A mortality model based on a mixture distribution function

Stefano Mazzuco *1, Bruno Scarpa †2, and Lucia Zanotto ‡3

[1] *Department of Statistical Sciences – University of Padova*
[2] *Department of Statistical Sciences – University of Padova*
[3] *Department of Statistical Sciences – University of Padova*

December 15, 2015

Abstract

A new mortality model based on a mixture distribution function is proposed. We mix a half-normal distribution with a generalization of the skew-normal distribution. As a result we get a six-paramters distribution function having a good fit with a wide variety of mortality patterns. This model is fitted to several mortality data and compared with Siler (five parameters) and Heligman-Pollard (eight parameters) models. The proposed model can be a convenient compromise between Heligman-Pollard model (which ensures a good fit with data but is often overparameterised) and Siler model (which is more compact but fail to capture accident humps).

Keywords: Mortality model, mixture distribution, skew bimodal normal distribution

1 Introduction

Human mortality modelling is an issue that attracted the attention of several demographers and scholars from other fields. However, in contrast to fertility

---

*mazzuco@stat.unipd.it
†scarpa@stat.unipd.it
‡zanotto@stat.unipd.it

modelling field, where several different models have been – even recently proposed (see, for instance, Chandola et al., 1999; Schmertmann, 2003; Peristera and Kostaki, 2007; Mazzuco and Scarpa, 2015), very few models are available for such analysis. The most popular alternative is the Gompertz (1825) model: the latter owes its popularity to its compactness and parsimony's, which make it particularly suited for comparative work. However Gompertz model does not incorporate all source of mortality, in particular infant and young adulthood mortality, while it is particularly focused on senescent mortality. Makeham (1860) generalized the Gompertz model, but such a generalization still disregards mortality at ages below 30. A further generalization has been provided by Siler (1979; 1983): a negative Gompertz function is added to Gompertz-Makeham model, so that infant mortality is also fitted. Finally, Heligman and Pollard (1980) proposed a 8-parameters models also incorporating what is usually referred to as "accident hump" (i.e. a hump often observed among males deaths function in correspondence of young adulthood ages). In developed countries this is interpreted as a peak in mortality due to a peak of male hormone production during puberty that increases the likelihood of dangerous behaviors by male adolescents (see Goldstein, 2011). However, in developing world, this hump is not necessarily confined to male population, and in case of high diffusion of HIV epidemic a much larger bulge can be found among female population. Sharrow et al. (2013) use Heligman-Pollard model to fit mortality data of a population with particularly high diffusion of HIV. However Rogers (1986) and Gage and Mode (1994) noted that estimation of Heligman-Pollard model is difficult due to model overparameterization and numerical issues of the commonly used estimation approach (i.e. weighted least squares). These estimation problems lead Dellaportas et al. (2001) and Sharrow et al. (2013) to use a Bayesian estimation approach. In this contribution, an additional model is proposed by mixing a half-normal and a Skew Bimodal Normal (Rocha et al., 2013) distribution. In this way we get a six-parameters model that is more parsimonious (thus with less identification issues) than the Heligman-Pollard model but accounts for all sources of mortality, including the "accident hump". It worths noting that in the proposed mixture model, deaths rather than rates (as in the Gompertz-Makeham model) or probabilities (as in the Heligman-Pollard model) are modeled. This choice stems from the fact that deaths of a life table can be seen as a probability density function. Therefore, this model can be easily estimated with a maximum likelihood approach, but other approaches such as weighted least squares or Bayesian inference can be used. The mixture model will be fitted with several mortality data of different shapes and compared with other models. We will follow the guidelines suggested by Congdon (1993) to evaluate model performance, which list the most important characteristics that a mortality model should have: smoothness, parsimony, interpolation, comparability, suited for trends and forecasting and analytic manipulation.

2 Modelling the "deaths" function

The model proposed by Heligman and Pollard (1980) represents the probability of dying at age x, q_x, using the following formula

$$\frac{q_x}{1-q_x} = A^{(x+B)^C} + D\exp\left[-E\left\{\log\left(\frac{x}{F}\right)\right\}^2\right] + GH^x \tag{1}$$

Basically, overall mortality is decomposed into three parts: the first one related to infant and child mortality (A represents the infant mortality rate and B mortality rate for one year old children, C is related to the rate of mortality decline after age 1), the second one is related to mid-life mortality with D, E and F representing the severity, the spread and the location of the accident hump, respectively; the third part reflects old age mortality being G the base level of old age mortality and H the rate of increase of mortality with age.

de Beer and Jansen (2014) define an extension of the (1). The model proposed by Siler (1979, 1983) represents the death rate at age x using the formula

$$m_x = a_1 e^{-b_1 x} + a_2 + a_3 e^{b_3 x} \tag{2}$$

This model also decomposes mortality into three components, but differently from Heligman-Pollard model accident humps are not allowed, a constant death rate in mid-life ages is assumed, instead. It might be noted that the (2) is a generalization of Gompertz-Makeham model, which is obtained by imposing $a_1 = 0$.

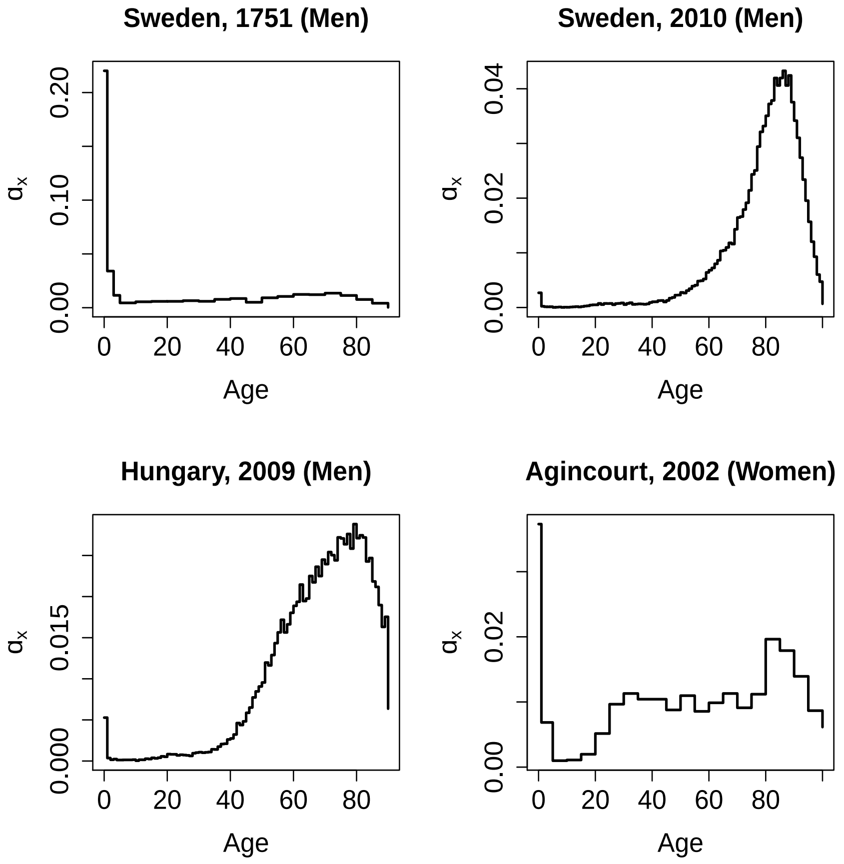
The underlying idea of our proposal is that deaths rather than rates (as in the Siler model) or probabilities (as in the Heligman-Pollard model) are modeled. This choice stems from the fact that deaths of a life table can be seen as a probability density function, thus maximum likelihood estimation becomes straightforward. Dellaportas et al. (2001) also uses deaths to fit Heligman-Pollard model, while Sharrow et al. (2013) use a Bayesian melding approach to death probabilities. We therefore need to find a density function that can have a good fit with deaths of a life table. This means finding a probability density function with at least two modes (one at birth and one at the Lexis point) but with possibly a third one, catching the so-called "accident hump". Figure 1 shows death functions from four different life tables: pre and post transitional Sweden, Hungary 2009 (with a particularly high skewness of adult age mortality) and Agincourt health and demographic surveillance system site in South Africa 2002. The first three life tables are provided by Human Mortality Database while the latter is provided by Sharrow et al. (2013). Sweden 1751 has a typical pre-transitional shape, with child mortality overwhelming the mortality at adult ages. In opposition, Sweden 2010 has a post-transitional shape, with a very limited infant mortality and a much more relevance of old age mortality. Hungary 2009 is a particular case with a particularly high skewness of adult deaths. Finally, Agincourt site mortality is affected by HIV diffusion, and an additional hump is observed at young adulthood ages. Siler model is likely to have a good fit in the first three cases but not in the Agincourt site, while Heligman-Pollard model is expected to have a good fit in Agincourt site, as also shown by Sharrow et al. (2013), but is also likely to have several identification problems in the other cases. We are looking for a model having a good fit in all cases without running into identification problems as the Heligman-Pollard model does. We, therefore decompose overall mortality into two parts: child mortality and adult mortality. Infant and child death function is modeled by means of a half-normal distribution, which is a normal distribution with mean 0 limited to the domain $(0, +\infty)$ and has the following distribution function

$$f_Y(y;\sigma) = \frac{\sqrt{2}}{\sigma\sqrt{\pi}} \exp\left(-\frac{y^2}{2\sigma^2}\right) \quad y > 0 \tag{3}$$

$f_Y(y;\sigma)$ has a maximum in zero, and then decreases as y increases. The rate to which it decreases depends on the value of the scale parameter σ: the higher σ the slower the mortality decline.

Adult death function is modeled using a generalization of the skew normal distribution. Skew normal distribution has been firstly introduced by Azzalini

Figure 1: Different deaths functions from Sweden (1751, 2010), Hungary (1994) and Agincourt health and demographic surveillance system site in South Africa (2002)



(1985) and its distribution function is given by

$$f_X(x;\xi,\omega,\alpha) = \frac{2}{\omega}\phi\left(\frac{x-\xi}{\omega}\right)\Phi\left(\alpha\frac{x-\xi}{\omega}\right) \tag{4}$$

where $\phi(\cdot)$ is the standard normal density function, $\Phi(\cdot)$ the standard normal pdf, ξ the location parameter, ω the scale parameter and α the shape parameter. It can be easily seen that if $\alpha = 0$ the (4) reduces to a normal density function. Since the the article by Azzalini (1985), extensions of skew-normal distribution received much attention in the literature (see Azzalini and Capitanio, 2013).

In our case, it is of particual interest that skew-normal family of distributions can be extended obtaining bi-modal skewed distributions. Rocha et al. (2013), for example, propose a generalization of the skew-normal distribution, called Skew Bimodal Normal (SBN), which characteristics are discussed by Elal-Olivero et al. (2009). Its probability density function is the following

$$f_X(x;\xi,\omega^2,\alpha,\lambda) = 2\omega^{-1}\left(\frac{1+\alpha\left(\frac{x-\xi}{\omega}\right)^2}{1+\alpha}\right)\phi\left(\frac{x-\xi}{\omega}\right)\Phi\left\{\lambda\left(\frac{x-\xi}{\omega}\right)\right\}. \quad (5)$$

The (5) can have - at most - two modes depending on the value of α. In particular if $\alpha \geq 0.5$, the resulting pdf has two modes, otherwise ($\alpha < 0.5$) the pdf is unimodal. Thus, by mixing the (3) and the (5) we get the mixture model

$$f_Z(z;m,\sigma,\xi,\omega^2,\alpha,\lambda) = m \cdot f_Y(z;\sigma) + (1-m)\cdot f_X(z;\xi,\omega^2,\alpha,\lambda) \quad (6)$$

where m is the mixture parameter (with support $[0,1]$), σ the parameter of the half-normal distribution and ξ, ω, α and λ the parameters of the SBN distribution.

Inference on model (6) is quite straightforward, considering that the death function is a density function determining the age distribution of deaths. Then, let us suppose we have the death function d_x and the total number of deaths occurred D, therefore the death occurred at age x are $D_x = d_x \cdot D$. The likelihood of vector $\mathbf{D_x}$ is given by a multinomial model

$$L(m,\sigma,\xi,\omega^2,\alpha,\lambda;\mathbf{D_x}) = \prod_{x=0}^{\Omega} p(x)^{D_x}. \quad (7)$$

Then, a parametric form can be specified for $p(x)$ and if we assume that deaths are distributed across age according density in (6), we can write

$$p(x;m,\sigma,\xi,\omega^2,\alpha,\lambda) = \int_x^{x+1} f_Z(z;m,\sigma,\xi,\omega^2,\alpha,\lambda)dz \quad (8)$$

and maximize the (7).

3 Interpreting parameters

Among the advantages of representing mortality through a parametric curve Congdon (1993) cites comparability and forecasting. Forecasting in particular can be done on the base of reasonable assumptions about parameters of the model. Thus, if a parametric model should be used for forecasting – and comparison, as well – we need a sensible demographic interpretation of all parameters of (6). Although, model formula might look awkward, it is not difficult to find an interpretation of parameters m, σ, ξ, ω, α and λ. This can be done in the light of the work by Cheung et al. (2005), who define three dimensions of the survival curve: horizontalization (i.e. how many survivors live before old age), verticalization (i.e. the concentration of aging-related deaths) and longevity extension (i.e. the relative length of the tail of longevity). We will refer parameters in (6) to these three dimensions showing the values they take in four different mortality regimes: a pre-transitional regime (Sweden, 1781), a post-transitional regime (Sweden, 2010), a developing country with high HIV prevalence (Agincourt, 2002, women) and a post-transitional regime with a relative high skewness of adult deaths (Hungary, 1994). Estimates are reported in table 1, together with standard errors, which have been computed via bootstrap. Estimates for Siler and Heligman-Pollard models are reported in Appendix B. Moreover, we fit mixture model to data from a long time series USA (1933–2010), so we can map parameter estimates for every with sensible life tables quantities. In this way, it should be easier to interpret parameters of mixture model. Moreover, following Congdon (1993), we bear in mind that goodness of fit should be considered together with statistical stability, a problem especially arising for overparameterized models. Thus, fitting data from a time series also allows us to assess statistical stability. USA mortality is relatively smooth over time, thus we should expect a smooth pattern also for model parameters. This data has been taken from Human Mortality Database and figures of parameters' estimates and life table quantities and of parameters' estimates trend over time in USA 1933-2010 are shown in appendix.

The mixture parameter m is a measure of the relative weight of the two mixing distributions, thus, in our model, can be interpreted as the relative weight of infant and child mortality, captured by $f_Y(z;\sigma)$ on overall mortality. From table 1 we can see that the highest value of m is found for Sweden 1781 and the lowest for Sweden 2010, as we would expect. The relation between mixture parameter and child mortality is even more clearly shown by figure 3 where mixture parameter is mapped with values of q_{0-5} from life tables.

The scale parameter of the half normal distribution σ is a measure of the extent to which the peak of mortality at birth is extended to ages greater than 0. Thus a high value of σ indicates that also child mortality (i. e. mortality between ages 1 and 5) and not only infant mortality (i.e. mortality at age 0) is high. In table 1 we can see that in pre-transitional Sweden and Agincourt site estimate σ is the highest, while in Hungary even though infant mortality is relatively high, child mortality is quite limited, as it can be seen from figure 1. Figure 4 shows that Half-normal scale parameter is strongly related with q_{1-5}. It should be noted that the rlation is less evident when the values of q_{1-5} are low. This reveals that this parameter is not well identified when infant and child mortality (we expect that when q_{1-5} is low also q_0 is low) is low. However the same problem arises for Siler model: when infant and child mortality becomes low, estimates of b_1 parameter becomes increasingly variable.

Mixture and half-normal scale parameters are related to the degree of horizontalization of the survival curve, as the higher child and infant mortality and the lower share of survivors live until old ages. A third parameter related to horizonatalization is α, which is the parameter determining whether the additional mode (the accident hump) exists or not. Rocha et al. (2013) and Elal-Olivero et al. (2009) show that if $\alpha \geq 0.5$ then the SBN density is bimodal. Therefore, if $\alpha \geq 0.5$ then an accident hump is detected by the mixture model and the higher the value of α the more pronounced the hump. In table 1 we can see that only for pre-transitional Sweden and Agincourt site α is greater than 0.5, revealing the hump that clearly appears in figure 1, while for the other countries there is no hump and $\alpha < 0.5$. Actually Hungary shows a small hump around 20, but it is quite negligible.

The SBN scale parameter ω indicates to what extent old age mortality is concentrated on modal age at death and can be taken as a measure of verticalization: the lower the variance, the steeper the decline of survival curve in old age. We can see a particularly high value of ω for Hungary and post-transitional Sweden and from figure 1 appears that in these countries we have a high variability around modal age at death. The low value of ω for Agincourt site should be interpreted keeping in mind the value of α: the latter parameter catches the premature deaths in the accident hump and therefore these are not considered by ω, which value decreases. We relate the values of SBN scale parameter with adult mortality (q_{5-50}) and it can be seen (see figure 5) that the highest the adult mortality indicators the highest the value of ω.

The other shape parameter of SBN distribution indicates the distribution skewness. In particular, if λ is positive then the pdf is skewed on the right, while if λ is negative the distribution is skewed on the left. From table 1 we notice a particularly low value of λ for Hungary, and this reflects the high skewness we noticed in figure 1. The value of λ for Agincourt site indicates a symmetric distribution, but, once again, this should be interpreted considering the value of α that captures death determining the accident hump. As for interpretation, distribution skewness indicates where the mass of distribution is more concentrated. Thus a left skewed distribution means that most of aging related death are concentrated after the modal age. Thus, we relate the skewness parameter to number of deaths occured in adult age (after 50) concentrated after the modal age at deaths (figure 7). The association is particularly clear: the lower the skewness parameter, the higher the share of adult deaths concentrated after the modal age at death.

Finally, ξ is the location parameter of SBN, and it is strictly connected to average life duration of adults. Considering that life expectancy at birth is particularly sensitive to infant mortality level, we relate SBN location parameter to life expectancy at 10. From figure 6, clearly appears that the higher the SBN location parameter the higher the life expectancy at 10.

Therefore, we can state that a demographic interpretation can be found for all parameters, even though it cannot be denied that interpretation becomes more difficult for some of them (ω, λ and ξ) when an accident hump exists ($\alpha \geq 0.5$). The latter is an aspect of the model that needs further exploration, in particular it would be useful to know how SBN location, scale and skewness parameters are associated with α. Figure 8 shows the trend of parameters' estimates over time. It might be noted that that the curves are quite smooth (there is an expectation for α that has a peak around 1940, but its value remains below 0.5 threshold) and give some indications on US mortality trend: infant (mixture parameter) and child (HN scale parameter) mortality has fallen in these decades, longevity extension (SBN location) has increased, adult (5-50) mortality has also fallen (SBN scale) and mortality after 50 has increasingly been more concentrated after the modal age at death (SBN skewness). No accident hump has been observed through these years (SBN shape).

Table 1: Parameters estimate for mixture models in 4 countries (bootstrapped standard errors between parentheses)

Country	m	σ	ω	α	λ	ξ
Sweden (1781)	0.289	0.833	16.487	1.947	0.271	42.504
	(0.007)	(0.161)	(0.108)	(0.081)	(0.021)	(0.538)
Sweden (2010)	0.010	0.465	19.884	0.000	-6.480	94.904
	(0.007)	(0.413)	(0.402)	(0.045)	(0.336)	(0.358)
Hungary (2009)	0.069	0.357	18.773	0.449	-10.000	88.533
	(0.021)	(1.269)	(0.824)	(0.143)	(0.050)	(0.219)
Agincourt (2002)	0.094	0.928	15.515	2.110	0.131	58.880
	(0.016)	(0.134)	(0.383)	(0.300)	(0.067)	(1.766)

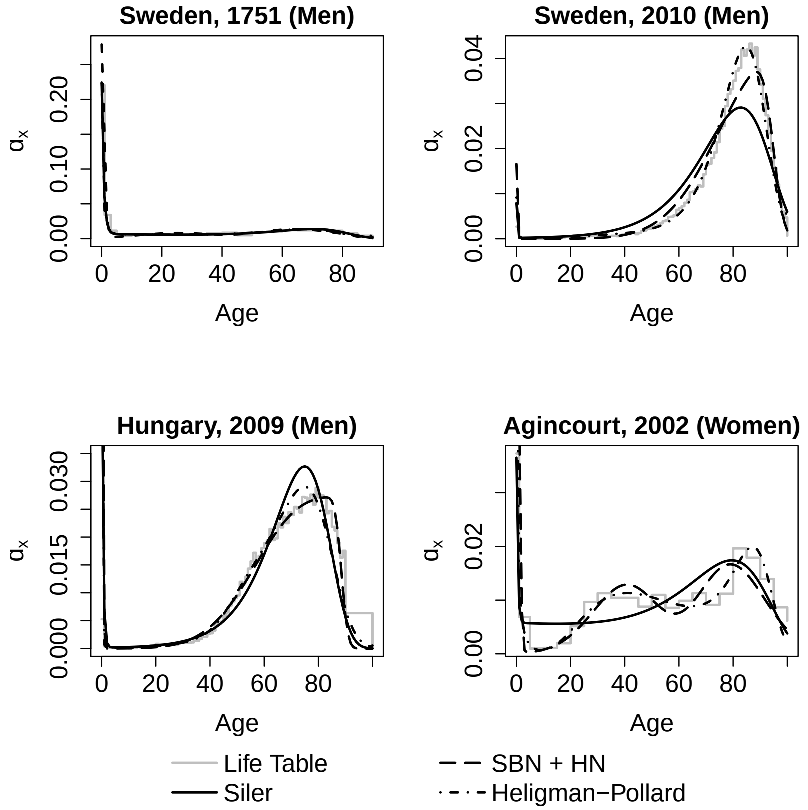
Table 2: Heligman-Pollard, mixture and Siler models AIC in 4 countries.

Country	Heligman-Pollard	SBN + HN	Siler
Sweden (1781)	182196.27	184246.40	182133.28
Sweden (2010)	337125.61	338787.65	343838.84
Hungary (2009)	543042.43	539920.98	544107.69
Agincourt (2002)	16279.43	16447.16	16623.39

4 Model fit

In this section, we compare the goodness of fit of model (6) and that of Siler and Heligman-Pollard models in the 4 cases exposed in figure 1. Table 1 shows the mixture model parameters' estimates for this data, while table 2 reports values of Akaike Information Criterion for all models in the 4 considered countries. Moreover, figure 2 shows the fit provided by mixture and Siler models to the four mortality regimes. Mixture model AIC is very close to Heligman-Pollard one in Sweden 2010 and even lower in Hungary and better than Siler model in Agincourt site. Its performance is not very good in Sweden 1751: in this case, since almost half of deaths occurs at age 0, it is crucial the estimate of infant mortality and Siler model provides the best fit for it. From curves in figure 2 it appears that all models have a reasonably good fit with data. There is one exception, though: Siler model does not capture the second hump for Agincourt site data, whereas Heligman-Pollard and mixture models do. This does not come as surprise, as Siler model cannot capture a second hump. Summarizing,

Figure 2: Different deaths functions from Sweden (1751, 2010), Hungary (2009) and Agincourt health and demographic surveillance system site in South Africa (2002) with fitted curves of Heligman-Pollard, Siler and SBN+HN models.



we can conclude that, mixture model has a good fit in all mortality age patterns also, contrary to Siler model, when there is an additional hump (the so-called "accident hump"). Heligman-Pollard has also a good fit in all the schedules we used but is less parsimonious than the mixture model, a characteristics that leads it to be often overparameterised, as noted by Congdon (1993). In addition, mixture model is also flexible enough to catch in a proper way mortality age distributions with a particularly high level of skewness, such as that shown by Hungary. Such a flexibility is lacking in both Siler and Hleigman-Pollard models, as hown by figure 2.

5 Conclusion

We defined a new mortality model based on a mixture of a half normal distribution and a Skew Bimodal Normal one. This model has been fitted to several mortality data with very different shapes of death counts curve, and a comparison has been made with the fit provided By Siler and Heligman-Pollard models. This model has shown several good properties. First it is flexible enough to have a relatively good fit with a wide range of mortality schedules, both pre and post transitional, with or without accident hump. The fit is good also with data with high skewness, where both Siler and Heligman-Pollard fail in catching data skewness. In no case we found a bad fit of the mixture model. Second, the mixture model is more parsimonious than the Heligman-Pollard model, and this makes it easier to fit. Moreover, since by modelling the death count function – which can be considered as a probability density function, we have a straightforward maximum likelihood estimate of parameters. Third, all the parameters have a sensible demographic interpretation and this makes it useful also for forecasting purposes.

Following the guidelines suggested by Congdon (1993) to evaluate model performance, we can say that mixture model is a valid choice because it is suited for interpolation, comparability, trends and forecasting and analytic manipulation. Given these first results, the analysis of the mixture model can be further extended in several directions. One possible extension is testing whether the model can be useful in cases where data are scarce and fragmentary. This problem has been faced by Wheldon et al. (2013) by means of a measurement error model, but assuming a parametric model on mortality, a Bayesian approach, specifying an informative prior distribution on parameters, can be considered to give a complete representation of mortality patterns. Bayesian inference on SBN model has already been discussed by Rocha et al. (2013), thus extending their suggested strategy to make inference on model (6) should be relatively easy. Another possible application of the mixture model is forecasting future mortality schedules. Figure 8 shows that the trend of mixture models parameters for USA is smooth enough to be extrapolated in future years. This can be done with simple time-series forecasting tool like, for instance, the Holt-Winters procedure (see, e.g., Holt, 1957; Winters, 1960).

References

Azzalini, A. (1985). A class of distributions which includes the normal ones. *Scandinavian Journal of Statistics 12*, 171–178.

Azzalini, A. and A. Capitanio (2013). *The Skew-Normal and Related Families.* Cambridge University Press. Cambridge Books Online.

Chandola, T., D. Coleman, and R. W. Hiorns (1999). Recent european fertility patterns: Fitting curves to 'distorted' distributions. *Population Studies 53*, 317–329.

Cheung, S. L. K., J.-M. Robin, E. J.-C. Tu, and G. Caselli (2005). Three Dimensions of the Survival Curve: Horizontalization, Verticalization, and Longevity Extensions. *Demography 42*, 243–258.

Congdon, P. (1993). Statistical graduation in local demographic analysis and projection. *Journal of the Royal Statistical Society: Series A 156*(2), 237–270.

de Beer, J. and F. Jansen (2014). The nidi mortality model. a new parametric model to describe the age pattern of mortality. Technical report, NIDI Working Paper.

Dellaportas, P., A. F. M. Smith, and P. Stavropoulos (2001). Bayesian analysis of mortality data. *Journal of the Royal Statistical Society. Series A (Statistics in Society) 164*(2), 275–291.

Elal-Olivero, D., H. Gómez, and F. Quintana (2009). Bayesian modeling using a class of bimodal skew-elliptical distributions. *Journal of Statistical Planning and Inference 139*(4), 1484–1492.

Gage, T. B. and C. J. Mode (1994). Some laws of mortality: How well do they fit? *Human Biology 65*(3), 445–461.

Goldstein, J. (2011). A secular trend toward earlier male sexual maturity: Evidence from shifting ages of male young adult mortality. *PLoS ONE 6*(8), e14826.

Gompertz, B. (1825). On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies. *Phil. Trans. R. Soc. 115*, 513–585.

Heligman, L. and J. Pollard (1980). The age pattern of mortality. *Journal of the Institute of Actuaries 107*(1), 49–80.

Holt, C. C. (1957). Forecasting trends and seasonals by exponentially weighted moving averages. ONR Research Memorandum, Carnegie Institute of Technology 52.

Human Mortality Database (2013). University of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany). Available at www.mortality.org or www.humanmortality.de (data downloaded on November 28, 2013).

Makeham, W. (1860). On the law of mortality and the construction of annuity tables. *J. Inst. Actuaries and Assur. Mag. 8*, 301–310.

Mazzuco, S. and B. Scarpa (2015). Fitting age-specific fertility rates by a flexible generalized skew normal probability density function. *Journal of the Royal Statistical Society: Series A*.

Peristera, P. and A. Kostaki (2007). Modelling fertility in modern populations. *Demographic Research 16*, 141–194.

Rocha, G. H., R. H. Loschi, and R. B. Arellano-Valle (2013). Inference in flexible families of distributions with normal kernel. *Statistics: A Journal of Theoretical and Applied Statistics 47*(6), 1184–1206.

Rogers, A. (1986). Parametrized multistate population dynamics and projection. *Journal of the American Statistical Association 81*, 48–61.

Schmertmann, C. P. (2003). A system of model fertilty schedules with graphically intuitive parameters. *Demographic Research 9*, 82–110.

Sharrow, D. J., S. J. Clark, M. A. Collinson, K. Kahn, and S. M. Tollman (2013, December). The age pattern of increases in mortality affected by hiv: Bayesian fit of the heligman-pollard model to data from the agincourt hdss field site in rural northeast south africa. *Demographic Research 29*, 1039–1096.

Siler, W. (1979). A competing-risk model for animal mortality. *Ecology 60*(4), 750–757.

Siler, W. (1983). Parameters of mortality in human populations with widely varying life spans. *Statistics in Medicine 2*, 373–380.

Wheldon, M. C., A. E. Raftery, S. J. Clark, and P. Gerland (2013). Estimating demographic parameters with uncertainty from fragmentary data. *Journal of the American Statistical Association 108*, 795–822.

Winters, P. R. (1960). Forecasting sales by exponentially weighted moving averages. *Management Science 6*, 324–342.

Appendix A: Estimates on USA 1933–2010

Figure 3: mixture parameter of model (6) and values of child mortality q_{0-5}

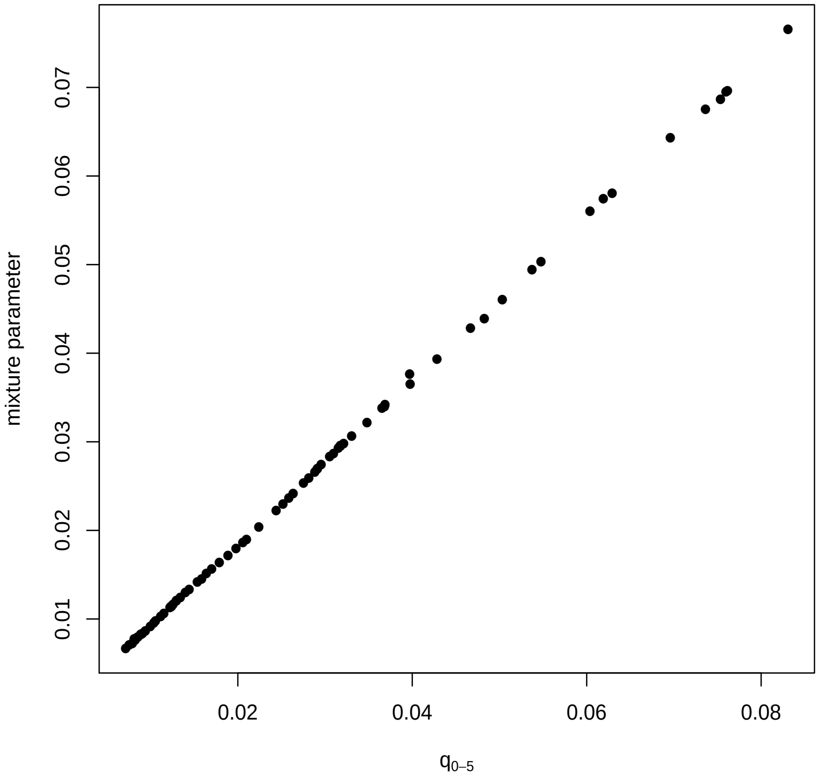


Figure 4: Half normal scale parameter of model (6) and values of q_{1-5}

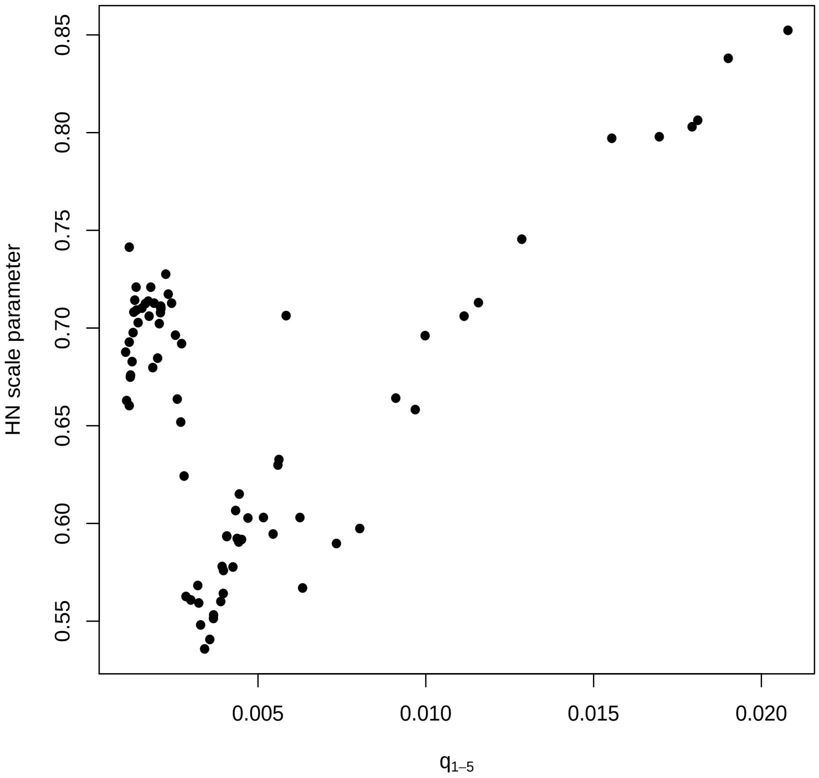


Figure 5: Skew bimodal normal scale parameter of model (6) and values of adult mortality q_{5-50}

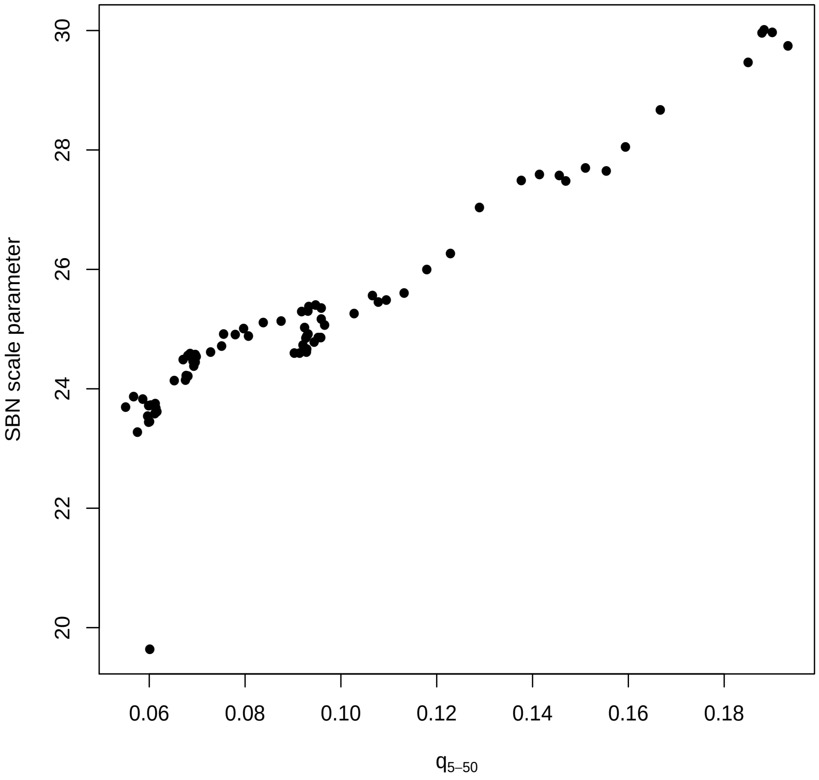


Figure 6: Skew bimodal normal location parameter of model (6) and values of life expectancy at 10 e_{10}

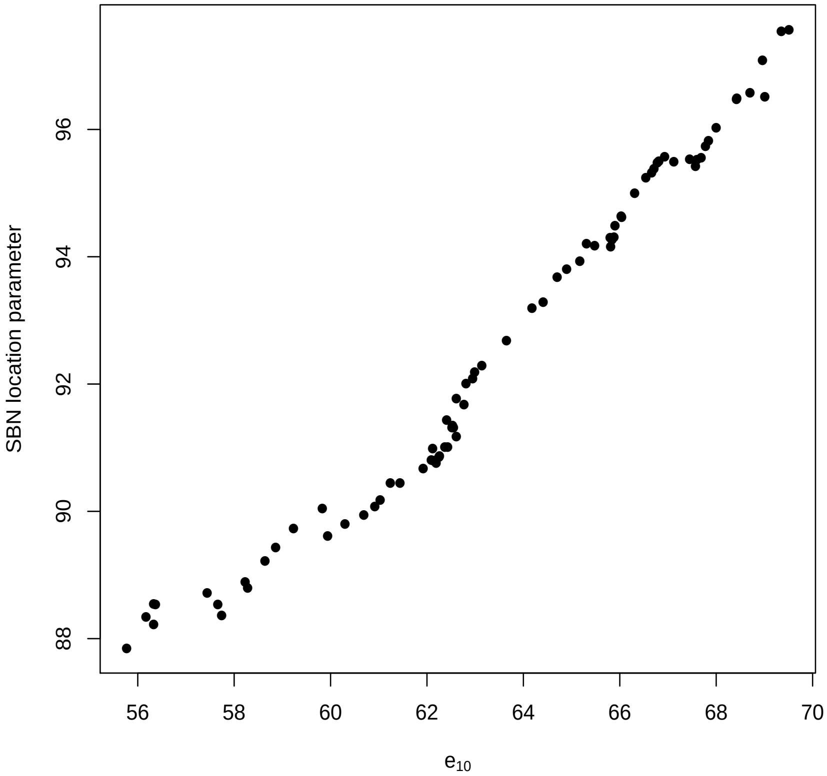


Figure 7: Skew bimodal normal skewness parameter of model (6) and share of adult (over 50) deaths occurring after modal age.

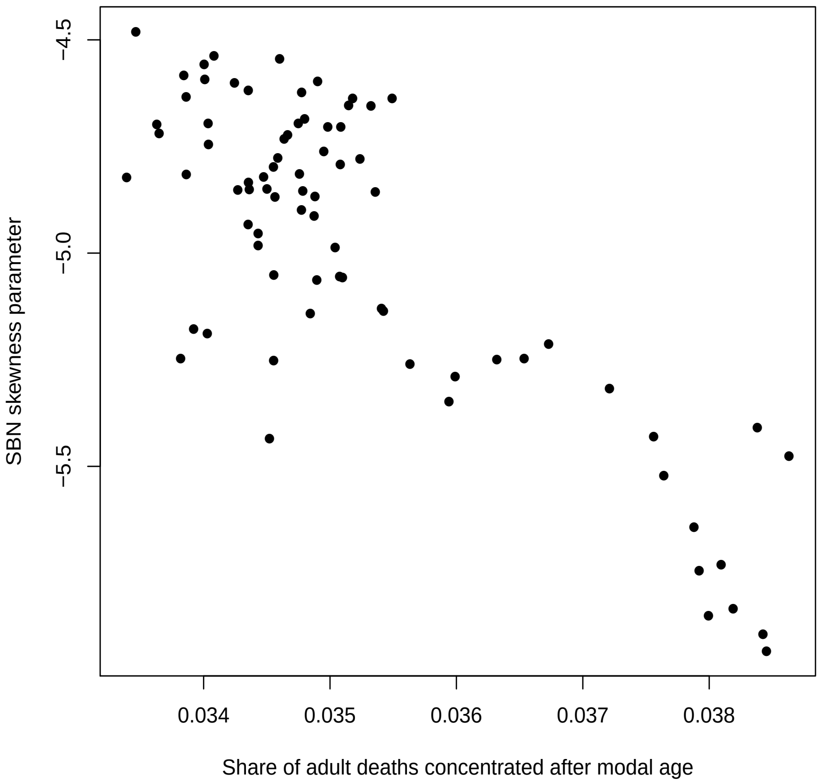
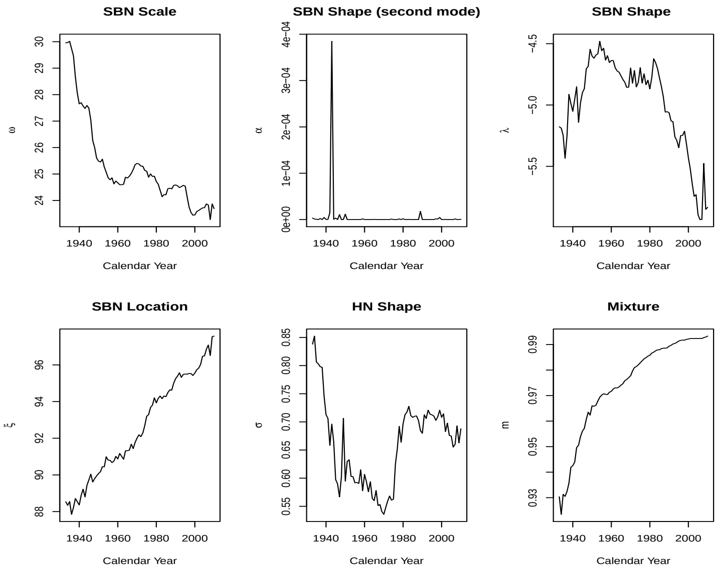


Figure 8: Mixture model parameters estimates. USA 1933–2010



Appendix B: Siler and Heligman-Pollard models likelihood and estimates

Siler and Heligman-Pollard models have been also estimated via maximum likelihood using multinomial formula in (7), where probabilities p_x are determined by the following formula for Siler model:

$$\begin{aligned} p_x &= S(x; a_1, b_1, a_2, a_3, b_3) - S(x+1; a_1, b_1, a_2, a_3, b_3) \qquad (9) \\ S(x; a_1, b_1, a_2, a_3, b_3) &= \exp\left(-\frac{a_1}{b_1} \cdot (1 - \exp(-b_1 x))\right) \cdot \exp(-a_2 t) \cdot \exp\left(\frac{a_3}{b_3} \cdot (1 - \exp(b3 * t))\right) \end{aligned}$$

The results of such a model are reported in table 3.

Table 3: Parameters estimate for Siler model in 4 countries (bootstrapped standard errors between parentheses)

Country	a_1	b_1	a_2	a_3	b_3
Sweden (1781)	0.447 (0.050)	0.833 (0.161)	16.487 (0.108)	1.947 (0.081)	0.271 (0.021)
Sweden (2010)	0.042 (0.028)	0.465 (0.413)	19.884 (0.402)	0.000 (0.045)	-6.480 (0.336)
Hungary (2009)	0.131 (0.046)	0.357 (1.269)	18.773 (0.824)	0.449 (0.143)	-10.000 (0.050)
Agincourt (2002)	0.078 (0.346)	0.928 (0.134)	15.515 (0.383)	2.110 (0.300)	0.131 (0.067)

The p_x for Heligman-Pollard model is defined by

$$\begin{aligned} p_x = S(x; A, B, C, D, E, F, G, H) - S(x+1; A, B, C, D, E, F, G, H) \qquad (10) \\ S(x; A, B, C, D, E, F, G, H) = \exp\left(\int_0^x q(t; A, B, C, D, E, F, G, H) dt\right) \end{aligned}$$

where $q(t; A, B, C, D, E, F, G, H)$ is defined by (1). Results of estimates are reported in table 4

Table 4: Parameters estimate for Heligman-Pollard model in 4 countries (bootstrapped standard errors between parentheses)

Country	A	B	C	D	E	F	G	H
Sweden (1781)	0.126 (0.003)	0.000 (0.022)	0.737 (0.052)	0.009 (0.000)	0.000 (0.025)	15.263 (0.979)	0.000 (0.000)	1.102 (0.001)
Sweden (2010)	0.001 (0.000)	0.015 (0.008)	0.280 (0.005)	0.001 (0.000)	0.971 (0.147)	55.000 (0.253)	0.000 (0.000)	1.141 (0.001)
Hungary (2009)	0.003 (0.001)	0.000 (0.084)	0.671 (0.010)	0.005 (0.001)	18.652 (1.290)	56.987 (1.114)	0.000 (0.000)	1.103 (0.018)
Agincourt (2002)	0.025 (0.002)	0.000 (0.056)	0.318 (0.036)	0.018 (0.001)	1.561 (0.340)	51.221 (1.709)	0.000 (0.000)	1.172 (0.005)