

Quadratic Stochastic Intensity and Prospective Mortality Tables

C., GOURIEROUX ⁽¹⁾ and A., MONFORT⁽²⁾

October, 2007

¹CREST, CEPREMAP, and University of Toronto.

²CNAM and CREST.

The first author gratefully acknowledges financial support of NSERC Canada and of the chair AXA : "Large Risks in Insurance".

Quadratic Stochastic Intensity and Prospective Mortality Tables

ABSTRACT

We consider a quadratic stochastic intensity model with Gaussian autoregressive factor, derive explicit formulas for the predictive mortality tables and provide the recursive updating formulas are also provided. We also explain how to use appropriately the Kalman filter to estimate the parameters of the model and to approximate the values of the underlying factor. This methodology is applied to the French human mortality tables.

Keywords : Mortality Table, Life Insurance, Longevity Risk, Mortality Linked Securities (MLS), Affine Model, Quadratic Model, Kalman Filter.

JEL number :

1 Introduction

It is often difficult to distinguish the products and methodologies in Finance and Insurance. A typical example concerns life insurance. Whereas life insurance contracts on persons are standard products for insurance companies, life insurance contracts on corporates, called Credit Default Swap (CDS), are proposed by banks and traded on financial markets. Recently, new methodologies have been developed for credit risk, that is for CDS [see e.g. Lando (1998), Gouriéroux, Monfort, Polimenis (2006)], based on the so-called affine or Compound Autoregressive (CaR) factor models [see Duffie, Kan (1996), Duffie, Filipovic, Schachermayer (2003), Darolles, Gouriéroux, Jasiak (2006)]. These models can be used in a risk-neutral framework for pricing mortality linked securities³(MLS), but also in the historical framework to predict future mortality, that is, to construct prospective mortality tables. We focus on this latter application.

The possibility to transfer the credit risk methodology to life insurance has already been noted by several authors [see e.g. Dahl (2004), Biffis (2004), Luciano, Vigna (2005), Dahl, Moller (2006), Schrage (2006), Cairns, Blake, Dowd (2006), Chen, Cox (2007) and the so-called Survivor Credit Offert Rate (SCOR) market models]. These authors consider a stochastic intensity model, with factors satisfying the Duffie, Kan standard affine dynamics. However, this standard dynamics is too restrictive, especially to ensure positive mortality intensities [see Gouriéroux, Sufana (2006)]. This leads to a new class of affine (CaR) models, called the quadratic class, which contains as special case the Cox, Ingersoll, Ross process and its multivariate extension, called the Wishart process [Constantinides (1992), Leippold, Wu (2002), Ahn, Dittmar, Gallant (2002), Chen, Filipovic, Poor (2004), Gouriéroux (2006), (2007), Gouriéroux, Jasiak, Sufana (2007)].

The aim of this paper is to introduce a quadratic stochastic intensity model for prospective mortality tables. The model is written in discrete time, since the discrete time approach is more appropriate in this application, where the time unit is either the month, or the year.

The model is introduced in Section 2. We discuss the parameter restrictions to ensure positive mortality rates and derive the closed form expressions of the conditional survivor function. The derivation of the prospective mor-

³Following a suggestion by Blake, Burrows (2001), the first 25-years longevity bonds have been issued in 2004 by the European Investment Bank with structuror BnpParibas.

tality tables is discussed in Section 3, the updating formulas. The quadratic stochastic intensity model is convenient for estimation purpose. We explain in Section 4 how to apply an extended Kalman filter on an approximate linear state space model in order to get asymptotically efficient estimators of the parameters. This model is estimated on French human mortality data in Section 5. Section 6 concludes. Proofs are gathered in Appendices.

2 The model

2.1 Quadratic stochastic intensity

We consider a discrete time model and focus on the lifetime of an individual with age x at date t . The distribution of his/her (residual) lifetime is characterized by the sequence of survivor rates at future dates. The survivor rates are denoted by ⁴

$$P[Y > t + 1 | Y > t, \text{ age at } t = x] = \exp[-\lambda(x; t + 1)], \forall x, t \geq 0, \quad (2.1)$$

where Y is the lifetime variable and $\lambda(x; t + 1)$ is a (continuously compounded) spot mortality rate.

We assume that

$$\lambda(x; t + 1) = \alpha a^x + \beta b^x Z_{t+1} + \gamma c^x Z_{t+1}^2, \quad (2.2)$$

where the process (Z_t) is a Gaussian autoregressive process with zero-mean⁵ :

$$Z_t = \rho Z_{t-1} + \sigma \varepsilon_t, \quad (2.3)$$

⁴We use below the standard statistical notation instead of the actuarial notation $q_{x,t}$.

⁵If the factor process (Z_t) has a nonzero mean, we can write $Z_t = \mu + Z_t^*$, where Z_t^* is zero mean. Then,

$$\begin{aligned} \lambda(x, t + 1) &= (\alpha a^x + \beta b^x \mu + \gamma c^x \mu^2) \\ &+ (\beta b^x + 2\gamma c^x \mu) Z_{t+1}^* + \gamma c^x (Z_{t+1}^*)^2. \end{aligned}$$

We get a model in which the sensitivity coefficients are combinations of exponential functions. This type of model can be treated in a similar way (see the end of Section 2.2).

(ε_t) is a standard Gaussian white noise and the parameters are such that $a > 0, b > 0, c > 0, \sigma > 0$.

The survivor rate depends both on the current period by means of Z_{t+1} and on the current age of the individual by means of x . If we disregard for a moment the time variability of the factor, the survivor rate is $\exp[-\alpha a^x - \beta b^x z - \gamma c^x z^2]$. It involves a combination of exponential functions of x within the exponential.⁶ This survivor rate can feature a shape with hump when some parameters a, b, c are larger than 1 and the others smaller than 1. The quadratic specification is introduced to ensure survivor rates between 0 and 1. More precisely, we have the following property (see Appendix 1 i) :

Proposition 1 : The survivor rates are strictly between 0 and 1 for any age x and factor value z , if and only, if $(\beta = 0, \alpha \geq 0, \gamma \geq 0$ with at least α or γ strictly positive), or $(\gamma > 0, \beta \neq 0, ac \geq b^2, 4\alpha\gamma \geq \beta^2)$.

Since the survivor rates are strictly positive for any x , the lifetime variable Y can take any positive value. In particular, the model does not assume a maximal admissible age for the individual, such as 120 years, say. However, if for instance $\gamma > 0, c > 1$, the survivor probability will tend quickly to zero, for large x .

2.2 The conditional survivor function

The conditional survivor function of an individual with age x at date t is :

$$\begin{aligned} S(x; t, t+h) &= P[Y > t+h | Y > t, \text{ age at } t = x] \\ &= E_t \left\{ \exp \left[- \sum_{u=1}^h \lambda(x+u-1; t+u) \right] \right\} \\ &= E_t \exp[-\Lambda(x; t, t+h)], \end{aligned}$$

where Λ denotes the integrated spot intensity for period $(t, t+h)$, and E_t the conditional expectation given the current and lagged values of the factor

⁶This is an alternative to standard p -splines, which involve polynomial functions of x [see e.g. Currie (2007)].

process. The integrated intensity is defined for any horizon $h \geq 1$, but also in the limiting case $h = 0$, where $S(x; t, t) = 1$, and $\wedge(x, t; t) = 0, \forall x, t$.

Proposition 2 : The integrated intensity is given by :

$$\begin{aligned} & \wedge(x; t, t+h) \\ = & \alpha a^{x-1} \frac{a^{h+1} - a}{a-1} + \beta b^{x-1} Z_t \frac{(\rho b)^{h+1} - \rho b}{\rho b - 1} + \gamma c^{x-1} Z_t^2 \frac{(c\rho^2)^{h+1} - c\rho^2}{c\rho^2 - 1} \\ + & \mu(x; t, h)' \varepsilon_{t,t+h} + \varepsilon'_{t,t+h} \Omega(x, h) \varepsilon_{t,t+h}, \end{aligned}$$

$$\text{where } \varepsilon_{t,t+h} = (\varepsilon_{t+1}, \dots, \varepsilon_{t+h})'$$

$$\mu(x; t, h) = (\mu_j(x; t, h)),$$

$$\begin{aligned} \text{with } \mu_j(x; t, h) = & \sigma \beta b^{x-1} \rho^{-j} \frac{(b\rho)^{h-1} - (b\rho)^j}{b\rho - 1} \\ & + 2\sigma\gamma Z_t c^{x-1} \rho^{-j} \frac{(c\rho^2)^{h+1} - (c\rho^2)^j}{c\rho^2 - 1}, \end{aligned}$$

$$\Omega(x, h) = (\omega_{ij}(x, h)),$$

$$\text{with } \omega_{ij}(x, h) = \sigma^2 \gamma^2 c^{x-1} \frac{(c\rho^2)^{h+1} - (c\rho^2)^{\max(i,j)}}{c\rho^2 - 1} \rho^{-(i+j)}.$$

Proof : See Appendix 2.

The computation of the conditional survivor function

$$S(x; t, t+h) = E_t \exp[-\wedge(x; t, t+h)],$$

is based on the following Lemma :

Lemma 1 : Let us consider a standard Gaussian vector ε , then

$$\begin{aligned}
& E \exp(\mu' \varepsilon + \varepsilon' \Omega \varepsilon) \\
&= \frac{1}{\det(Id - 2\Omega)} \exp\left(\frac{1}{2} \mu' (Id - 2\Omega)^{-1} \mu\right).
\end{aligned}$$

By applying Lemma 1, we get the following Proposition :

Proposition 3 : The conditional survivor function is given by

$$\begin{aligned}
& S(x, t, t + h) \\
&= \exp\left[-\alpha a^{x-1} \frac{a^{h+1} - a}{a - 1} - \beta b^{x-1} Z_t \frac{(\rho b)^{h+1} - \rho b}{\rho b - 1} - \gamma c^{x-1} Z_t^2 \frac{(c\rho^2)^{h+1} - c\rho^2}{c\rho^2 - 1}\right. \\
&\quad \left. - \frac{1}{2} \log \det(Id + 2\Omega(x, h)) + \frac{1}{2} \mu(x; t, h)' [Id + 2\Omega(x, h)]^{-1} \mu(x; t, h)\right].
\end{aligned}$$

It is interesting to compare this result with the literature on affine models for credit risk. It is known that the conditional log-survivor function is an affine function ⁷ of the bivariate factor (Z_t, Z_t^2) . However, the coefficients of this affine representation have no explicit expressions in general. An explicit form is usually derived in term structure models for corporate bond in the special case of age independent coefficients (i.e. $a = b = c = 1$). Proposition 2 extends this result to age dependent coefficients of exponential form [see also Maghsoodi (1996) in the CIR framework].

The computations are valid for any choice of coefficients. When $\gamma = 0$, we get a model in which the spot mortality rate is an affine function of the Gaussian $AR(1)$ process (Z_t) . From Property 1 the positivity condition is not satisfied in this framework. Nevertheless, this model is simple and can be considered in a first step. We get the extension of Makeham model [Makeham (1860)] considered in Schrage (2006).

Finally, the model is easily extended to include more factors or more exponential terms in the sensitivity coefficients. For instance, we can assume :

$$\lambda(x; t) = \sum_{k=1}^K [\alpha_k a_k^x + \beta_k b_k^x Z_{k,t} + \gamma_k c_k^x Z_{k,t}^2],$$

⁷Indeed, the model is quadratic in (Z_t) , but affine in (Z_t, Z_t^2) .

where the $(Z_{k,t}), k = 1, \dots, K$ processes are independent such that $Z_{k,t} = \rho_k Z_{k,t-1} + \sigma_k \varepsilon_{k,t}$. By the independence property, the conditional mortality function has a closed form expression, where the terms indexed by k are summed up within the exponential in the expression of Proposition 2.

2.3 Longevity risk

The stationarity assumption of survivor rate in Section 2.1 is not compatible with the observed increasing longevity risk. Nonstationary feature can be introduced either by considering time dependent sensitivity coefficients of the factor, or by introducing a unit root factor process.

We will follow the first approach and consider mortality rates :

$$\lambda(x; t) = \alpha a^x \delta_a^t + \beta b^x \delta_b^t Z_t + \gamma c^x \delta_c^t Z_t^2, \quad (2.4)$$

where $\delta_a, \delta_b, \delta_c$ are positive real numbers. If $\gamma > 0, \beta \neq 0, \delta_b > 0, \delta_c > 0$, the condition for positive mortality rates becomes [see Appendix 1, ii)] :

$$ac \geq b^2, \delta_a \delta_c \geq \delta_b^2, 4\alpha\gamma\delta_a\delta_c \geq \beta^2\delta_b^2.$$

The advantage of an exponential trend specification [also introduced in Dahl, Moller (2006), Section 3.1.2] is the following : Let us consider the expectation of the "mortality intensity", equal to : $\alpha a^x \delta_a^t + \gamma c^x \delta_c^t E Z_t^2$, in the special case $\delta_a = 1, \delta_c < 1, a > 1, c > 1, \alpha > 0, \gamma > 0$. The "expected mortality intensity" is an increasing function of age x and a decreasing function of time (longevity effect). When t tends to infinity, the expected mortality intensity tends to the limiting function αa^x . Thus, contrary to a specification with a linear trend introduced by means of either the sensitivity coefficient, or a unit root factor [see e.g. Lee, Carter (1992), Renshaw, Haberman (2003)a,b, Delwarde (2004), or the model estimated in Shrager (2006)]⁸, the exponential trend specification ensures an expected mortality intensity, which is both positive and decreasing function of time.

Another advantage of the exponential trend specification is the possibility to capture components of the type a^x (when $\delta_a = 1$), δ_a^t (when $a = 1$) and δ_a^{t-x} (when $a = 1/\delta_a$), that are age, period and cohort effects, respectively [see e.g. Renshaw, Haberman (2003)a), Sunamoto (2005), Currie et al. (2006), Currie (2007)].

⁸When $\rho = 1$, the quadratic stochastic intensity model contains a linear trend since $E(Z_t^2)$ is proportional to t .

The extended specification (2.4) leads also to closed form expressions for the conditional survivor functions. [see Appendix 4].

Corollary 1 : For the quadratic stochastic intensity model with exponential trend, we get :

$$\begin{aligned}
& S^*(x; t, t+h) \\
= & \exp\left[-\alpha a^{x-1} \delta_a^t \frac{(a\delta_a)^{h+1} - (a\delta_a)}{a\delta_a - 1} - \beta b^{x-1} \delta_b^t Z_t \frac{(\rho b \delta_b)^{h+1} - \rho b \delta_b}{\rho b \delta_b - 1}\right. \\
& - \left. \gamma c^{x-1} \delta_c^t Z_t^2 \frac{(c\rho^2 \delta_c)^{h+1} - c\rho^2 \delta_c}{c\rho^2 \delta_c - 1}\right. \\
& \left. - \frac{1}{2} \log \det[Id + 2\Omega^*(x; t, h)] + \frac{1}{2} \mu^*(x; t, h)' [Id + 2\Omega^*(x, t, h)]^{-1} \mu^*(x; t, h)\right]
\end{aligned}$$

where :

$$\begin{aligned}
\mu_j^*(x; t, h) &= \sigma \beta b^{x-1} \delta_b^t \rho^{-j} \frac{(b\rho \delta_b)^{h+1} - (b\rho \delta_b)^j}{b\rho \delta_b - 1} \\
&+ 2\sigma \gamma Z_t c^{x-1} \delta_c^t \rho^{-j} \frac{(c\rho^2 \delta_c)^{h+1} - (c\rho^2 \delta_c)^j}{c\rho^2 \delta_c - 1}, \\
w_{ij}^*(x; t, h) &= \sigma^2 \gamma^2 c^{x-1} \delta_c^t \frac{(c\rho^2 \delta_c)^{h-1} - (c\rho^2 \delta_c)^{\max(i,j)}}{c\rho^2 \delta_c - 1} \rho^{-(i+j)}.
\end{aligned}$$

The factor model above involves 10 parameters, that are $\alpha, \beta, \gamma, a, b, c, \delta_a, \delta_b, \delta_c$ and ρ , which is much more than the standard models usually considered in the literature. This explains its flexibility.

3 Prediction

Let us now explain how the quadratic intensity model can be used for prediction purpose.

At a given date t , we have observed the past survivor (or mortality) rates, and this information can be used to calibrate the parameters and to get

approximations of the current and lagged factor values (see Section 4). We also have to predict the residual lifetime for the different classes of age. Two approaches can be followed that are the simulation and analytic approaches of prospective tables, respectively.

3.1 Simulation approach

This approach has been suggested for the computation of the reserves for a portfolio of CDS by the financial regulator [see e.g. Gouriou, Jasiak (2007), Chapter 11]. The idea is to approximate the distribution at date t of the residual lifetime by simulation as follows :

For each age x and replication s ,

- i) Draw a sequence of innovations $\varepsilon_{t+1}^s, \dots, \varepsilon_{t+u}^s$, in the standard normal;
- ii) Deduce the simulated factor values through the recursive formulas : $Z_{t+u}^s = \rho Z_{t+u-1}^s + \sigma \varepsilon_{t+u}^s, u = 1, \dots$, with initial condition $Z_t^s = Z_t$;
- iii) Deduce the simulated future spot intensities $\lambda^s(x + u - 1; t + u)$ by substituting Z_{t+u}^s to Z_{t+u} and $x + u - 1$ to x in formula (2.2).
- iv) Perform a succession of Bernoulli trials with the simulated survivor probabilities $\exp[-\lambda^s(x + u - 1; t + u)]$ to get the simulated residual lifetime Y^s ;
- v) Replicate $s = 1, \dots, S$; the sample distribution of Y^1, \dots, Y^S provides a good estimator of the distribution of the residual lifetime given the age at t , the factor value Z_t and the fact that the individual is still alive at t .

In particular, several summary statistics of the residual lifetime distribution can be deduced by averaging on the replications, such as the prospective mortality tables, or the expected residual lifetime.

3.2 Analytic approach

Direct computation of prospective survivor (or mortality) tables are generally performed in Insurance [see e.g. Denuit, Robert (2007)]. They rely on models providing closed form expressions of the conditional survivor function. The quadratic stochastic intensity model of Section 2 is a new example of such models. Given the information available at date t , the prospective survivor rate (also called forward survivor rate) is defined by :

$$\exp[-\lambda^P(x; t, t+h)] = \frac{S(x; t, t+h)}{S(x; t, t+h-1)}, \quad (3.1)$$

where x denotes the age at time t . For instance for the model without longevity, we directly deduce from Proposition 3 the following result :

Corollary 2 : In the quadratic stochastic intensity model (without longevity) the prospective table is given by

$$\begin{aligned} \lambda^P(x; t, t+h) &= \alpha a^{x+h-1} + \beta b^{x+h-1} \rho^h Z_t + \gamma c^{x+h-1} \rho^{2h} Z_t^2 \\ &+ \frac{1}{2} \log \det[Id + 2\Omega(x, h)] - \frac{1}{2} \log \det[Id + 2\Omega(x, h-1)] \\ &- \frac{1}{2} \mu(x; t, h)' [Id + 2\Omega(x, h)]^{-1} \mu(x; t, h) + \frac{1}{2} \mu(x; t, h-1)' [Id + 2\Omega(x, h-1)]^{-1} \mu(x; t, h-1) \end{aligned}$$

The prospective survivor rate $\exp[-\lambda^P(x; t, t+h)]$ can be written as :

$$\exp[-\lambda^P(x; t, t+h)] = \frac{E_t\{[S(x; t, t+h-1) \exp[-\lambda(x-h+1; t+h)]]\}}{S(x; t, t+h-1)}. \quad (3.2)$$

It is the expectation of the underlying spot survivor rate $\exp[\lambda(x+h-1; t+h)]$, with respect to a modified probability. This prediction interpretation can be seen in the formula of Corollary 2, where the first three terms of the decomposition can be written as:

$$\alpha a^{x+h-1} + \beta b^{x+h-1} E_t Z_{t+h} + \gamma c^{x+h-1} (E_t Z_{t+h})^2.$$

They involve the prediction of Z_{t+h} , whereas the remaining terms are correcting for risk included in the future factor values.

A major difference between affine models applied to term structure and to mortality is the following : whereas a real (or nominal) interest rate can take any sign, a mortality rate has to be nonnegative. For instance, a Gaussian affine model with $\gamma = 0$ [see e.g. Schrage (2006)] will provide meaningless negative prospective mortality rates for some environment Z_t .

3.3 The affine property and the updating formulas

It is known that the bivariate process (Z_t, Z_t^2) , where (Z_t) is a Gaussian vector autoregressive process, is CaR [see e.g. Darolles, Gourieroux, Jasiak (2006)].

This explains why the conditional survivor function and the prospective survivor rate are affine functions of Z_t, Z_t^2 for any x, h [note that $\mu(x; t, h)$ is affine in Z_t]. Therefore, we can write :

$$S(x; t, t + h) = \exp[a(x, h) + b(x, h)Z_t + c(x, h)Z_t^2]. \quad (3.3)$$

It is usual in affine models to emphasize the updating formulas of the predictive tables when t increases, or equivalently the recursive formulas for coefficients a, b, c . The updating step can avoid the use of the explicit formula in Corollary 1, which requires the inversion of a matrix of dimension h , where h can be large (up to 120 in our problem). The recursive formulas are obtained as follows :

By definition we have :

$$\begin{aligned} S(x; t, t + h) &= E_t \exp[-\wedge(x; t, t + h)] \\ &= E_t \{\exp[-\lambda(x; t + 1) - \wedge(x + 1; t + 1, t + h)]\} \\ &= E_t E_{t+1} \{\exp[-\lambda(x; t + 1) - \wedge(x + 1; t + 1, t + h)]\}, \\ &= E_t \{\exp[-\lambda(x; t + 1)] S(x + 1; t + 1, t + h)\}, \end{aligned}$$

by the law of iterated expectations.

We deduce that :

$$\begin{aligned} S(x; t, t + h) &= E_t \{\exp[-\alpha a^x - \beta b^x Z_{t+1} - \gamma c^x Z_{t+1}^2 \\ &\quad + a(x + 1, h - 1) + b(x + 1, h - 1)Z_{t+1} + c(x + 1, h - 1)Z_{t+1}^2]\}. \end{aligned}$$

The right hand side can be computed explicitly (see Appendix 3).

Proposition 4 : For the quadratic stochastic intensity model without longevity risk, the sensitivity coefficients satisfy the recursive relationship :

$$\begin{aligned}
c(x, h) &= \rho^2[-\gamma c^x + c(x+1, h-1)] \\
&+ 2\sigma^2 \rho^2 \frac{[-\gamma c^x + c(x+1, h-1)]^2}{1 - 2\sigma^2[-\gamma c^x + c(x+1, h-1)]^2}, \\
b(x, h) &= \rho[-\beta b^x + b(x+1, h-1)] \\
&+ 2\sigma^2 \rho \frac{[-\beta b^x + b(x+1, h-1)][-\gamma c^x + c(x+1, h-1)]}{1 - 2\sigma^2[-\gamma c^x + c(x+1, h-1)]}, \\
a(x, h) &= -\alpha a^x + a(x+1, h-1) + \frac{\sigma^2}{2} \frac{[-\beta b^x + b(x+1, h-1)]^2}{1 - 2\sigma^2(-\gamma c^x + c(x+1, h-1))} \\
&- \frac{1}{2} \log[1 - 2\sigma^2(-\gamma c^x + c(x+1, h-1))],
\end{aligned}$$

with initial conditions : $a(x, 0) = b(x, 0) = c(x, 0) = 0$.

For expository purpose, we have presented the recursive formulas for a model without longevity risk, but similar formulas can be deduced for model with longevity risk in which :

$$S(x, t, t+h) = \exp\{a(t; x, h) + b(t; x, h)Z_t + c(t, x, h)Z_t^2\},$$

[see Appendix 3]. In the case with longevity risk, the recursive formulas have to be applied for each cohort separately.

4 Estimation

The estimation procedure can be based on a multistep approach described in Section 5.2. The main steps of this approach are :

- i) the estimation of the marginal function $\bar{\lambda}(x) = \frac{1}{T} \sum_{t=1}^T \lambda(x, t)$;
- ii) the extraction of an unconstrained factor;
- iii) the separation of a deterministic trend component and of a zero-mean stationary factor.

Moreover, it is possible, to use the extended Kalman filter, to derive more efficient estimators.⁹

For expository purpose, the results are presented in a stationary framework, that is without longevity effect. If longevity is introduced, the approximate linear state space representation can still be used, but the asymptotic distribution of the estimator will be modified.

4.1 The observations

Let us consider different generations (cohorts) of individuals indexed by the birth year, t_0 , say. The number of individuals in generation t_0 is denoted by N_{t_0} and the lifetimes of the individuals are $Y_{i,t_0}, i = 1, \dots, N_{t_0}$, t_0 varying. We make the following assumptions :

Assumption 1 : There exists a common factor (Z_t).

Assumption 2 : The lifetime variables $Y_{i,t_0}, i = 1, \dots, N_{t_0}$, t_0 varying, are independent conditional on the factor path.

Assumption 3 : The lifetime variables satisfy the same quadratic stochastic intensity model (2.2) conditional on the factor path.

The above set of assumptions implies that the lifetime variables have different distributions and are dependent. Indeed,

- i) the lifetime distribution depends on the generation due to the cross-effects age \times date in the intensity;
- ii) the lifetime variables are dependent, since they depend on the unobservable common factor Z_t .

In practice, the raw observations of the lifetime variables are not available, but published mortality tables can be used. They provide the empirical counterpart of the survivor (mortality) rates for different years and ages. Let us denote these data by

$$\hat{\Pi}(x; t + 1), x = 0, 1, \dots, \bar{x}, t = 1, \dots, T. \quad (4.1)$$

⁹But a numerical instability may be encountered when switching regimes necessitate a severe reduction of the estimation period (as it is the case for the data used in Section 5).

If the sizes of the different generations are large, it is known that [see e.g. Anderson, Goodman (1957)] :

i) The variables $\hat{\Pi}(x; t + 1)$, $x = 0, \dots, \bar{x}$, $t = 1, \dots, T$ are asymptotically independent, conditional on (Z_t) .

ii) The variables are asymptotically Gaussian,¹⁰ conditional on (Z_t) :

$$\begin{aligned} & \sqrt{n_{x,t}}[\hat{\Pi}(x; t + 1) - \exp[-\lambda(x; t + 1)]] \\ & \xrightarrow{d} N[0, \exp[-\lambda(x; t + 1)](1 - \exp(-\lambda(x; t + 1)))], \end{aligned}$$

where $n_{x,t}$ is the number of individuals of age x alive at date t , that is, the size of the so-called Population-at-Risk (PaR).

4.2 Approximate linear state space representation

The asymptotic behavior can be used to linearize the model. The empirical estimator of $\lambda(x; t + 1)$ is defined by :

$$\hat{\lambda}(x; t + 1) = -\log \hat{\Pi}(x; t + 1). \quad (4.2)$$

By applying the delta-method, we deduce :

$$\sqrt{n_{x,t}}[\hat{\lambda}(x; t + 1) - \lambda(x, t + 1)] \xrightarrow{d} N[0, \exp(\lambda(x; t + 1)) - 1].$$

Thus, we can write the approximate state space model :

$$\begin{cases} \hat{\lambda}(x; t + 1) \simeq \alpha a^x + \beta b^x Z_{t+1} + \gamma c^x Z_{t+1}^2 + \eta(x; t + 1), x = 0, \dots, \bar{x}, \\ Z_{t+1} = \rho Z_t + \sigma \varepsilon_{t+1}, \end{cases} \quad (4.3)$$

where the error terms are Gaussian, independent, with zero mean and variance $V\varepsilon_t = 1$, $V[\eta(x; t + 1)] = \frac{\exp \lambda(x; t + 1) - 1}{n_{x,t}}$.

¹⁰We assume that the population size is sufficiently large to get the asymptotic Gaussian approximation. Otherwise, a Poisson approximation can be preferred [Brouhns, Denuit, Vermunt (2002)].

The state space model above involves a quadratic measurement equation. It can be transformed into a linear state space as follows. By taking the square of the transition equation, we get :

$$Z_{t+1}^2 = \sigma^2 + \rho^2 Z_t^2 + \varepsilon_{t+1}^*,$$

$$\text{where } \varepsilon_{t+1}^* = 2\sigma\rho Z_t \varepsilon_{t+1} + \sigma^2(\varepsilon_{t+1}^2 - 1).$$

Thus, the initial state space model can be rewritten as :

$$\begin{cases} \hat{\lambda}(x; t+1) &= \alpha a^x + \beta b^x Z_{t+1} + \gamma c^x Z_{t+1}^2 + \eta(x; t+1), x = 0, \dots, \bar{x}, \\ Z_{t+1} &= \rho Z_t + \sigma \varepsilon_{t+1}, \\ Z_{t+1}^2 &= \sigma^2 + \rho^2 Z_t^2 + \varepsilon_{t+1}^*, \end{cases} \quad (4.4)$$

where the bivariate error term of the transition system is no longer Gaussian, but is still a martingale difference sequence with :

$$V_t \begin{pmatrix} \sigma \varepsilon_{t+1} \\ \varepsilon_{t+1}^* \end{pmatrix} = \begin{pmatrix} \sigma^2 & 2\sigma^2 \rho Z_t \\ 2\sigma^2 \rho Z_t & 4\sigma^2 \rho^2 Z_t^2 + 2\sigma^4 \end{pmatrix}.$$

We get a linear state space model with a bivariate state variable $Z_{1,t} = Z_t, Z_{2,t} = Z_t^2$.

Then, we can apply a standard extended Kalman filter to linear state space model (4.4) in order to get the quasi-maximum likelihood estimator of the unknown parameters and the filtered factor value $\hat{Z}_{1,t} = \hat{Z}_t, \hat{Z}_{2,t} = (\hat{Z}_t^2)$. for $t = 1, \dots, T$. Such an approximate Kalman filter approach has already been applied¹¹ to stochastic migration models [Gagliardini, Gourieroux (2005), Feng, Gourieroux, Jasiak (2007)]. Under the stationarity assumption, that is without longevity effect, this approach provides consistent asymptotically Gaussian estimators, whenever $n_{x,t}$ is large for any x, t [see Gagliardini, Gourieroux, Monfort (2007)]. If longevity risk were introduced, this approach will still provide consistent estimators, but the asymptotic distribution and speed of convergence will be modified.

¹¹It has also been used in Lazar (2004), De Jong, Tickle (2006), Schrage (2006), without taking into account the heteroscedasticity of the error terms $\eta(x; t, t+1)$, that is, without correcting for either the number of surviving people, or the value of the survivor rate.

Finally, additional outputs of the Kalman filter are the predicted factor values $\hat{Z}_{1,t} = \hat{E}_t Z_{t+1}$, $\hat{Z}_{2,t} = \hat{E}_t(Z_{t+1}^2)$.

5 Application

5.1 Data and model

The database used in the application is the human mortality database (<http://www.mortality.org>) developed by the University of Berkeley. This base provides in particular the disaggregate mortality rates by gender, and the total mortality rates (male plus female) in France for ages x , $0 \leq x \leq 110$, and years t , $1899 \leq t \leq 2004$. We present in this section an application to total mortality rates; similar results disaggregated by gender are provided in Appendix 5. The data can be represented as a surface giving the mortality rate as a function of age and time (see Figure 1).

[Insert Figure 1 : Mortality Rates in France : Females + Males]

As expected the rate decreases with time for fixed age, and increases with age for sufficiently large age and fixed time.

We consider the estimation of a 1-factor quadratic intensity model of the type¹² :

$$\lambda(x, t) = \alpha_1 \exp(a_1 x + b_1 t) + \alpha_2 Z_t \exp(a_2 x + b_2 t) + \alpha_3 Z_t^2 \exp(a_3 x + b_3 t), \quad (5.1)$$

where (Z_t) is a zero-mean stationary process¹³. This 1-factor model can provide a good fit if it is applied to a part of the mortality table, but more factors would be needed for the analysis of the complete table. In the rest of the paper, we consider the mortality rates corresponding to age¹⁴ $50 \leq x \leq 90$, and year $1949 \leq t \leq 2004$.

¹²which is a special case of 2-factor affine model with (Z_t, Z_t^2) as factor.

¹³The change of parameter $a^x \delta_a^t \rightarrow \exp(a_1 x + b_1 t)$ is introduced to facilitate the numerical optimizations.

¹⁴For the same reason, Cairns, Blake, Dowd (2006) have considered the limited range $60 \leq x \leq 90$, when estimating a 2-factor stochastic intensity model, and White (2002) the range $55 \leq x \leq 96$. Moreover, the pattern of mortality rate changes around the mid-twentieth century [Lee, Miller (2001)].

Indeed, the behaviour of $\lambda(x, t)$ as a function of x , for a given t , is different for small values of age (including the infant mortality $x = 0$) and for the medium or large values of x (see Figure 1). Moreover, by considering the average mortality rates per year $\bar{\lambda}(t) = \frac{1}{\bar{x}} \sum_{x=0}^{\bar{x}} \lambda(x, t)$ (see Figure 2), we note a change of regime after the second world war. This explains the range of years which has been retained.

[Insert Figure 2 : Average Mortality Rate per Year : Females + Males]

Even on this subset of mortality rates, the direct joint estimation of parameters $\alpha_i, a_i, b_i, i = 1, \dots, 3$ is not an easy task. Some reasons are the following :

i) For a given age x , it is difficult to disentangle the deterministic and stochastic time components, in particular since the number of observation dates is rather small.

ii) The stationarity conditions for the factor process are difficult to impose a priori in the estimation procedure.

iii) The number of unknowns is large, in particular due to the unobservable factors values.

Before applying directly a joint estimation approach such as the approximate Kalman filter, it is preferable to develop a simpler, more robust multistep estimation method, whose output can be used later on as starting value in the approximated Kalman filter approach described in Section 4.

5.2 A multistep approach

Let us now describe such a multistep approach.

Step 1 : Consistent estimations of a_1, a_2, a_3 .

Let us consider the mortality rate averaged by age :

$$\bar{\lambda}(x) = \frac{1}{T} \sum_{t=1}^T \lambda(x, t). \quad (5.2)$$

Under specification (5.1), we get a function of age of the type :

$$\bar{\lambda}(x) = \tilde{\alpha}_1 \exp(a_1 x) + \tilde{\alpha}_2 \exp(a_2 x) + \tilde{\alpha}_3 \exp(a_3 x). \quad (5.3)$$

The specification (5.3) can be fitted by nonlinear least squares to the observed averaged mortality rates. For the considered range of age, an excellent fit is obtained with a function of the form $\alpha^* + \beta^* \exp(a^* x)$, with $a^* = 1.03 \cdot 10^{-2}$. The observed and fitted averaged values are reported on Figure 3.

[Insert Figure 3 : Average Mortality Rates per Age : Observed and Fitted].

Step 2 : 2-step determination of unconstrained factor values

In the second step, we are interested in the time factor without disentangling the deterministic and stochastic components. Among the 1-factor specifications providing an aggregate mortality rate of the type $\alpha^* + \beta^* \exp(a^* x)$, the one providing the best fit is :

$$\alpha + F_t + \beta \exp(a^* x) F_t^2.$$

The 2-step approximations of the unconstrained factor values are derived by minimizing :

$$\Sigma_x \Sigma_t [\hat{\lambda}(x, t) - \alpha - F_t - \beta \exp(a^* x) F_t^2]^2, \quad (5.4)$$

where $a^* = 1.03 \cdot 10^{-2}$. This optimization is performed with respect to α, β and F_t, t varying.

Step 3 : Global determination of the unconstrained factor

From a theoretical point of view, it is preferable to perform a joint minimization of :

$$\min_{a, \alpha, \beta, F_t} \Sigma_x \Sigma_t [\lambda(x, t) - \alpha - F_t - \beta \exp(ax) F_t^2]^2, \quad (5.5)$$

with respect to α, β, F_t, t varying, and a . However, the optimization could be numerically difficult, if the starting values in the minimization algorithm are not well-chosen. The aim of steps 1 and 2 was to provide such starting values, especially reasonable ones for the unobservable factor path. The approximated factor values $\hat{F}_t, t = 1, \dots, T$, derived from the optimization (5.5),

are provided in Figure 4. We observe the decreasing trend corresponding to the longevity effect.

[Insert Figure 4 :Extracted Factor : Females + Males]

Similar patterns are obtained for the factors derived from disaggregation by gender (see Figures 4.m-4.f).

Step 4 : Disentangling the deterministic and stochastic time components

Once the unconstrained factor value have been approximated in a reasonable way, we can consider the decomposition between deterministic and stochastic components. This step is the most important one in the sense that a small error in this decomposition can have severe implications for long run forecast of mortality rates.

To get a specification of the type (5.1), we can adjust \hat{F}_t to an exponential trend and get :

$$\hat{F}_t = \gamma_0 + \gamma_1 \exp(\gamma_2 t) + Z_t, \quad (5.6)$$

where (Z_t) is a zero-mean stationary process. A nonlinear least squares adjustment provides residuals : $\hat{Z}_t = \hat{F}_t - \hat{\gamma}_0 - \hat{\gamma}_1 \exp(\hat{\gamma}_2 t)$ (see Figure 5), which are clearly nonstationary with different behaviours before and after the early eighties. This regime change has also been detected on UK mortality data by Cairns et al. (2006).

[Insert Figure 5 : Factor Residuals : Females + Males]

This regime switching is also observed at the disaggregate level (see Figures 5.m-5.l) and is especially clear for male.

Taking into account this regime change leads to a readjustment of the exponential trend on the last period 1982-2004, only. The estimated values are (with $t = 0$ in 1982) :

$$\begin{aligned} \hat{\gamma}_0 &= 1.21 \cdot 10^{-2}, & \hat{\gamma}_1 &= 6.88 \cdot 10^{-3}, & \hat{\gamma}_2 &= -4.4 \cdot 10^{-2} \\ &(13.4) & &(16.4) & &(-3.1) \end{aligned}$$

$$\begin{aligned} \hat{\alpha} &= -1.53 \cdot 10^{-2}, & \hat{\beta} &= 8.18 \cdot 10^{-2} & \text{and} & \hat{a} &= 1.02 \cdot 10^{-2}; \\ &(-14.6) & &(5.8) & & &(5.35) \end{aligned}$$

t -ratios are given in parentheses.

Finally, the associated factor residuals can be fitted to a zero-mean AR(1) process :

$$Z_{t+1} = \rho Z_t + \sigma \varepsilon_{t+1}, E\varepsilon_{t+1} = 0, V\varepsilon_{t+1} = 1. \quad (5.7)$$

We get : $\hat{\rho} = -0.32, \hat{\sigma}^2 = 2.5 \cdot 10^{-8}$.
(-1.37)

To summarize the estimated specification of the mortality rate is of the form,

$$\lambda(x, t) = \alpha + \gamma_0 + \gamma_1 \exp(\gamma_2 t) + Z_t + \beta \exp(ax) [\gamma_0 + \gamma_1 \exp(\gamma_2 t) + Z_t]^2, \quad (5.8)$$

where (Z_t) is the AR(1) process (5.7) and the parameters are replaced by their estimations.

It is interesting to note that the autoregressive coefficient is estimated and not fixed a priori, contrary to what is done in the basic models (see e.g. Lee (2000), Yang (2001)], in which the introduced factors are either random walk, $\rho = 1$, or white noise, $\rho = 0$.

5.3 Predictions

Standard output of dynamic state space model are the future predictions of the variable of interest $\lambda(x, t)$. These predictions are provided in Figure 6 for $x = 50, 60, 70, 80, 90$, up to $t = 2050$. This horizon is sufficiently large to be compatible with the residual maturities 20, 25, 30 years of the longevity bonds currently issued. These predictions have a smooth pattern, which is a consequence of the selected model and of the interest in longevity only. Indeed, we were not interested in morbidity and in the future occurrence of some epidemy [see e.g. Chen, Scott (2007)].

[Insert Figure 6 : Historical Data and Prediction of Mortality Rates]

However, these predictions are performed under the historical distribution, and, as noted in Section 3.2, they do not correspond to the prospective tables, which are expectations under a modified probability. In a second step, we have applied the simulation approach of Section 3.1 to several generations and derive the distribution of residual lifetime. Moreover, we have considered

successively the global populations, the females only and the males only (see Appendix 5).

[Insert Figure 7 : Distribution of Residual Lifetime, Cohort 50-year old in 2004]

[Insert Figure 8 : Distribution of Residual Lifetime, Cohort 65-year old in 2004]

[Insert Figure 9 : Distribution of Residual Lifetime, Cohort 80-year old in 2004]

We observe skewed distributions, with a skewness depending on the cohort. They are left skewed for cohort 50-year old, right skewed for cohort 80-year old and almost symmetric for cohort 65-year old.

Then, we can compute expected residual lifetimes (see Table 1)

Table 1 : Expected residual lifetime (in years)

cohort	Total	Females	Males
50-year in 2004	33.4	36.2	31.2
65-year in 2004	19.9	21.7	17.9
80-year in 2004	9.2	10.0	7.9

Even, if data on expected residual lifetimes are regularly diffused, these summary statistics are not completely relevant due to the high skewness of the distributions.

It is also possible to compute prospective (that is forward) mortality tables. For instance, Figure 10 displays prospective mortality rates for the cohort with age 80-year in 2004, for the total population, the females only and the males only.

[Insert Figure 10 : Prospective of Forward Mortality Rate, Cohort 80-year old in 2004 : Total, Males, Females].

6 Concluding Remarks

In this paper, we have introduced a quadratic stochastic intensity model, which ensures nonnegative mortality rates, accounts for cross-effects of age and date, is flexible, allows for closed form expressions of the prospective mortality table and provides a convenient modelling of longevity risk. This model is simple to implement, since the standard Kalman filter can be applied on an approximated linear state space model without loss of information, after an appropriate choice of starting values for parameters and factors.

All the results of the paper have been presented with the perspective of deriving prospective mortality tables. For this reason they have been written in terms of historical probability. But, clearly the explicit prediction formulas or the updating equations can also be applied to a risk-neutral quadratic stochastic intensity model, that is, for mortality pricing purpose. The existence of explicit formulas means that explicit pricing formulas will be available for a large set of mortality linked securities (MLS).

Appendix 1

Positivity conditions

i) Model without longevity risk

The positivity condition for λ is :

$$\alpha a^x + \beta b^x z + \gamma c^x z^2 > 0, \forall x, z.$$

This condition is equivalent to :

$$(\gamma > 0, \beta^2 b^{2x} - 4\alpha\gamma(ac)^x \leq 0, \forall x \geq 0) \text{ or } (\gamma = \beta = 0, \alpha > 0).$$

Since the inequalities $\beta^2 b^{2x} - 4\alpha\gamma(ac)^x \leq 0, \forall x \geq 0$, are equivalent to the inequalities $4\alpha\gamma(ac/b^2)^x \geq \beta^2, \forall x$, we get the result in Proposition 1.

ii) Model with longevity risk

In this case the condition on the discriminant becomes :

$$ac > b^2 \text{ and } 4\alpha\gamma(\delta_a\delta_c)^{t+1} \geq \beta^2(\delta_b^2)^{t+1}, \forall t \geq 0.$$

It is equivalent to :

$$ac > b^2, \delta_a\delta_c \geq \delta_b^2, 4\alpha\gamma\delta_a\delta_c \geq \beta^2\delta_b^2,$$

whenever $\beta \neq 0, \delta_b > 0$.

Appendix 2

Integrated Intensity

The future factor values can be written in terms of the current factor value and future innovations as :

$$Z_{t+u} = \rho^u Z_t + \sigma \sum_{j=1}^u \rho^{u-j} \varepsilon_{t+j}.$$

We deduce the expression of the spot intensity :

$$\begin{aligned} \lambda(x+u-1; t+u) &= \alpha a^{x+u-1} + \beta b^{x-1} (b\rho)^u Z_t + \gamma c^{x-1} (c\rho^2)^u Z_t^2 \\ &+ \sigma \beta b^{x-1} (b\rho)^u \sum_{j=1}^u \rho^{-j} \varepsilon_{t+j} \\ &+ 2\sigma \gamma Z_t c^{x-1} (c\rho^2)^u \sum_{j=1}^u \rho^{-j} \varepsilon_{t+j} \\ &+ \sigma^2 \gamma^2 c^{x-1} (c\rho^2)^u \left(\sum_{j=1}^u \rho^{-j} \varepsilon_{t+j} \right)^2. \end{aligned}$$

The integrated intensity is given by :

$$\begin{aligned}
\wedge(x; t, t+h) &= \sum_{u=1}^h \lambda(x+u-1; t+u) \\
&= \alpha a^{x-1} \sum_{u=1}^h a^u + \beta b^{x-1} Z_t \sum_{u=1}^h (b\rho)^u + \gamma c^{x-1} Z_t^2 \sum_{u=1}^h (c\rho^2)^u \\
&+ \sigma \beta b^{x-1} \sum_{j=1}^h \left\{ \left[\sum_{u=j}^h (b\rho)^u \right] \rho^{-j} \varepsilon_{t+j} \right\} \\
&+ 2\sigma \gamma Z_t c^{x-1} \sum_{j=1}^h \left\{ \left[\sum_{u=j}^h (c\rho^2)^u \right] \rho^{-j} \varepsilon_{t+j} \right\} \\
&+ \sigma^2 \gamma^2 c^{x-1} \sum_{i=1}^h \sum_{j=1}^h \left\{ \left[\sum_{u=\max(i,j)}^h (c\rho^2)^u \right] \rho^{-(i+j)} \varepsilon_{t+i} \varepsilon_{t+j} \right\} \\
&= \alpha a^{x-1} \frac{a^{h+1} - a}{a-1} + \beta b^{x-1} Z_t \frac{(b\rho)^{h+1} - b\rho}{b\rho-1} + \gamma c^{x-1} Z_t^2 \frac{(c\rho^2)^{h+1} - c\rho^2}{c\rho^2-1} \\
&+ \sigma \beta b^{x-1} \sum_{j=1}^h \left[\frac{(b\rho)^{h+1} - (b\rho)^j}{b\rho-1} \rho^{-j} \varepsilon_{t+j} \right] \\
&+ 2\sigma \gamma Z_t c^{x-1} \sum_{j=1}^h \left[\frac{(c\rho^2)^{h+1} - (c\rho^2)^j}{c\rho^2-1} \rho^{-j} \varepsilon_{t+j} \right] \\
&+ \sigma^2 \gamma^2 c^{x-1} \sum_{i=1}^h \sum_{j=1}^h \left[\frac{(c\rho^2)^{h+1} - (c\rho^2)^{\max(i,j)}}{c\rho^2-1} \rho^{-(i+j)} \varepsilon_{t+i} \varepsilon_{t+j} \right].
\end{aligned}$$

Appendix 3

Recursive formulas

Let us consider the general framework including longevity risk. In this case, the survival function can be written as :

$$S(x; t, t + h) = \exp\{a(t; x, h) + b(t; x, h)Z_t + c(t; x, h)Z_t^2\}.$$

We have :

$$\begin{aligned} & S(x; t, t + h) \\ = & E_t \left\{ \exp[-\alpha a^x \delta_a^{t+1} + a(t + 1, x + 1, h - 1) + [-\beta b^x \delta_b^{t+1} + b(t + 1; x + 1, h - 1)]] Z_{t+1} \right. \\ & \left. + [-\gamma c^x \delta_c^{t+1} + c(t + 1; x + 1, h - 1)] Z_{t+1}^2 \right\} \\ = & \exp\{-\alpha a^x \delta_a^{t+1} + a(t + 1; x + 1, h - 1) + [-\beta b^x \delta_b^{t+1} + b(t + 1; x + 1, h - 1)] \rho Z_t \\ & + [-\gamma c^x \delta_c^{t+1} + c(t + 1; x + 1, h - 1)] \rho^2 Z_t^2\} \\ & E_t \left\{ \exp[\sigma[-\beta b^x \delta_b^{t+1} + b(t + 1; x + 1, h - 1) \right. \\ & \left. + 2\rho Z_t[-\gamma c^x \delta_c^{t+1} + c(t + 1; x + 1, h - 1)] \varepsilon_{t+1} \right. \\ & \left. + \sigma^2[-\gamma c^x \delta_c^{t+1} + c(t + 1; x + 1, h - 1)] \varepsilon_{t+1}^2] \right\} \\ = & \exp\{-\alpha a^x \delta_a^{t+1} + a(t + 1; x + 1, h - 1) + [-\beta b^x \delta_b^{t+1} + b(t + 1; x + 1, h - 1)] \rho Z_t \\ & + [-\gamma c^x \delta_c^{t+1} + c(t + 1; x + 1, h - 1)] \rho^2 Z_t^2\} \\ & + \frac{\sigma^2 (-\beta b^x \delta_b^{t+1} + b(t + 1; x + 1, h - 1) + 2\rho Z_t[-\gamma c^x \delta_c^{t+1} + c(t + 1; x + 1, h - 1)])}{2 (1 - 2\sigma^2[-\gamma c^x \delta_c^{t+1} + c(t + 1, x + 1, h - 1)]^2)} \\ & - \frac{1}{2} \log[1 - 2\sigma^2(-\gamma c^x \delta_c^{t+1} + c(t + 1; x + 1, h - 1))^2]. \end{aligned}$$

The recursive formulas are deduced by identifying the coefficient of 1, Z_t , Z_t^2 .

Appendix 4

Nonstationary case

We get :

$$\begin{aligned}\lambda(x+u-1, t+u) &= \alpha a^{x-1} \delta_a^t (a \delta_a)^u + \beta b^{x-1} \delta_b^t (b \rho \delta_b)^u Z_t + \gamma c^{x-1} \delta_c^t (c \rho^2 \delta_c)^u Z_t^2 \\ &+ \sigma \beta b^{x-1} \delta_b^t (b \rho \delta_b)^u \sum_{j=1}^u \rho^{-j} \varepsilon_{t+j} \\ &+ 2\sigma \gamma Z_t c^{x-1} \delta_c^t (c \rho^2 \delta_c)^u \sum_{j=1}^u \rho^{-j} \varepsilon_{t+j} \\ &+ \sigma^2 \gamma^2 c^{x-1} \delta_c^t (c \rho^2 \delta_c)^u \left(\sum_{j=1}^u \rho^{-j} \varepsilon_{t+j} \right)^2.\end{aligned}$$

This formula is similar to the formula in Appendix 2 after replacing $a \rightarrow a \delta_a, b \rightarrow b \delta_b, c \rightarrow c \delta_c, a^{x-1} \rightarrow a^{x-1} \delta_a^t, b^{x-1} \rightarrow b^{x-1} \delta_b^t, c^{x-1} \rightarrow c^{x-1} \delta_c^t$.

Appendix 5

Mortality Analysis by Gender

This appendix provides the estimation results when the gender is taken into account. The models are similar to the model considered in Section 5 at the aggregate level.

i) Male Mortality

The estimated coefficients are :

$$\hat{\alpha} = -0.022, \quad \hat{\beta} = 0.104, \quad \hat{a} = 0.0093, \\ (-9.3) \qquad (3.5) \qquad (30.1)$$

$$\hat{\gamma}_0 = 0.015, \quad \hat{\gamma}_1 = 0.011, \quad \hat{\gamma}_2 = -0.025, \\ (4.8) \qquad (4.2) \qquad (-2.1)$$

$$\hat{\rho} = -0.08, \quad \hat{\sigma}^2 = 3.7 \cdot 10^{-8} \\ (-0.36)$$

The analogues of Figures 1-6 are provided below and denoted Figures 1.m-6.m.

[Insert Figure 1.m : Mortality Rate in France : Males]

[Insert Figure 2.m : Average Mortality Rate per Year : Males]

[Insert Figure 3.m : Average Mortality Rate per Age (Males) : Observed-Fitted]

[Insert Figure 4.m : Extracted Factor, Global Method (Males)]

[Insert Figure 5.m : Factor Residuals (Males)]

[Insert Figure 6.m : Data and Predictions of Mortality Rates (Males)]

ii) Female Mortality

The estimated coefficients are :

$$\hat{\alpha} = -0.020, \quad \hat{\beta} = 0.026, \quad \hat{a} = 0.111,$$

$$(-13.8) \quad (3.4) \quad (34.0)$$

$$\hat{\gamma}_0 = 0.0144, \quad \hat{\gamma}_1 = 0.0084, \quad \hat{\gamma}_2 = -0.052,$$

$$(18) \quad (24) \quad (-3.6)$$

$$\hat{\rho} = -0.38, \quad \hat{\sigma}^2 = 3.8 \cdot 10^{-8}$$

$$(-1.7)$$

The analogues of Figure 1-6 are provided below and denoted Figures 1.f-6.f

[Insert Figure 1.f : Mortality Rate in France : Females]

[Insert Figure 2.f : Average Mortality Rate per Year : Females]

[Insert Figure 3.f : Average Mortality Rate per Age (Females) : Observed-Fitted]

[Insert Figure 4.f : Extracted Factor, Global Method (Females)]

[Insert Figure 5.f : Factor Residuals (Females)]

[Insert Figure 6.f : Data and Predictions of Mortality Rates (Females)]

REFERENCES

Ahn, D., Dittmar, R., and A., Gallant (2002) : "Quadratic Term Structure Models : Theory and Evidence", *Review of Financial Studies*, 15, 243-255.

Alho, J. (2000) : "Discussion of Lee (2000)", *North-American Actuarial Journal*, 4, 91-93.

Anderson, T., and L., Goodman (1957) : "Statistical Inference About Markov Chains", *Journal of the American Mathematical Society*, 28, 89.

Biffis, E. (2004) : "Affine Processes for Dynamic Mortality and Actuarial Valuations", forthcoming, *Insurance : Mathematics and Economics*.

Biffis, E., and P., Millosovich (2004) : "The Fair Value of Guaranteed Annuity Options", DP Bocconi University.

Blake, D., and W., Burrows (2001) : "Survivor Bonds : Helping to Hedge Mortality Risk", *Journal of Risk and Insurance*, 68, 339-348.

Blake, D., Cairns, A., and K., Dowd (2006) : "Living with Mortality : Longevity Bonds and Other Mortality Linked Securities", forthcoming *British Actuarial Journal*.

Booth, H., De Jong, P., Hyndman, R., and L., Tickle (2006) : "Lee-Carter Mortality Forecasting : A Multi-Country Comparison of Variants and Extension", *Demographic Research*, 15, 289-310.

Brouhns, N., Denuit, M., and J., Vermunt (2002) : "A Poisson log-bilinear Regression Approach to the Construction of Projected Lifetables", *Insurance : Mathematics and Economics*, 31, 373-393.

Cairns, A., Blake , D., and K., Dowd (2004) : "Pricing Death : Frameworks for the Valuation and Securitization of Mortality Risk", DP Actuarial Mathematics and Statistics, Heriot-Watt University.

Cairns, A., Blake, D. and K., Dowd (2006) : "A Two Factor Model for Stochastic Mortality with Parameter Uncertainty : Theory and Calibration",

forthcoming Journal of Risk and Insurance.

Cairns, A., Blake, D., Dowd, K., Coughlan, G., Epstein, D., Ong, A. and I., Balevich (2007) : "A Quantitative Comparison of Stochastic Mortality Models Using Data from England & Wales and the United States", Heriot-Watt University DP.

Chen, H., and S., Cox (2007) : "Modeling Mortality with Jumps : Transitory Effects and Pricing Implication to Mortality Securitization", DP Georgia State University.

Chen, L., Filipovic, D., and V., Poor (2004) : "Quadratic Term Structure Models for Riskfree and Defaultable Rates", *Mathematical Finance*, 14, 515-536.

Constantinides, G. (1992) : "A Theory of the Nominal Term Structure of Interest Rates", *Review of Financial Studies*, 5, 531-552.

Cox, S., and Y., Lin (2005) : "Securitization of Mortality Risks in Life Annuities", *Journal of Risk and Insurance*, 72, 227-252.

Currie, I. (2007) : "Smoothing and Forecasting Mortality Rates with p-Splines", DP Heriot Watt University.

Currie, I., Durban, M., and P., Eilers (2004) : "Smoothing and Forecasting Mortality Tables", *Statistical Modelling*, 4, 279-298.

Currie, I., Kirkby, J., and S., Richard (2006) : "The Importance of Year of Birth in Two-Dimensional Mortality Rate", *British Actuarial Journal*, 12, 5-61.

Dahl, M. (2004) : "Stochastic Mortality in Life Insurance : Market Reserves and Mortality Linked Insurance Contracts", *Insurance : Mathematics and Economics*, 35, 113-136.

Dahl, M., and T., Moller (2004) : "Valuation and Hedging of Life Insurance Liabilities with Systematic Mortality Risk", *Insurance : Mathematics and Economics*, 39, 193-217.

Darolles, S., Gouriéroux, C., and J., Jasiak (2006) : "Structural Laplace

Transform and Compound Autoregressive Models”, *Journal of Time Series Analysis*, 27, 477-503.

De Jong, F., (2000) : ”Time Series and Cross-Section Information in Affine Term Structure Models”, *Journal of Business and Economic Statistics*, 18, 300-314.

De Jong, F. and L., Tickle (2006) : ”Extending Lee-Carter Mortality Forecasting”, *Mathematical Population Studies*, 13, 1-18.

Delwarde, A. (2004) : ”Modèle Log-bilinéaire pour l’élaboration de tables de mortalité prospectives”, PhD Thesis, Louvain University.

Denuit, M., and C., Robert (2007) : ”Actuariat des assurances des personnes”, *Economica*.

Duffee, G., and R., Stanton (2004) : ”Estimation of Dynamic Term Structure Models”,

Duffie, D., Filipovic, D., and W., Schachermayer (2003) :” Affine Processes and Application in Finance”, *Annals of Applied Probability*, 13, 984-1053.

Duffie, D., and R., Kan (1996) : ”A Yield Factor Model of Interest Rates”, *Mathematical Finance*, 6, 379-406.

Feng, D., Gouriéroux, C., and J., Jasiak (2007) : ”The Ordered Qualitative Model for Credit Rating Transitions”, forthcoming *Journal of Empirical Finance*.

Gagliardini, P., and Gouriéroux, C. (2005) : ”Stochastic Migration Models”, *Journal of Financial Econometrics*, 3, 188-226.

Gagliardini, P., Gouriéroux, C., and A., Monfort (2007) : ”Error in Factor Model”, CREST DP.

Gallop, A., Pasdika, U., and J., Wolff (2005) : ”Increased Longevity and the Challenge of Determining q_x at Extreme Ages”, Presented at ”Living to 100 Symposium”, Orlando.

Gompertz, B. (1825) : ”On the Nature of the Function Expressive of the

Law of Human Mortality”, *Philosophical Transactions*, 27, 513-585.

Gourieroux, C. (2006) : ”Continuous Time Wishart Process for Stochastic Risk”, *Econometric Reviews*, 25, 177-217.

Gourieroux, C. (2007) : ”Car and Affine Processes”, *Lecture Notes in Time Series*”, Springer, forthcoming.

Gourieroux, C., and J., Jasiak (2007) : ”The Econometrics of Individual Risk : Credit, Insurance and Marketing”, Princeton University Press.

Gourieroux, C., Jasiak, J., and R., Sufana (2007) : ”The Wishart Autoregressive Process for Stochastic Volatility”, forthcoming *Journal of Econometrics*.

Gourieroux, C., Monfort, A., and V., Polimenis (2006) : ”Affine Term Structure Models for Credit Risk”, *Journal of Financial Econometrics*, 4, 494-530.

Gourieroux, C., and R., Sufana (2006) : ”Classification of Affine Term Structure Models”, *Journal of Financial Econometrics*, 4, 31-52.

Human Mortality Database, University of California, Berkeley (USA), Available at <http://www.mortality.org>.

Hyndman, R., and M., Ullah (2005) : ”Robust Forecasting of Mortality and Fertility Rates : A Functional Data Approach”, DP 2/05, Monash University.

Lando, D. (1998) : ”On Cox Processes and Credit Risky Securities”, *Review of Derivative Research*, 2, 99-120.

Lazar, D. (2004) : ”On Forecasting Mortality Using Lee-Carter Method”, Babes-Bolyai University DP.

Lee, R. (2000) : ”The Lee-Carter Method of Forecasting Mortality, with Various Extensions and Applications”, *North-American Actuarial Journal*, 4, 80-93.

Lee, R., and L., Carter (1992) : ”Modeling and Forecasting the Time

Series of US Mortality”, *Journal of the American Statistical Association*, 87, 659-671.

Lee, R., and T., Miller (2001) : ”Evaluating the Performance of the Lee-Carter Method for Forecasting Mortality”, *Demography*, 38, 537-549.

Leippold, M., and L., Wu (2002) : ”Asset Pricing Under the Quadratic Class”, *Journal of Financial and Quantitative Analysis*, 37, 374-295.

Luciano, E., and E., Vigna (2005) : ”A Note on Stochastic Survival Probabilities and Their Calibration”, DP University of Turin.

Maghsoodi, Y. (1996) : ”Solution of the Extended CIR Term Structure and Bond Option Valuation”, *Mathematical Finance*, 6, 89-109.

Makeham, W. (1860) : ”On the Law of Mortality and the Construction of Mortality Tables”, *Journal of the Institute of Actuaries*, 8,

Milevsky, M., and S., Promislow (2001) : ”Mortality Derivatives and the Option to Annuitize”, *Insurance : Mathematics and Economics*, 29, 299-318.

Olivieri, A. (2001) : ”Uncertainty in Mortality Projections : an Actuarial Perspective”, *Insurance : Mathematics and Economics*, 29, 231-245.

Pitacco, E. (2004) : ”Survival Models in a Dynamic Context : a Survey”, *Insurance : Mathematics and Economics*, 35, 279-298.

Renshaw, A., and S., Haberman (2003)a : ”Lee-Carter Mortality Forecasting with Age Specific Enhancement”, *Insurance : Mathematics and Economics*, 33, 255-272.

Renshaw, A., and S., Haberman (2003)b : ”On the Forecasting of Mortality Reduction Factors”, *Insurance : Mathematics and Economics*, 32, 379-401.

Schrager, D. (2006) : ”Affine Stochastic Mortality”, *Insurance : Mathematics and Economics*, 38, 81-97.

Sunamoto, N. (2005) : ”Cohort Effect Structure in the Lee-Carter Residual Term”, DP Fukoku Mutual Life Insurance Company.

Vallin, J. (1973) : "La mortalité par génération en France depuis 1899", INED, Paris.

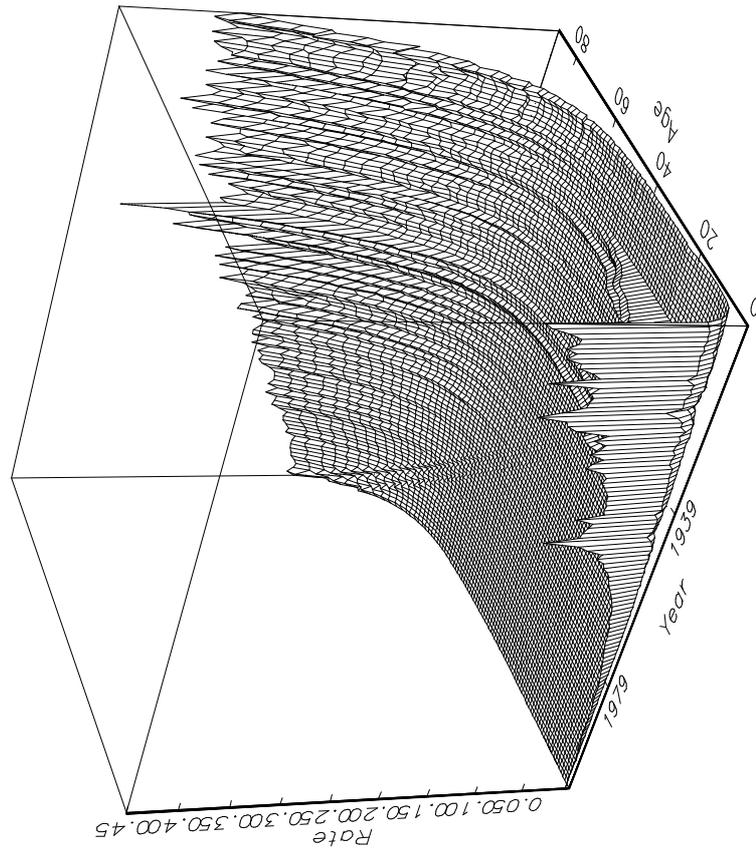
Wang, J. (2007) : "Fitting and Forecasting Mortality for Sweden : Applying the Lee-Carter Model", DP 2007:1, Stockholm University.

White (2002) :

Willets, R. (2004) : "The Cohort Effect : Insights and Explanations", British Actuarial Journal, 10,

Yang, S. (2001) : "Reserving, Pricing and Hedging for Guaranteed Annuity Options", Ph.D. Thesis, Heriot-Watt University, Edinburgh.

FIGURE 1 : Mortality Rate in France:Females:Males +Males



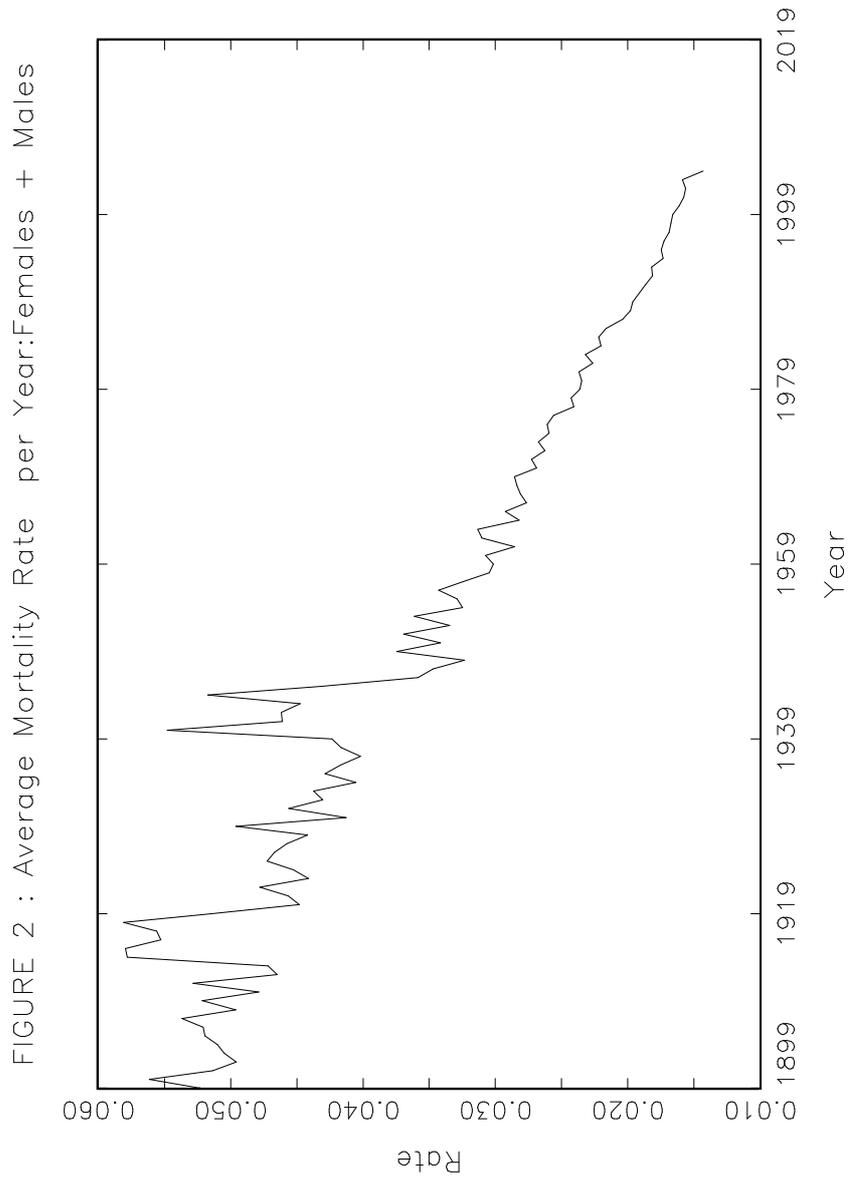


FIGURE 3 : Average Mortality Rate per Age (Females+Males)
Observed (solid line), Fitted (dotted line)

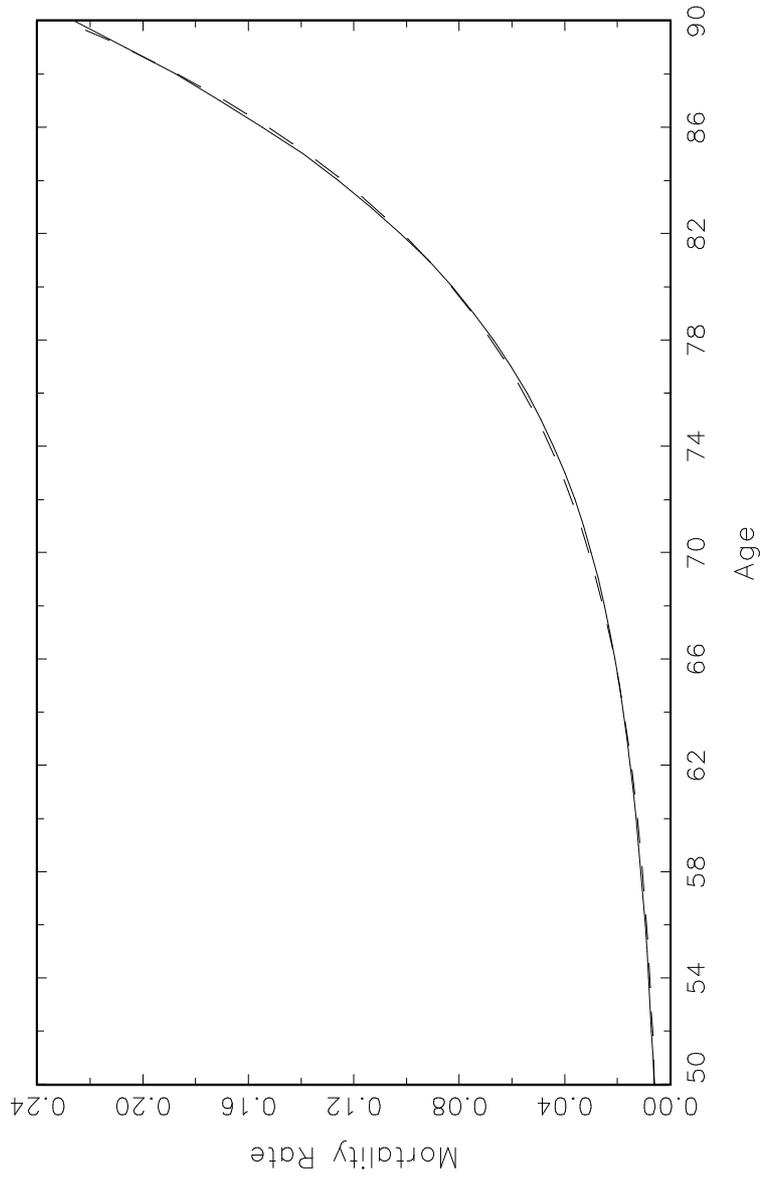


FIGURE 4 : Extracted Factor , Global Method (Females+Males)

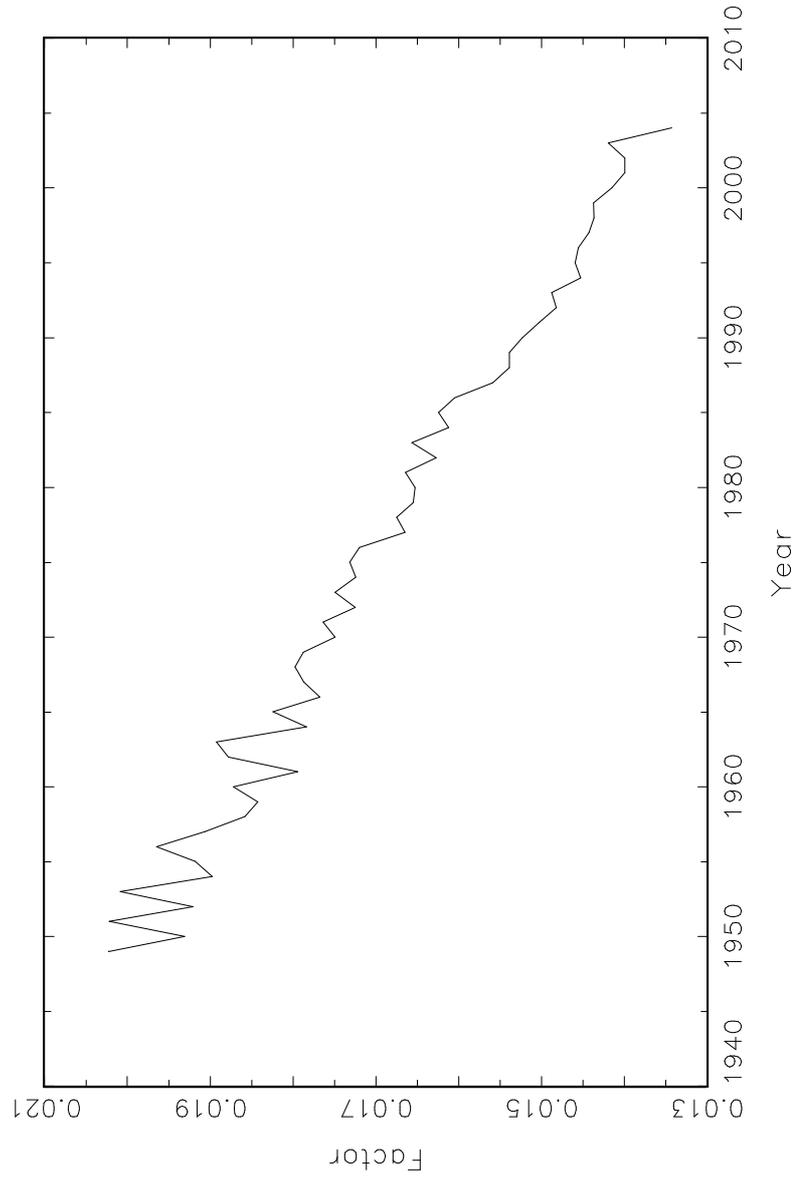


FIGURE 5 :Factor Residuals (Females+Males)

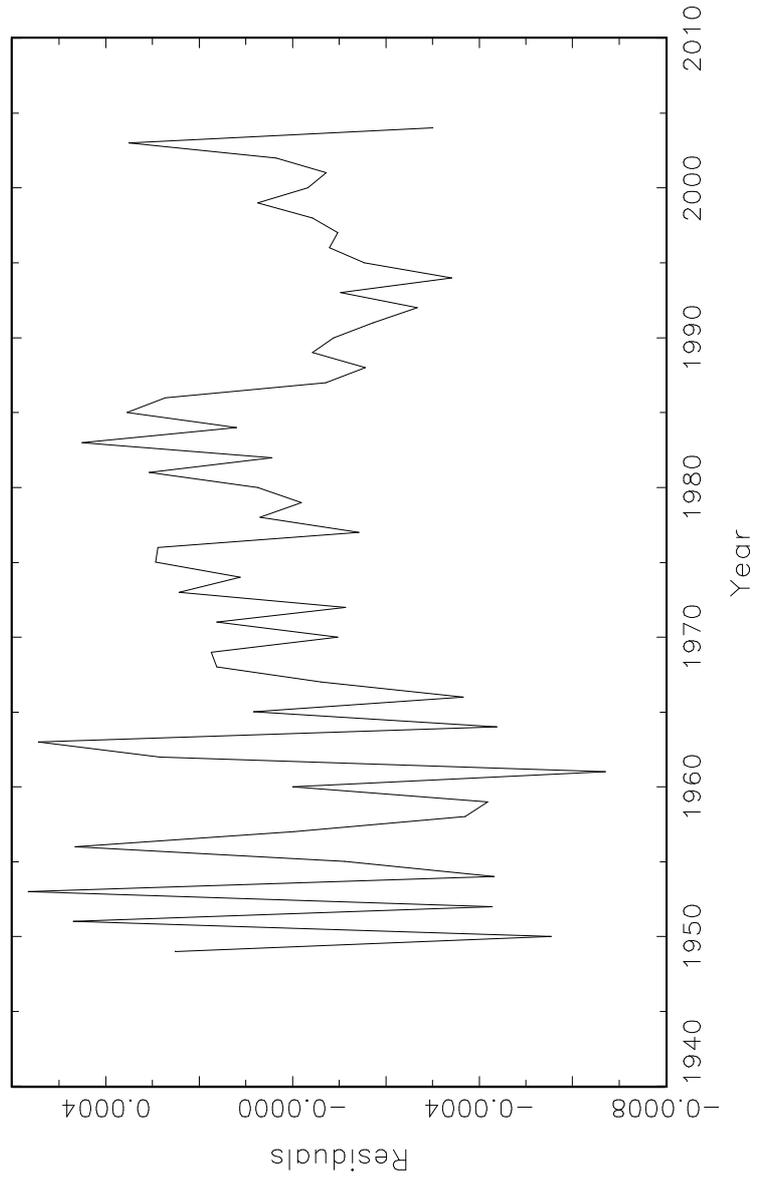


FIGURE 6: Data and Predictions of Mortality Rates(Females+Males)

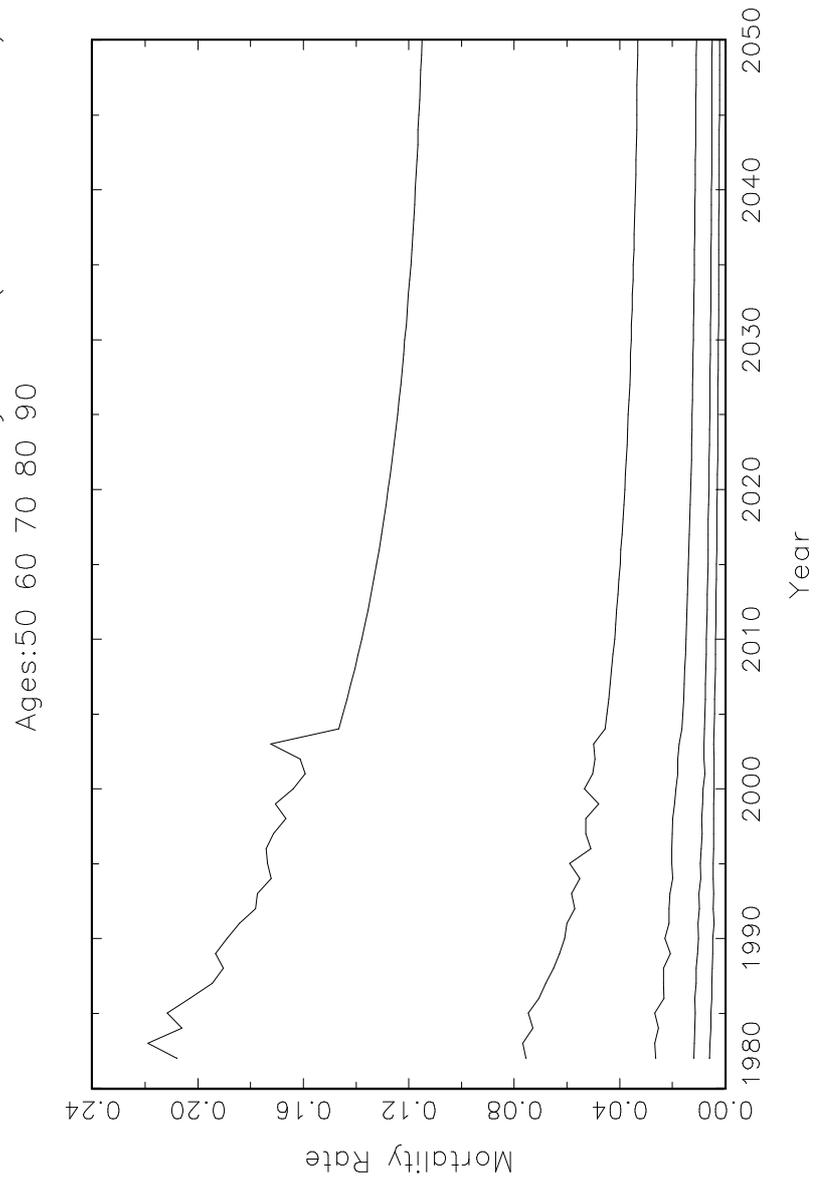


FIGURE 7: Distribution of Residual Life Time
Cohort 50 years in 2004
Solid line:Total, Dashes:Males, Short dashes:Females

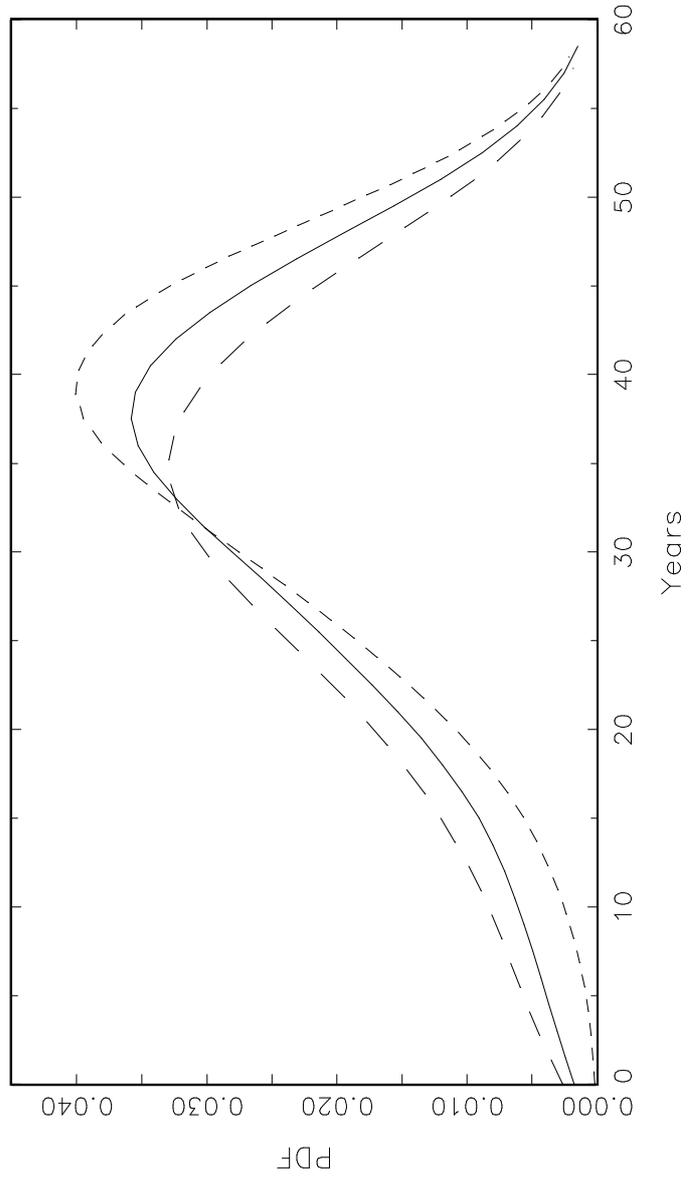


FIGURE 8: Distribution of Residual Life Time
Cohort 65 years in 2004
Solid line:Total, Dashes:Males, Short dashes:Females

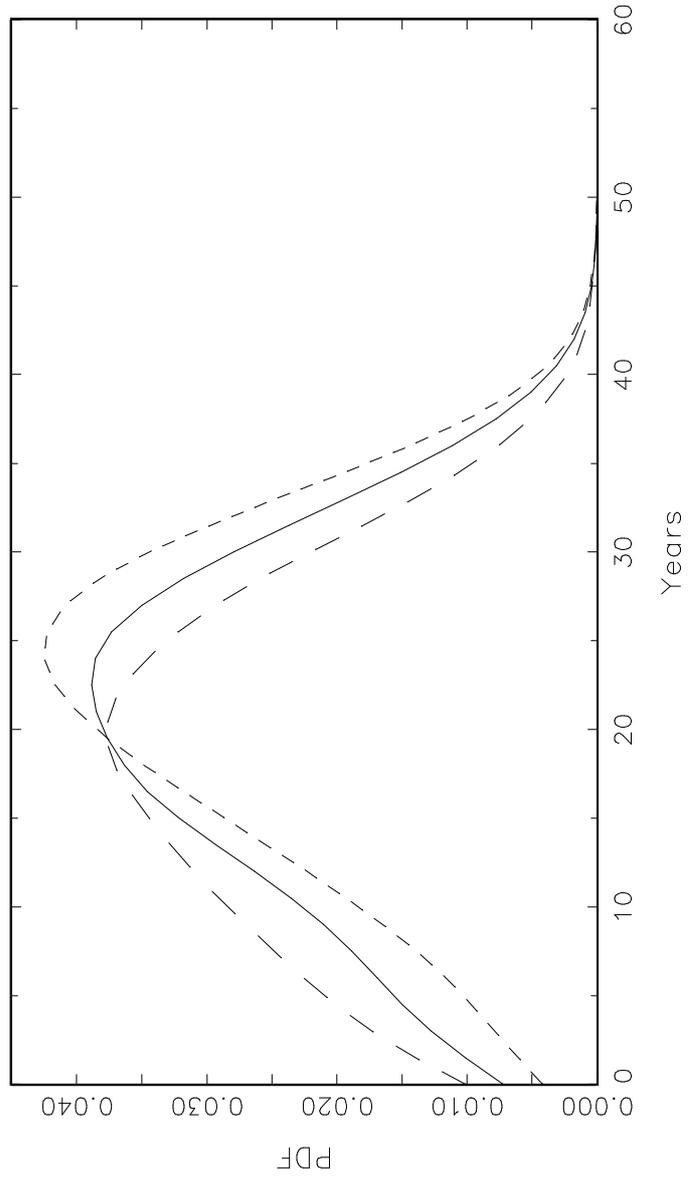


FIGURE 9: Distribution of Residual Life Time
Cohort 80 years in 2004
Solid line:Total, Dashes:Males, Short dashes:Females

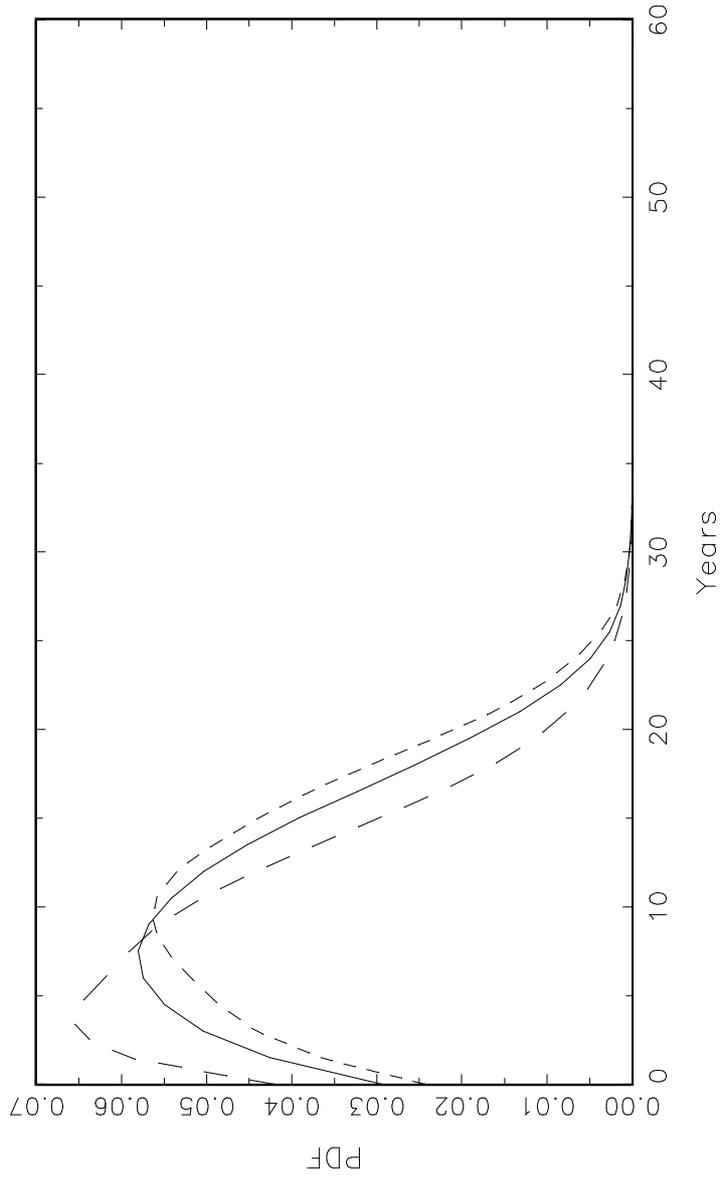


FIGURE 10 Prospective or Forward Mortality Rate
 Cohort 80 years in 2004
 Total Males Females

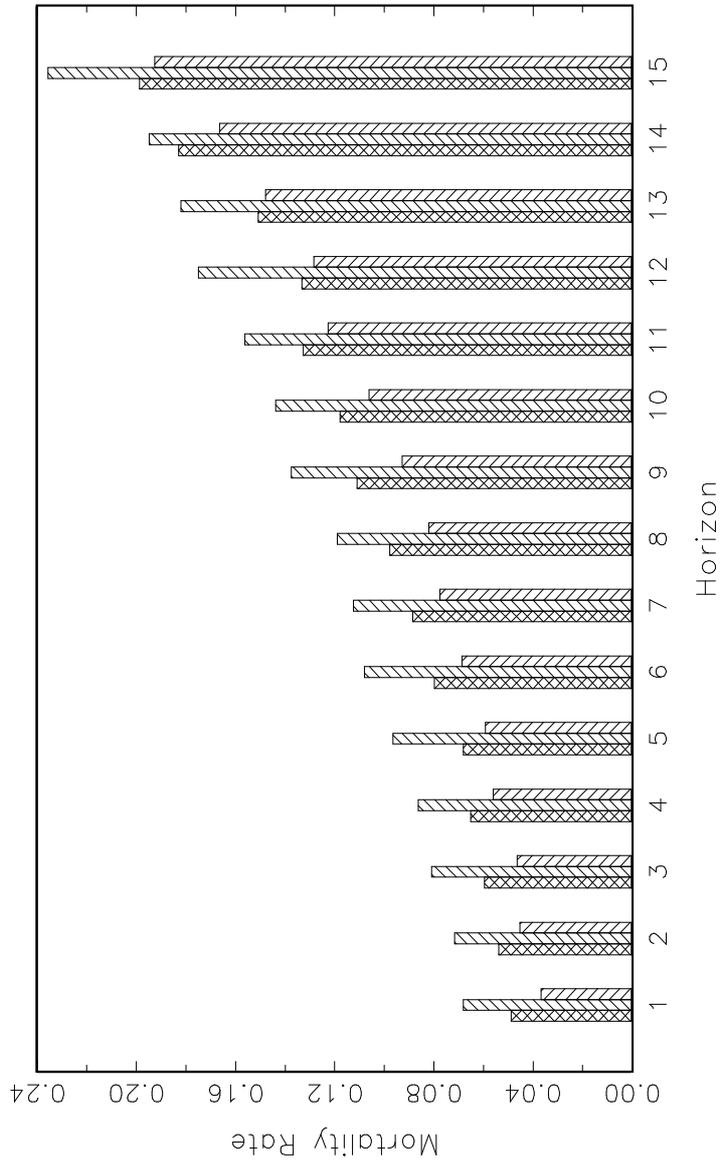


FIGURE 1—m : Mortality Rate in France:Males

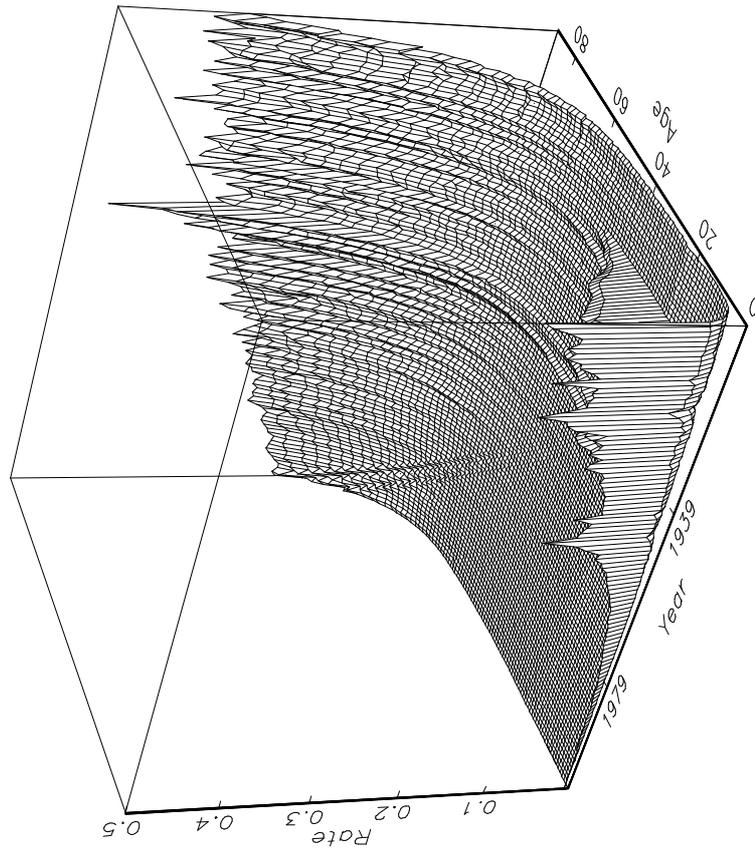


FIGURE 2-m : Average Mortality Rate per Year: Males

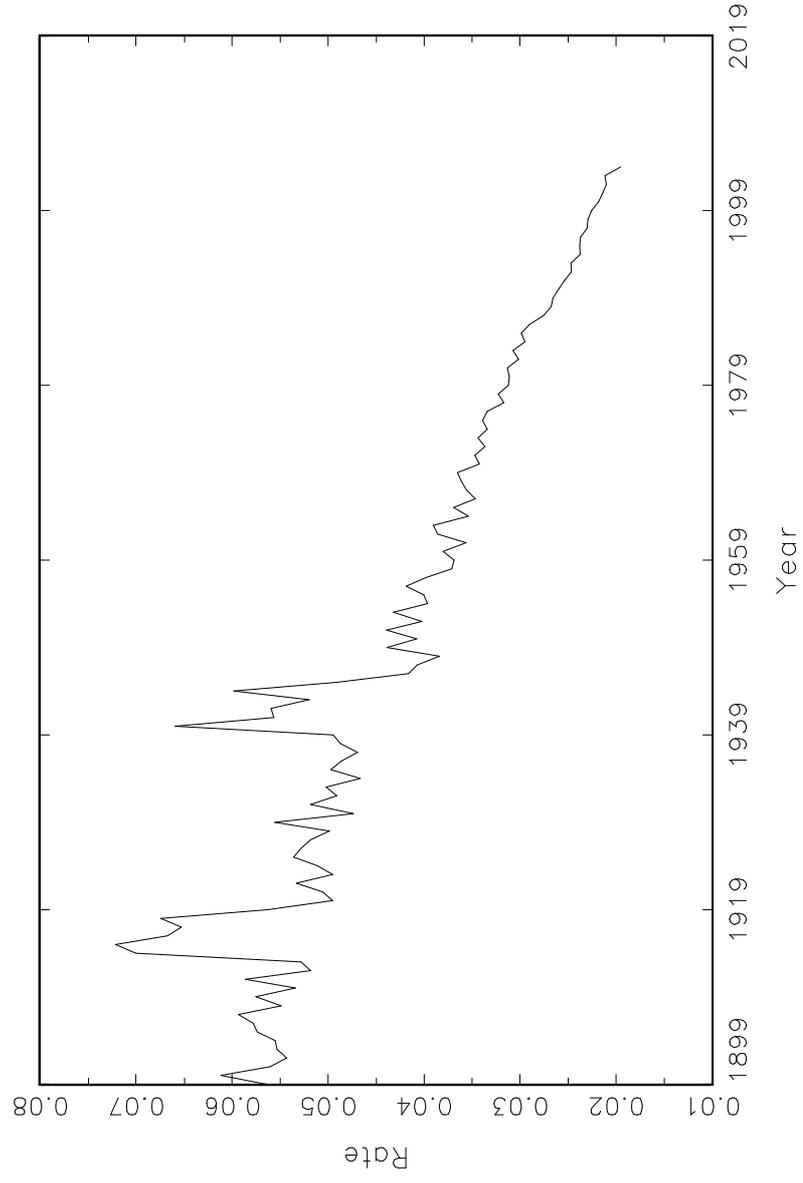


FIGURE 3—m : Average Mortality Rate per Age (Males)
Observed (solid line), Fitted (dotted line)

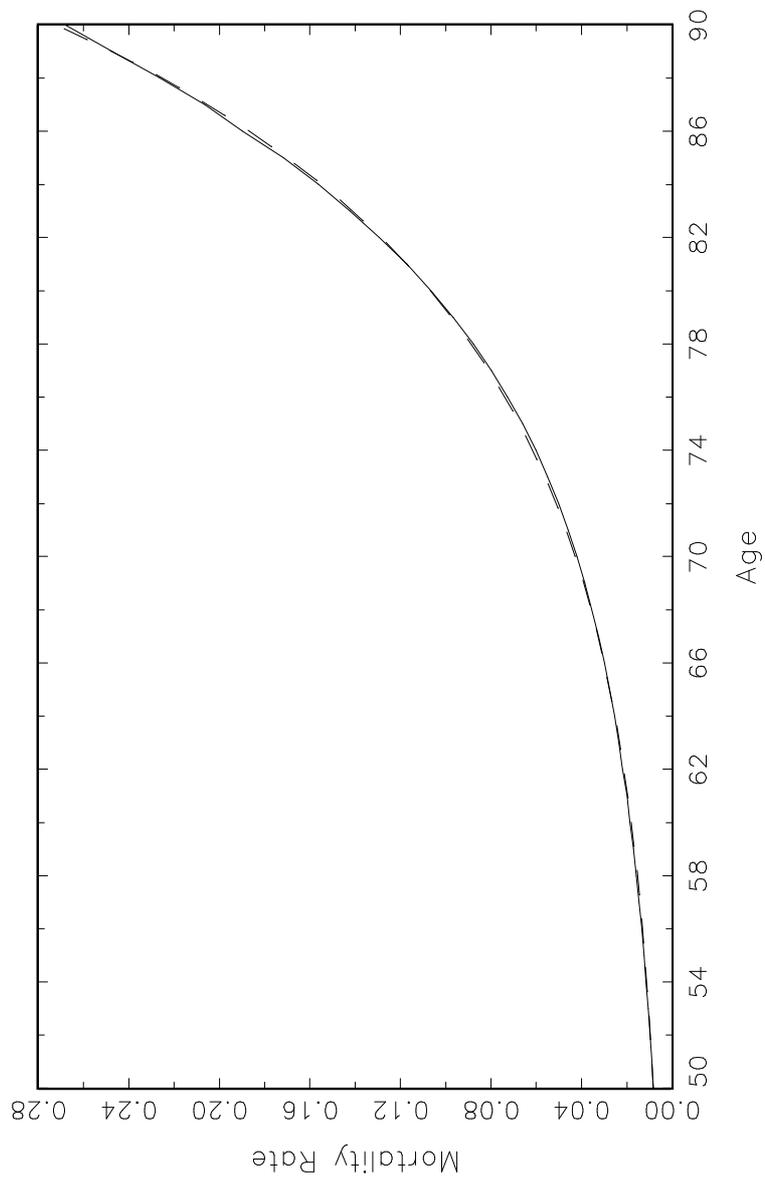


FIGURE 4—m : Extracted Factor , Global Method (Males)

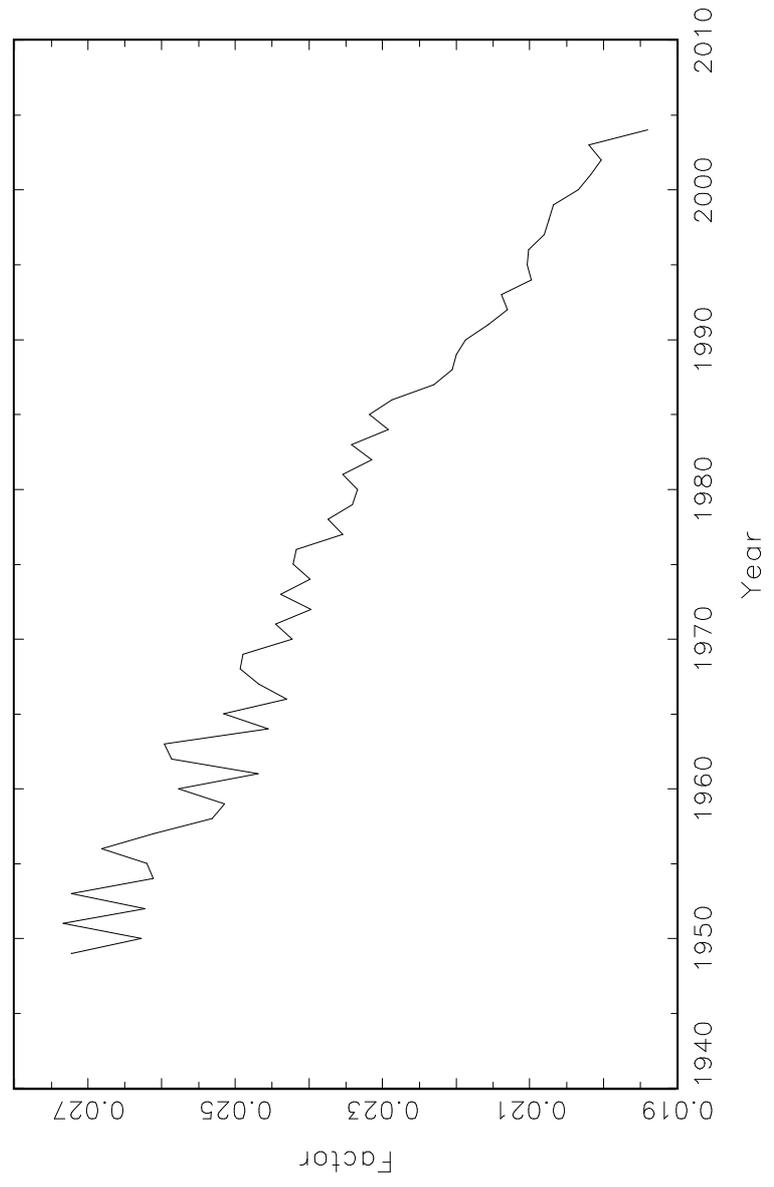


FIGURE 5-m : Factor Residuals (Males)

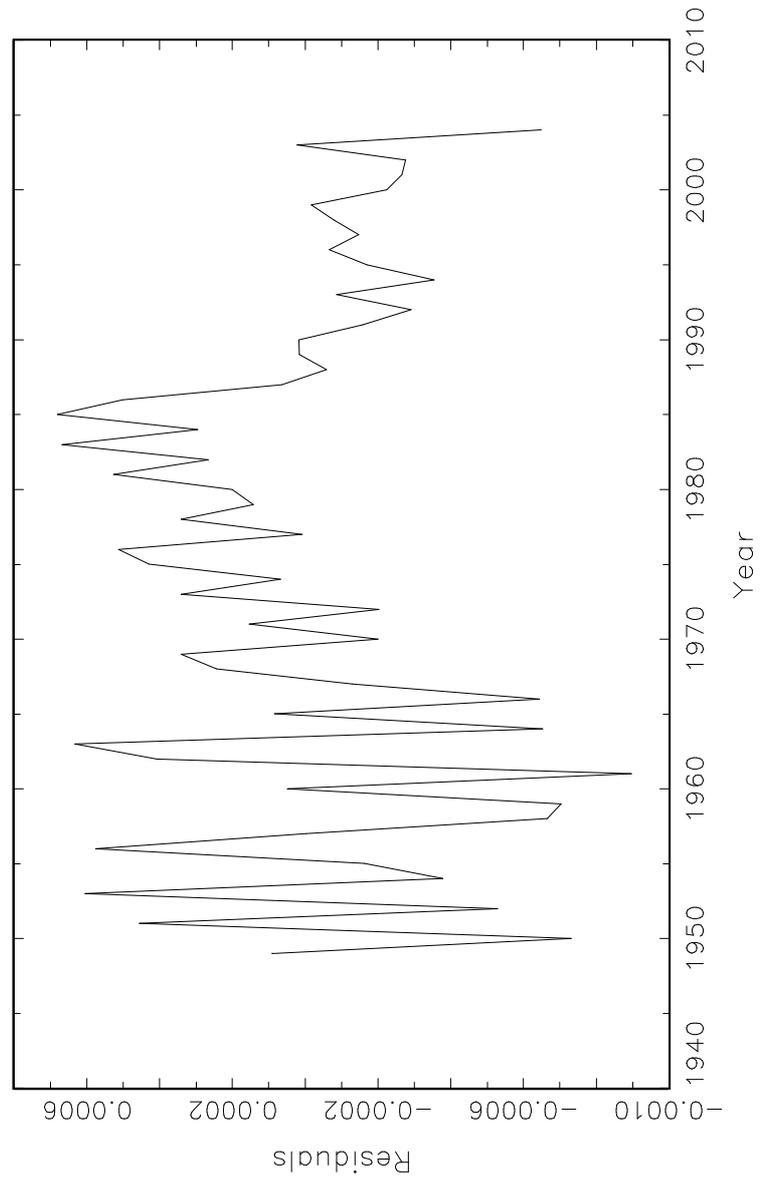


FIGURE 6—m;Data and Predictions of Mortality Rates(Males)
Ages:50 60 70 80 90

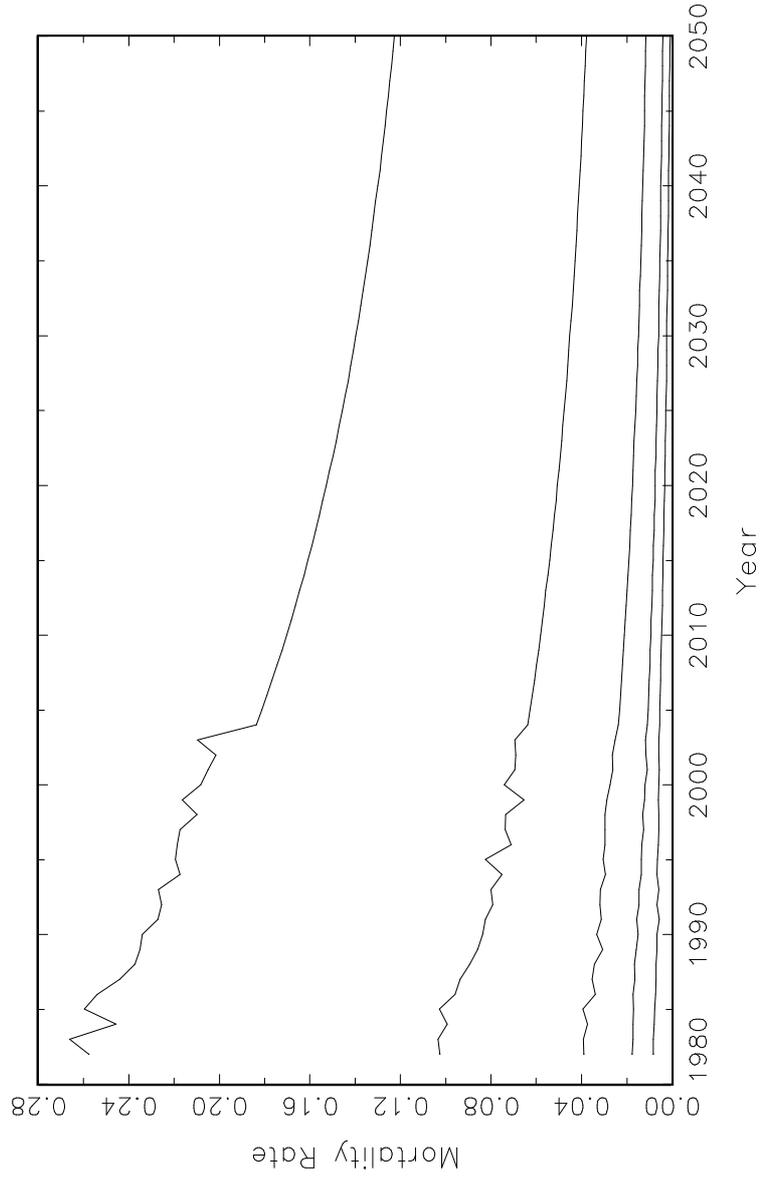


FIGURE 1-f : Mortality Rate in France:Females

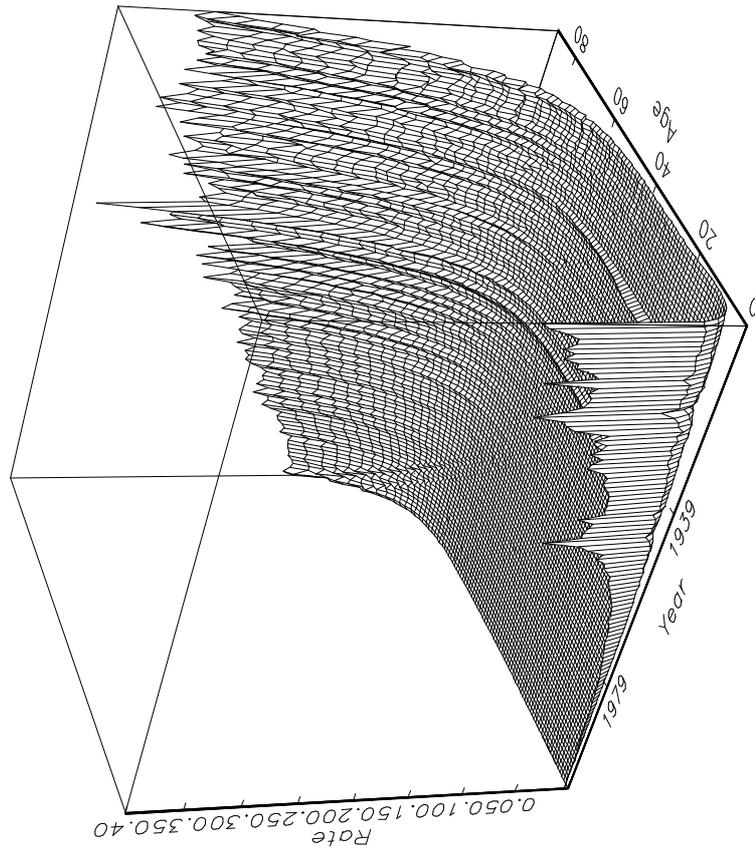


FIGURE 2-f : Average Mortality Rate per Year: Females

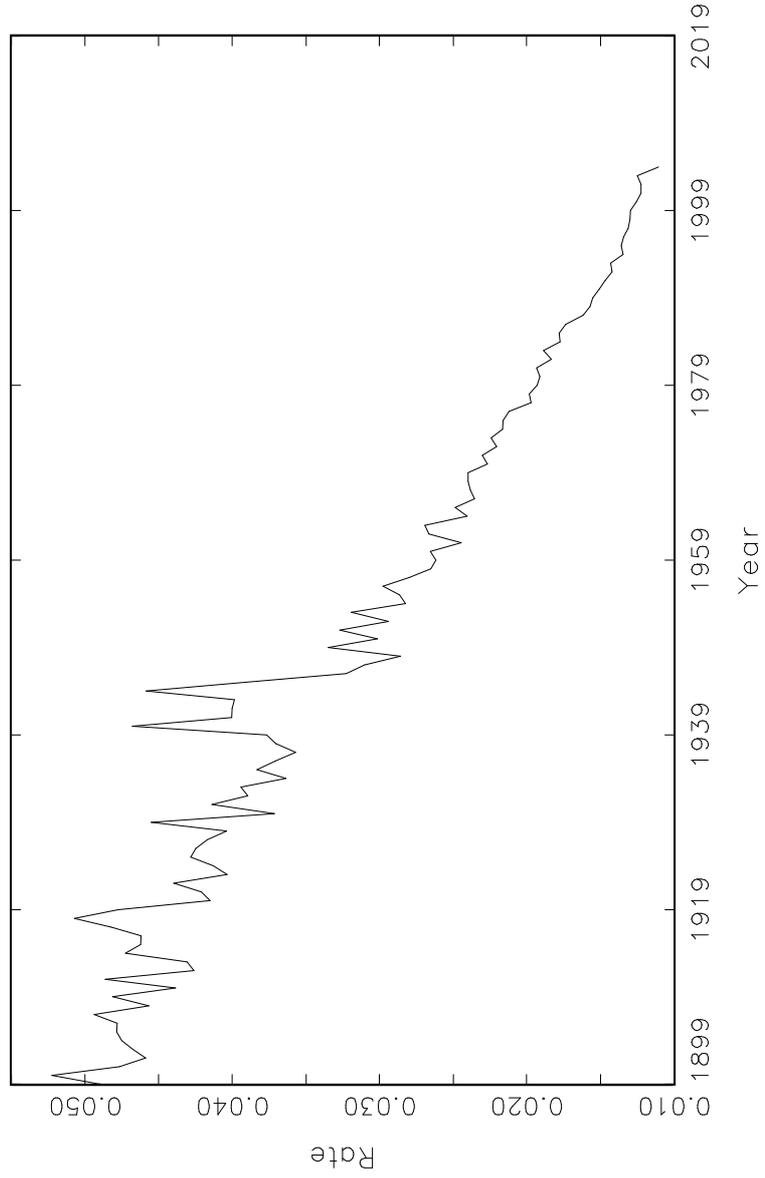


FIGURE 3-f : Average Mortality Rate per Age(Females)
Observed (solid line), Fitted (dotted line)

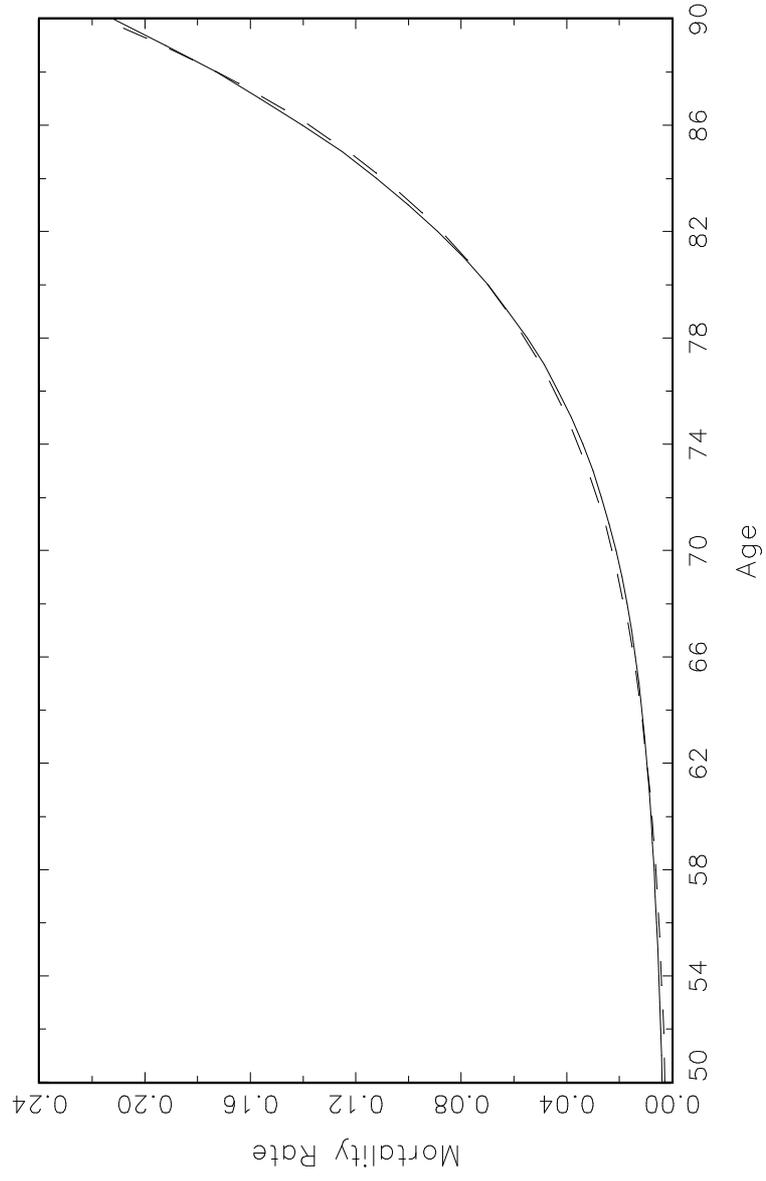


FIGURE 4-f : Extracted Factor , Global Method(Females)

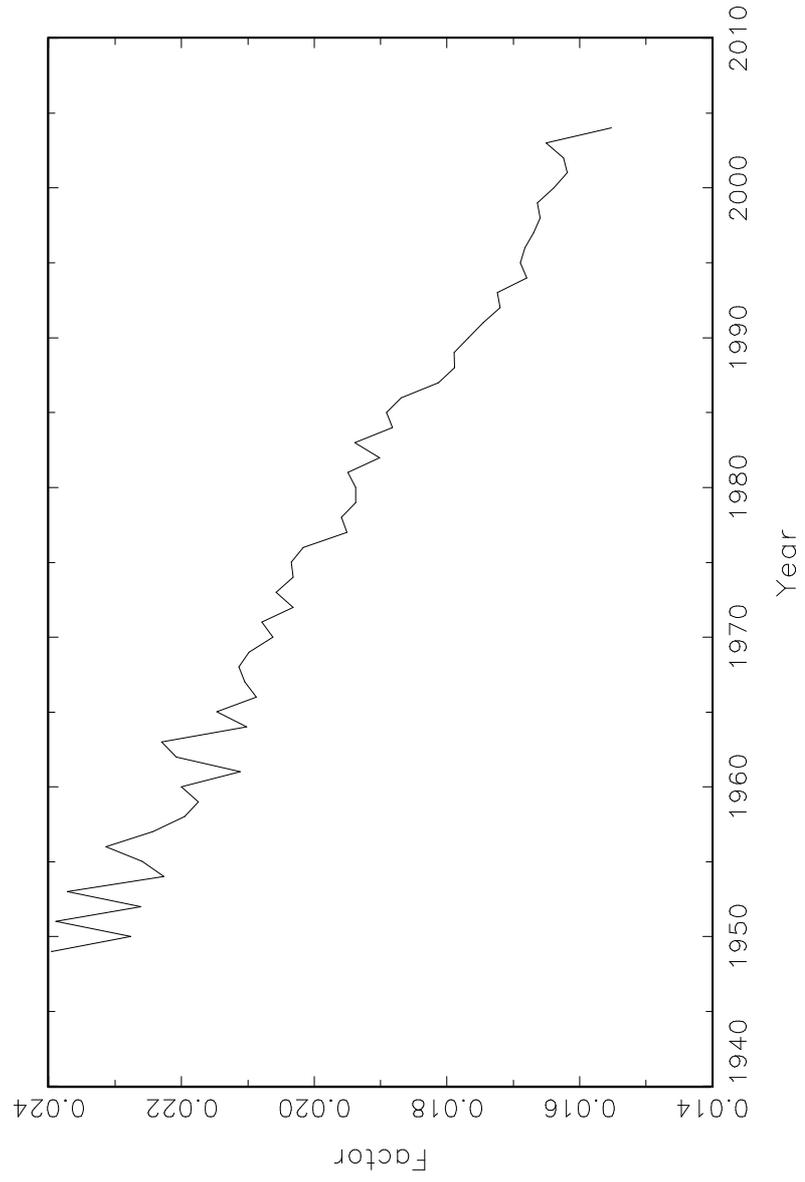


FIGURE 5—f : Factor Residuals(Females)

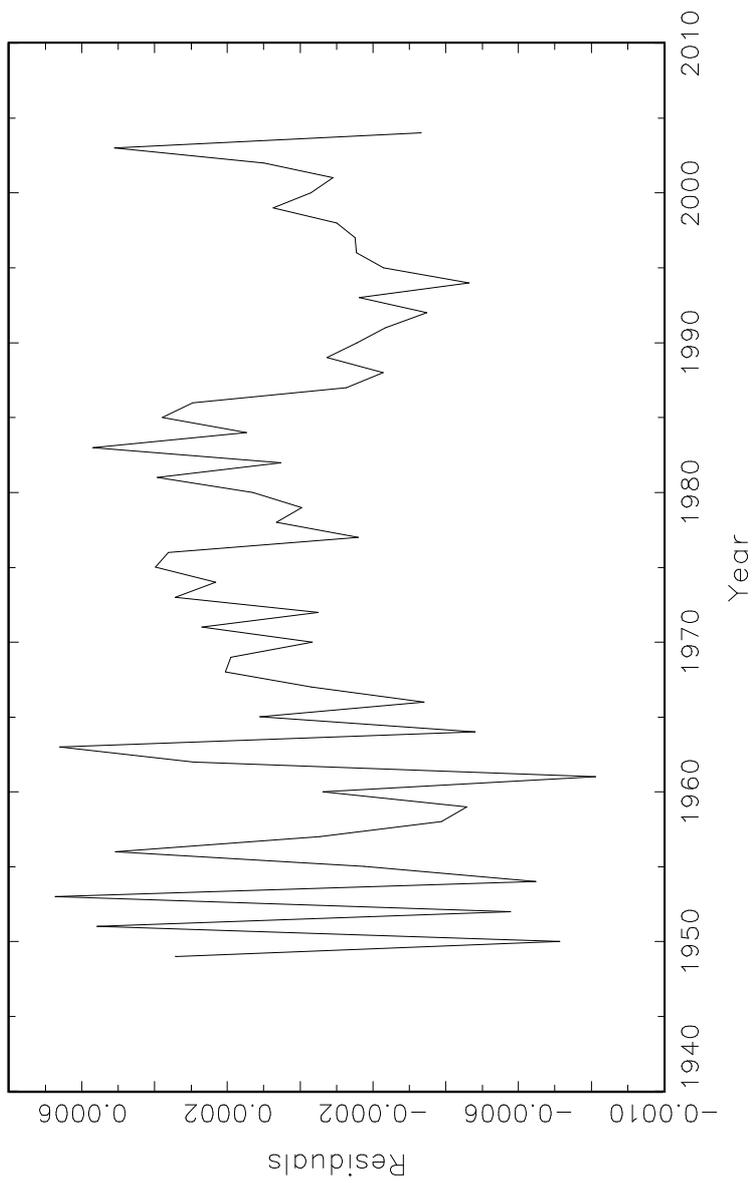


FIGURE 6-f: Data and Predictions of Mortality Rates(Females)
Ages:50 60 70 80 90

