	0.919
USA	male	1.015	0.968
JPN	female	0.968	0.925
JPN	male	0.973	0.898
FRA	female	0.993	0.915
FRA	male	1.011	0.988

Table 2. The coefficient of determination R^2 for CLC and DLC. The estimates are made for log mortality measured in deviations from the mean and in deviations from mean and trend. The column "Det" reports the fit in a detrending regression with a constant and trend.

Country	Gender	CLC	DLC	Det	CLC	DLC
		De-meaned $\log m_{xt}$			De-trended $\log m_{xt}$	
USA	female	0.966	0.976	0.949	0.337	0.520
USA	male	0.951	0.970	0.915	0.421	0.646
USA	total	0.965	0.975	0.946	0.349	0.541
JPN	female	0.970	0.994	0.925	0.594	0.925
JPA	male	0.975	0.988	0.949	0.502	0.767
JPN	total	0.974	0.991	0.940	0.564	0.857
FRA	female	0.965	0.980	0.955	0.235	0.552
FRA	male	0.941	0.971	0.901	0.402	0.705
FRA	total	0.956	0.978	0.932	0.355	0.681

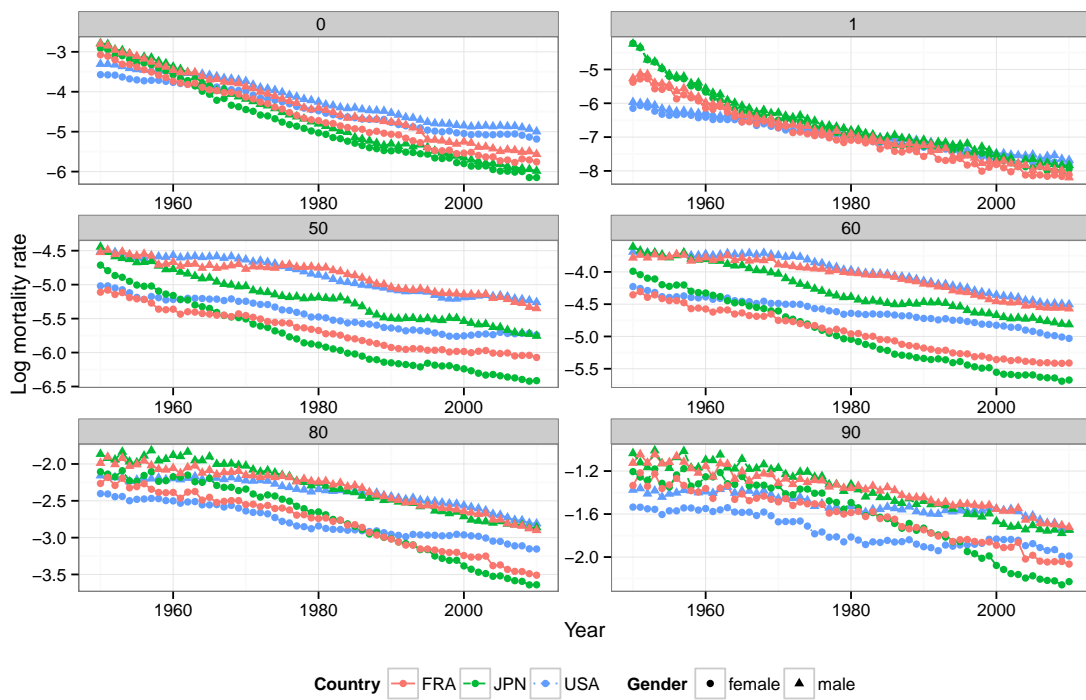


Figure 1: Log (all cause) mortality, men and women, for selected age groups $x=(0,1,50,60,80,90)$, for France, Japan, and USA.

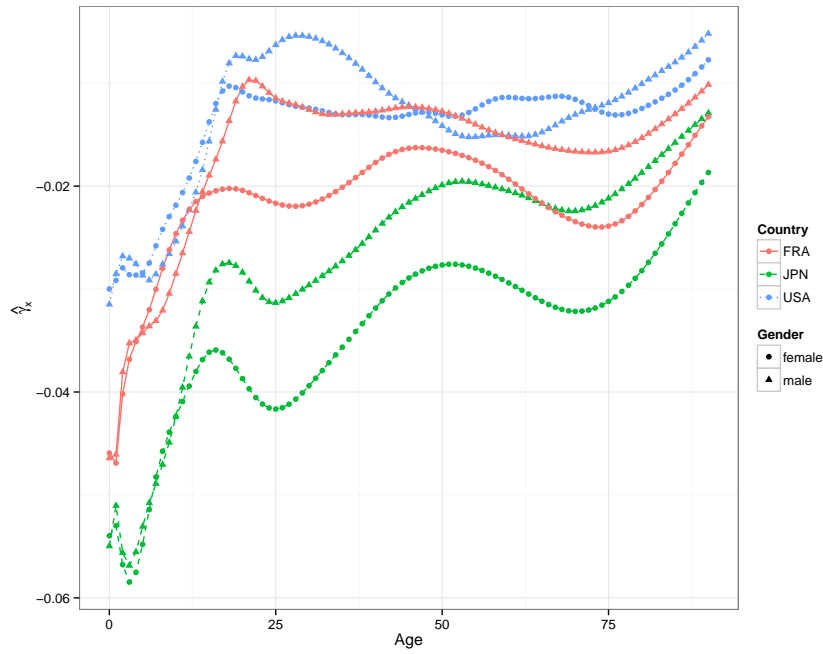


Figure 2: The trend slope from linear detrending of log mortality as a function of age.

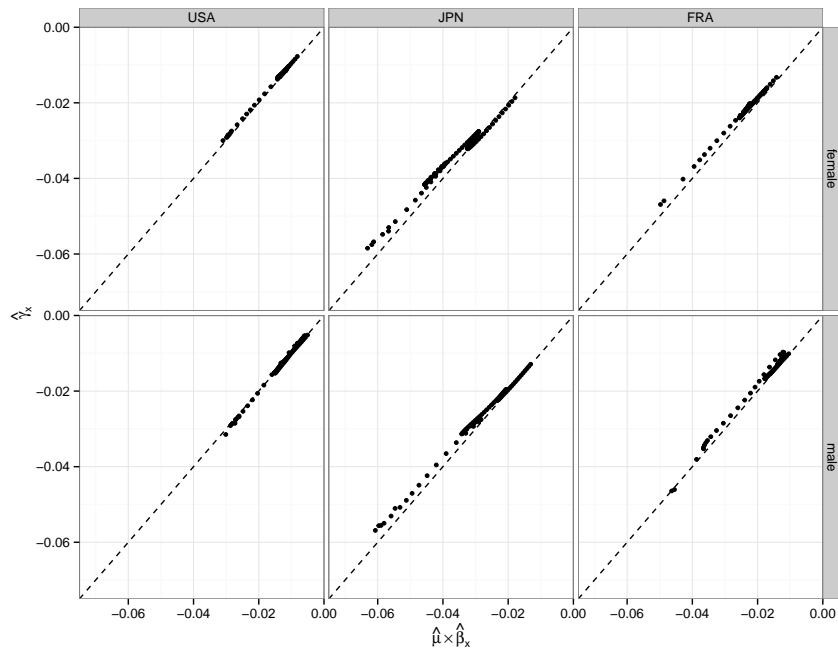


Figure 3: Cross plot of the trend slope estimate from the one step DLC model and the two-step CLC model.

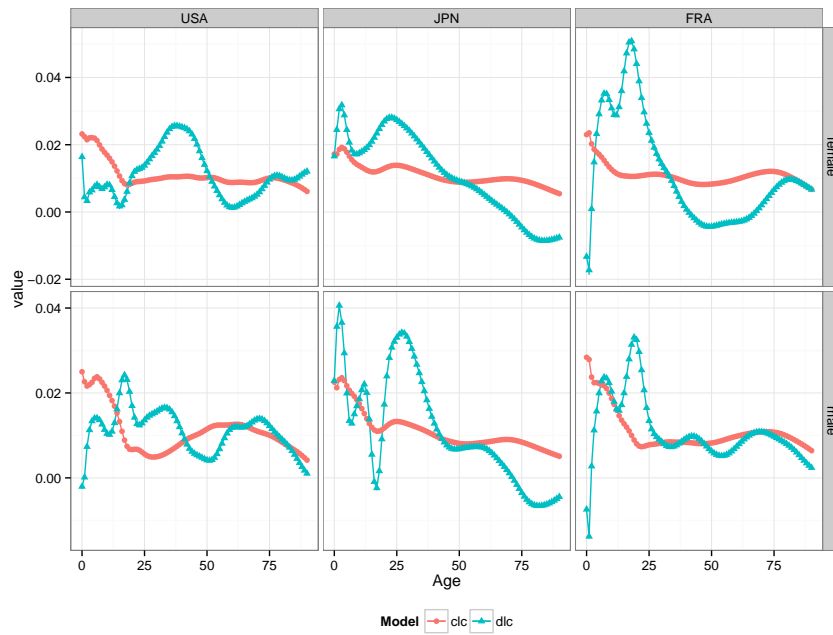


Figure 4: The estimate of β_x for the CLC model (red) and $\tilde{\beta}_x$ for the DLC model (blue).

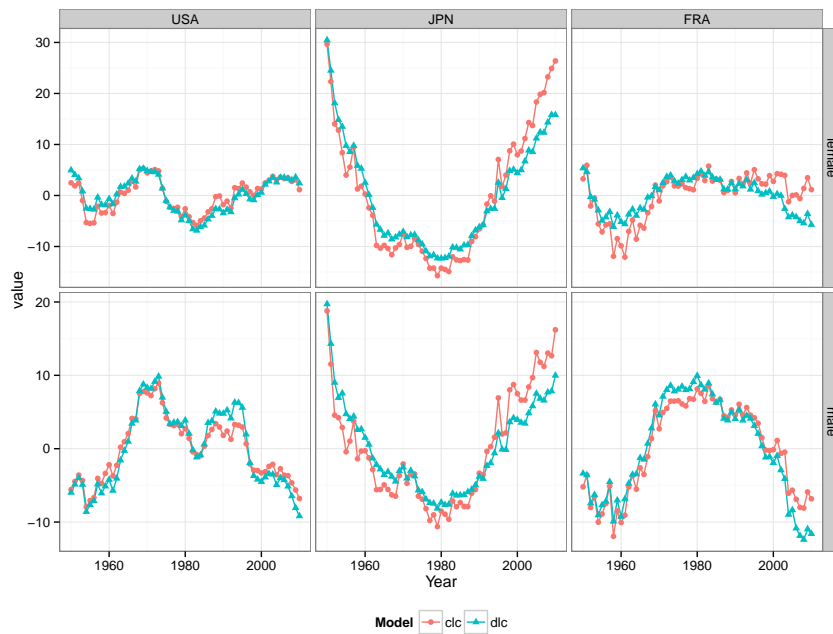


Figure 5: The estimate of the mortality index k_t of the CLC model corrected for the trend, $k_t - k_1 - \hat{\mu}(t - 1)$, (red) and \tilde{k}_t estimated from the DLC model (blue).

Research Papers 2013



- 2014-26: Markku Lanne, Jani Luoto and Henri Nyberg: Is the Quantity Theory of Money Useful in Forecasting U.S. Inflation?
- 2014-27: Massimiliano Caporin, Eduardo Rossi and Paolo Santucci de Magistris: Volatility jumps and their economic determinants
- 2014-28: Tom Engsted: Fama on bubbles
- 2014-29: Massimiliano Caporin, Eduardo Rossi and Paolo Santucci de Magistris: Chasing volatility - A persistent multiplicative error model with jumps
- 2014-30: Michael Creel and Dennis Kristensen: ABC of SV: Limited Information Likelihood Inference in Stochastic Volatility Jump-Diffusion Models
- 2014-31: Peter Christoffersen, Asger Lunde and Kasper V. Olesen: Factor Structure in Commodity Futures Return and Volatility
- 2014-32: Ulrich Hounyo: The wild tapered block bootstrap
- 2014-33: Massimiliano Caporin, Luca Corazzini and Michele Costola: Measuring the Behavioral Component of Financial Fluctuations: An Analysis Based on the S&P 500
- 2014-34: Morten Ørregaard Nielsen: Asymptotics for the conditional-sum-of-squares estimator in multivariate fractional time series models
- 2014-35: Ulrich Hounyo: Bootstrapping integrated covariance matrix estimators in noisy jump-diffusion models with non-synchronous trading
- 2014-36: Mehmet Caner and Anders Bredahl Kock: Asymptotically Honest Confidence Regions for High Dimensional
- 2014-37: Gustavo Fruet Dias and George Kapetanios: Forecasting Medium and Large Datasets with Vector Autoregressive Moving Average (VARMA) Models
- 2014-38: Søren Johansen: Times Series: Cointegration
- 2014-39: Søren Johansen and Bent Nielsen: Outlier detection algorithms for least squares time series regression
- 2014-40: Søren Johansen and Lukasz Gatarek: Optimal hedging with the cointegrated vector autoregressive model
- 2014-41: Laurent Callot and Johannes Tang Kristensen: Vector Autoregressions with Parsimoniously Time Varying Parameters and an Application to Monetary Policy
- 2014-42: Laurent A. F. Callot, Anders B. Kock and Marcelo C. Medeiros: Estimation and Forecasting of Large Realized Covariance Matrices and Portfolio Choice
- 2014-43: Paolo Santucci de Magistris and Federico Carlini: On the identification of fractionally cointegrated VAR models with the F(d) condition
- 2014-44: Laurent Callot, Niels Haldrup and Malene Kallestrup Lamb: Deterministic and stochastic trends in the Lee-Carter mortality model