



Forecasting and simulating mortality tables

Miklós Arató^a, Dávid Bozsó^{a,b}, Péter Elek^{a,c}, András Zempléni^{a,*}

^a Department of Probability Theory and Statistics, Eötvös Loránd University, Budapest, Pázmány P. sétány 1/C. H-1117, Hungary

^b ING Hungary Insurance Pte. Co. Ltd., Budapest, Dózsa György út 84/b. H-1068, Hungary

^c Ministry of Finance, Budapest, Józsefnádor tér 2-4. H-1051, Hungary

ARTICLE INFO

Article history:

Received 27 December 2007

Accepted 25 January 2008

Keywords:

Azbel-model

Forecasting

Hypothesis testing

Mortality table

Simulation

ABSTRACT

In this paper we suggest solutions to the actuaries, facing the problem of estimating future mortality tables, especially in cases where there is a lack of relevant data and where the tendencies are not easy to estimate directly. We propose the utilization of external sources of information in the form of other, published mortality tables and use formal statistical tests to decide among these possible candidates. The procedure can also be applied for checking e.g. the goodness of mortality selection factors. We suggest the use of parametric families in modelling; for example the simple 2-parameter Azbel model. We conclude the paper by a simulation study which allows for the quantification of the possible risks related to unforeseen changes in the mortality tables in the future. To calibrate the variances of these models, initial estimates are needed, which we get by the Lee–Carter method.

© 2008 Elsevier Ltd. All rights reserved.

1. Introduction

An actuary regularly faces the problem of calculating annuities or insurance premiums, for which the exact values depend on future changes in mortality tables. As this is not known at the moment of the calculations, one usually uses the latest available tables for the given (or an analogous) population.

This problem is especially challenging in the new EU member states, which have recently experienced quick changes in mortality patterns and one may expect the changes to continue. As the level of the health care as well as the health-consciousness lag behind the current western European or American situation, it is logical to assume a development analogous to what was observed in these countries. The application of this method can be summarized as follows:

1. Find a historical mortality table that resembles well the current table of the population
2. The future development of the mortality table is supposed to be analogous to the development of the historical table, which is available (at least for a certain time period)
3. Several possible scenarios may be simulated alongside the evolution of the main table of point 2, allowing for the investigation of longer series as well as the estimation of the variance of the parameter under investigation (annuity, premium etc.)

The suggested methods can be applied in two main situations. In the first case there is enough data for the important age groups. Then the suggested method of analogy can be a relatively simple procedure for checking the classical demographic or more modern forecast methods such as the Lee–Carter method [1]. Such a situation is typical in the case of population mortality.

* Corresponding author.

E-mail addresses: aratonm@ludens.elte.hu (M. Arató), bozso.david@ing.hu (D. Bozsó), elekpeti@cs.elte.hu (P. Elek), zempleni@ludens.elte.hu (A. Zempléni).

Another – probably more important – area of application is the case when we have only a short period of observations and these do not cover the whole range of ages. In Hungary (and several other countries in Central Europe) this is true for insurance companies providing life insurance or annuities. The private insurance companies started their operation in the early 1990s and they have sold hardly any annuities since then. Pension funds have large membership, but the financial services itself will start only in the next couple of years. So there is no mortality data of the over 60s available for the insurance companies. The proposed method of analogy helps us to model the mortality of future annuitants without any previous local experience.

The paper is organized as follows. In Section 2 we investigate the similarity measures for mortality tables. Then we propose a new method for hypothesis testing based on these measures in Section 3. Section 4 deals with fitting parametric models, while in Section 5 we investigate forecasting and simulation.

2. The method of analogy

2.1. Measures for similarity

There is an extensive database of mortality tables at the homepage of the Society of Actuaries¹ but one may use other sources as well. So the main question to be answered here is how we should measure the similarity between mortality tables. The first attempt to test the similarity was the work of Walsh [2].

Let us denote the table, which we want to estimate by q_1 and its elements by q_{i1} . The first statistics we propose is the weighted quadratic deviation (*QDEV*):

$$QDEV = \sum_{i=K}^N \frac{T_i(q_{i1} - q_{i0})^2}{q_{i0}}, \quad (1)$$

where T_i is the number of exposures in year i and q_0 is the table chosen as a candidate for the approximator. (Its elements are denoted by q_{i0} .) The starting and finishing years K and N can be chosen appropriately to the problem we want to solve and to the available data.

In formula (1) the weights are selected proportionally to the reciprocal of the approximate variance q_{i0}/T_i of the estimator q_{i1} under the null hypothesis that the true mortality rate is q_{i0} . Thus the limit distribution is chi-squared, so a statistical test for the equality of the two tables can be based on the critical values for the chi-squared distribution with $N - K + 1$ degrees of freedom. However, for population tables these values are always statistically significant. But this is not exactly the same as practical significance, which would mean that the two tables are not exchangeable for practical calculations. The following measures are more suitable for the latter purpose.

The first alternative is the so-called *A/E* statistics, defined as

$$A/E = 100 \cdot \frac{\sum_{i=K}^N l_{i0} q_{i1}}{\sum_{i=K}^N l_{i0} q_{i0}}, \quad (2)$$

where l_{i0} is the probability to be alive at age i according to the base table, i.e. $l_{i+1,0} = l_{i0}(1 - q_{i0})$ and $l_{K0} = 1$. (This and the next statistics can be found in the paper of Mitchell and McCarthy [3].) Thus the summands in the denominator are proportional to the estimated number of deaths in the given year in the population, while the numerator gives the similar quantity, based on the population distribution of the base table and the risks defined by the investigated table. This can also be interpreted as a Laspeyres index of the two sets of probabilities q_i , taking weights from the basis population.

The expected remaining lifetime can be defined as follows:

$$ERL = 100 \cdot \frac{\sum_{i=K}^N l_{i1} - 0.5}{\sum_{i=K}^N l_{i0} - 0.5}, \quad (3)$$

where $l_{K0} = l_{K1} = 1$. (3) gives the proportion of the lifetimes (starting from year K and ending at year N) in the two tables. The multiplier 100 allows for an expression of the ratio in percentages.

In spite of the fact that often the main interest (especially for annuities) lies in the mortality for higher ages, we suggest to apply the above coefficients for the choice $K = 20$ or $K = 30$ and $N = 60$ or 70 as a typical insurance company has the bulk of its data from this interval. In our applications we have actually chosen $K = 30$ and $N = 70$. The interested reader may find further test statistics in the well-known book of Benjamin and Pollard [4].

¹ <http://www.soa.org/ccm/content/areas-of-practice/special-interest-sections/computer-science/table-manager/>.

Table 1

The values of different statistics for approximating the Hungarian male and female mortality of year 2000 with US mortalities of the past

Hungarian table	US table	QDEV (30, 70)	A/E (30, 70)	ERL (30, 70)	ERL1 (20, 70)	ERL2 (45, 70)	ERL3 (60, 70)
Male, 2000	Male, 1950	185 019	95.64	98.70	97.86	99.70	100.50
Female, 2000	Female, 1970	29 755	101.81	105.23	103.02	107.88	107.10

The numbers below the names of the statistics indicate the lower and upper end points, K and N .

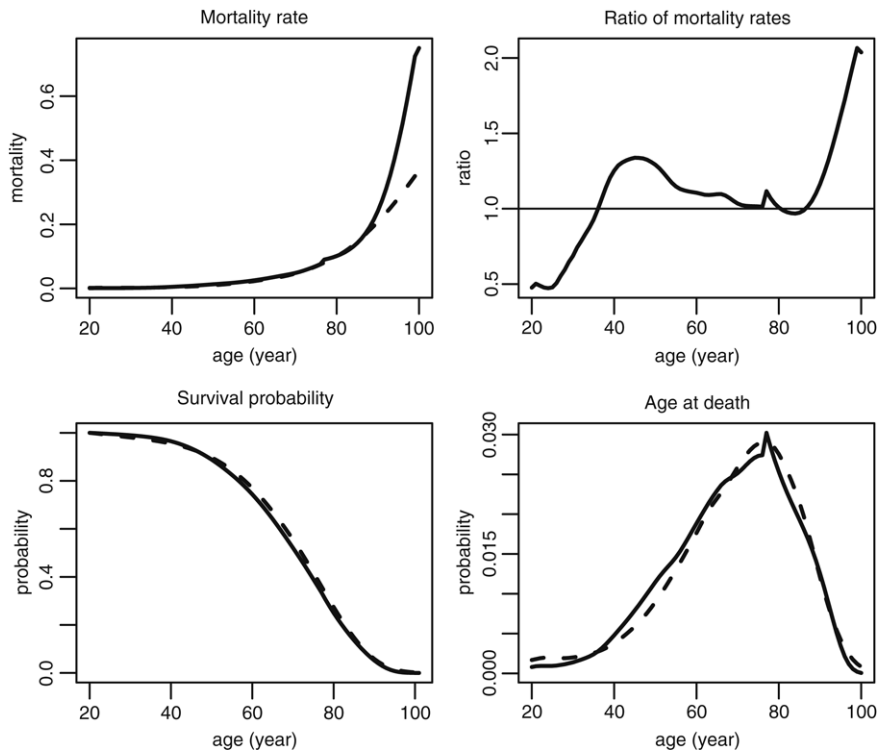


Fig. 1. Comparison of Hungarian male mortality (year 2000, solid) and US male mortality (1950, broken).

Table 2

Critical values for the different statistics, based on the Hungarian male mortality of year 2000, $K = 30$ and $N = 70$

Quantiles	0.005	0.05	0.95	0.995
QDEV	0	0	55.58	66.08
A/E	98.10	98.77	101.21	101.92
ERL	99.76	99.85	100.15	100.23

2.2. Applications

Here we show the application of the above methods to the case of the Hungarian population mortality tables as compared to past US tables. The fit here is usually not as good as for annuitant tables of insurance companies, but we still get reasonable fits.

Fig. 1 displays the comparison of Hungarian male mortality to US male mortality in 1950. We see – especially from the upper right panel – that the Hungarian data show lower mortality between 20 and 40 years, while display higher mortality between 40 and 60 and beyond 90 years. However, between the ages of 60 and 90 the fit looks very good – and for example in case of pension calculation these are the most important ages. The values of the statistics are given in the first row of Table 1.

Similar comparisons for the female mortality show a smaller time lag between the Hungarian and the American tables, here the best fit is given by the 1970 US mortality data. (See Fig. 2 for the graphs and the second row of Table 1 for the statistics.)

Table 2 shows the critical values of the above statistics developed with the method of Section 3 for the Hungarian male mortality table. Of course, none of the statistical tests accepts formally the fit of the US table but this was to be expected because of the huge number of observations (Hungary has a population of about 10 millions). Hypothesis testing – especially for real-life insurance portfolios – is examined in detail in the next section.

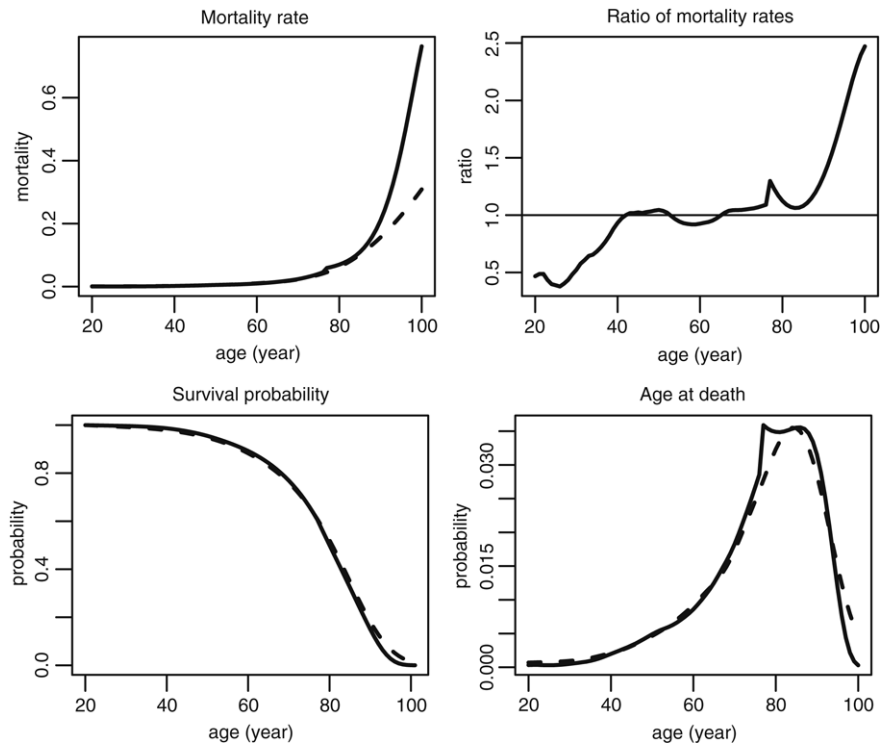


Fig. 2. Comparison of Hungarian female mortality (year 2000, solid) and US female mortality (1970, broken).

Table 3

95% confidence intervals of the test statistics in case of two different sample sizes and age distributions

Statistics	Distributed as the insured population		With equal population in each age	
	Lower	Upper	Lower	Upper
50 000 exposed people at ages 30–70				
QDEV	0	58.14	0	57.98
A/E	67.11	142.36	91.13	108.77
ERL	98.45	101.49	98.90	101.07
500 000 exposed people at ages 30–70				
QDEV	0	56.46	0	56.14
A/E	88.59	112.28	97.26	102.79
ERL	99.53	100.48	99.65	100.34

3. Hypothesis testing

In this section, we analyze the possibility of using the introduced statistics for hypothesis testing. We may use the above statistics for formal tests as follows. Let us suppose that the true mortality is given by the reference table q_0 . In the case when the data is considered as a realization of a sample we would like to decide whether it could originate from this table. In order to create an empirical test, we take simulation results 10 000 times from the same population, based on mortalities from the reference table. The 10 000 values of the statistics allow us to calculate critical values with high accuracy. The advantage of this procedure is that it is not an asymptotic result but an exact one using the age distribution of our population.

To see the effects of the particular form of the age distribution and of the population size on the critical values of the test statistics, we calculated 95% confidence intervals for QDEV, A/E and ERL assuming either a uniform age distribution or the age distribution of a real-world insured portfolio (see Fig. 3 about the applied distribution), and choosing the population size as either 50 000 or 500 000 people, respectively. Table 3 shows that the lengths of the confidence intervals are much shorter for the uniformly distributed case and also for the larger population case.

We may also examine the accuracy of the asymptotic confidence intervals of the test statistics. As $K = 30$ and $N = 70$, QDEV is asymptotically distributed as χ^2 with 41 degrees of freedom and thus the asymptotic 95%-confidence interval is (0, 56.94), which is notably shorter than the confidence intervals for the cases of 50 000 sample size in Table 3. Generally, the deviation from the asymptotic result is increasing when the number of insured people is decreasing or the age distribution becomes non-uniform.

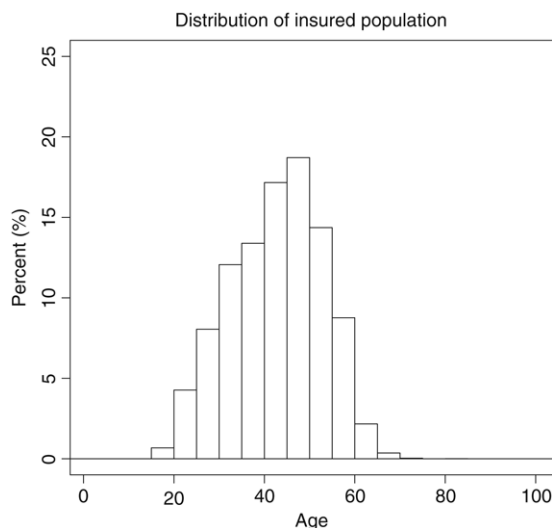


Fig. 3. The age distribution of the insured population.

Table 4
95% normal confidence intervals for the A/E statistic

Number of exposed people	Distributed as the insured population		With equal population at each age	
	Lower	Upper	Lower	Upper
50 000	99.62	100.38	99.91	100.09
500 000	99.88	100.12	99.97	100.03

In principle, one may use a normal approximation for the A/E statistics. Then, the approximate confidence interval at level $(1 - \alpha)$ is the following:

$$\left(100 - 100 \cdot u_{\alpha/2} \sqrt{\frac{\sum_{i=K}^N l_{i0}^2 q_{i0} (1 - q_{i0}) / T_i}{\left(\sum_{i=K}^N l_{i0} q_{i0}\right)^2}}, 100 + 100 \cdot u_{\alpha/2} \sqrt{\frac{\sum_{i=K}^N l_{i0}^2 q_{i0} (1 - q_{i0}) / T_i}{\left(\sum_{i=K}^N l_{i0} q_{i0}\right)^2}} \right).$$

Table 4 shows that there is a huge difference between these asymptotic intervals and the intervals obtained by simulation in Table 3. Therefore, especially when the population size is not large enough, we do not recommend to use the approximations based on the asymptotic results for either QDEV or A/E.

The simulation method also allows us to analyze the power of our tests when the true mortality is the hypothetical mortality multiplied by a constant. We estimated the power using again 10 000 iterations. It can be seen from Fig. 4 that both for the portfolios with 50 000 and 500 000 insured people the power function at the uniform age distribution differs very much from the function at the insured age distribution. We can also observe that the QDEV and A/E statistics are not reliable enough for the smaller portfolio with the insured age distribution, while the ERL statistics yields generally the best results.

4. Parametric models

The aim of this section is to model mortality with fitted parametric tables in order to forecast its development by these models. We propose to model the dynamics of the parameters instead of using mortality itself, this being much simpler while providing reasonable flexibility. The following questions arise naturally:

- which parametric model to use,
- how to estimate the parameters and
- how to measure the goodness of fit.

The last question has already been answered in the previous section, where we introduced a wide choice of statistics. The traditional answer for the first question is the Gompertz–Makeham-model, which is a well-known, more than a century old solution. See Gompertz [5] for its first publication, and Makeham [6]. We also present an alternative, known as the

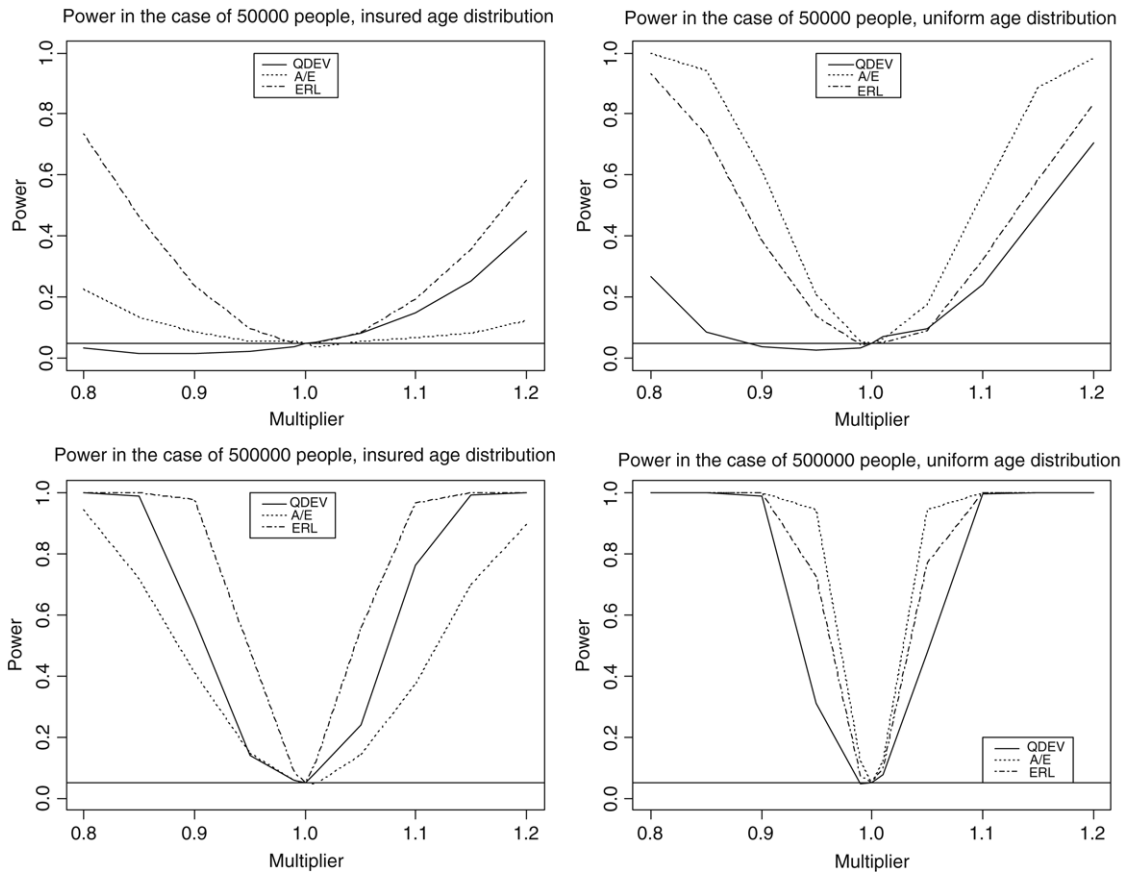


Fig. 4. The power of the test for different population sizes and age distributions.

Azbel-model [7], which is in most cases as good and even simpler, with readily interpretable parameters. It is given by the formula

$$q_x = Ab \exp[b(x - X)]$$

where A , b and X are the parameters. Although this formula is already simpler than the Gompertz–Makeham model we can further simplify it by a reparametrization. Let $T = X - \ln(Ab)/b$, then we have $q_x = \exp(b(x - T))$. This model will be our main tool in forecasting. Its parameters are the following: b is the shape parameter and parameter T is the end point as $q_T = 1$ (at this age the population dies out).

How to fit these models? An initial approach would be to use unweighted or weighted least squares estimates where the minimization is done by numerical methods and in the weighted case the weights can be \hat{l}_x/\hat{q}_x [4]. However, it turned out that better results are obtained in practice when we also approximate the expected remaining lifetimes at each age, i.e. we minimize

$$\sum_{x=K}^N \frac{\hat{l}_x}{\hat{q}_x} \left[(\hat{q}_x - q_x)^2 + (\hat{e}_x - e_x)^2 \right] \tag{4}$$

where q_x and e_x are the mortality and expected remaining lifetime at age x of the parametric model, and \hat{q}_x and \hat{e}_x denote the corresponding empirical values.

Another, simpler procedure in the case of the Azbel-model is a loglinear regression-type estimator. Using the fact that $\log q_x = b(x - T)$, a simple (or weighted) linear regression can be performed with the logarithm of q_x as the dependent and age as the independent variable. Then the regression coefficients can be directly used to obtain the parameters of the model.

Azbel-models for the q_x 's and for the l_x 's were fitted with these methods and in order to check the differences between the approaches we have calculated the q_x 's from l_x 's and vice versa and plotted them. The corresponding Gompertz–Makeham model was only fitted by minimizing expression (4). As our main aim is to model the death rates for pensioners, all models were estimated on ages between 60 and 90.

Fig. 5 displays some fitted Azbel and Gompertz–Makeham models, while Table 5 shows the calculated A/E and ERL statistics of the previous section both for ages 15–105 and for 60–90. It turns out that the weighted fit with expected

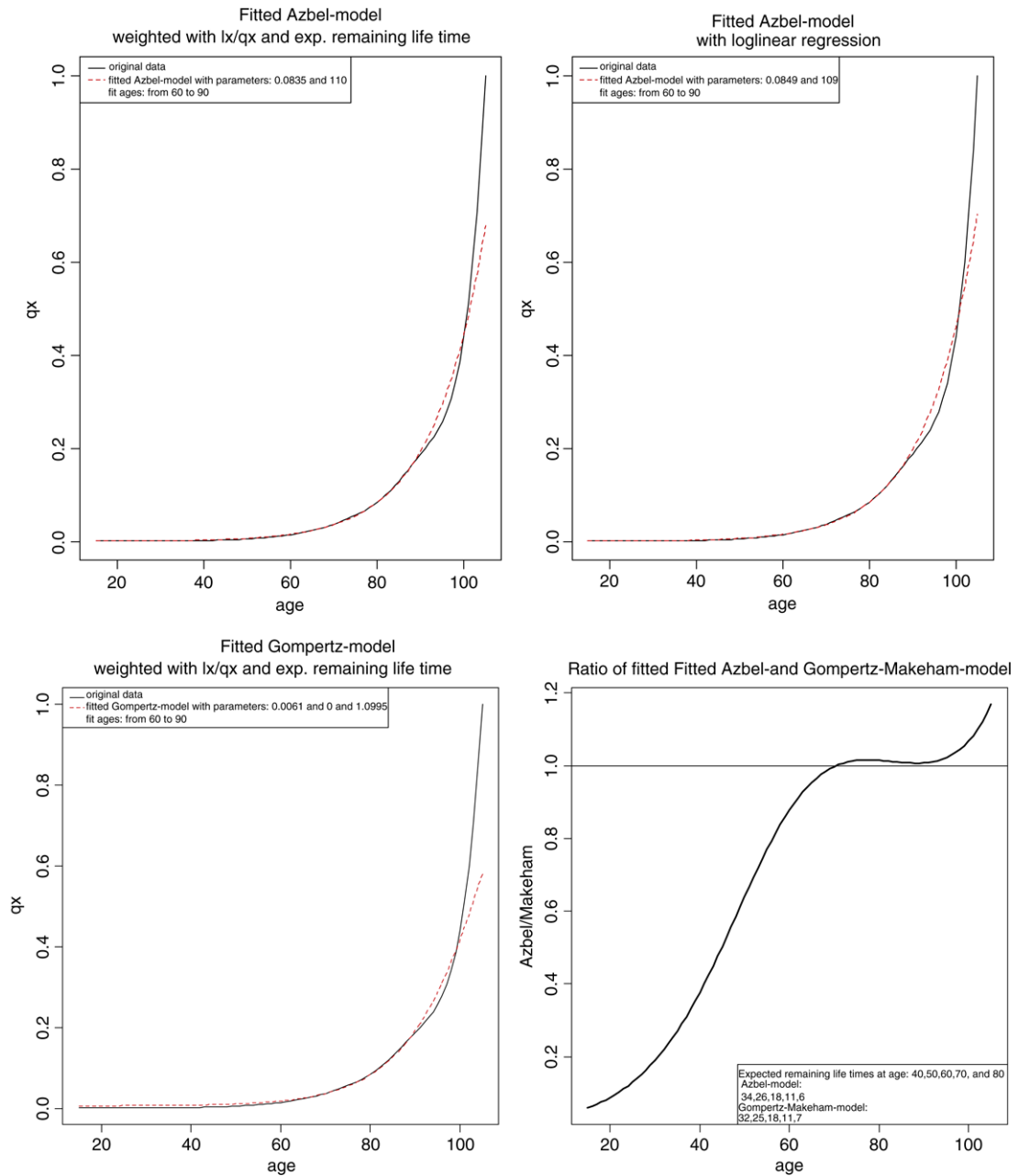


Fig. 5. Fitted Azbel- and Gompertz-models.

Table 5

A/E and ERL statistics corresponding to the models fitted to ages 60–90

Method	Test interval	A/E	ERL
Weighted with l_x/q_x and expected remaining lifetime (Gompertz-model)	Ages 15–105	102.82	96.94
	Ages 60–90	100.54	99.02
Weighted with l_x/q_x and expected remaining lifetime (Azbel-model)	Ages 15–105	105.24	97.97
	Ages 60–90	100.04	99.84
Loglinear regression (Azbel-model)	Ages 15–105	104.35	98.28
	Ages 60–90	99.51	100.21
Weighted loglinear regression (Azbel-model)	Ages 15–105	105.29	98.89
	Ages 60–90	102.12	99.66

Table 6
95% confidence intervals of the test statistics in case of various population sizes (fit for ages 60–90)

Population size	A/E		ERL	
	Lower	Upper	Lower	Upper
1 000	94.23	106.56	96.83	102.99
2 000	95.49	104.45	97.83	102.33
4 000	97.13	103.29	98.52	101.63
8 000	97.87	102.23	98.88	101.11
16 000	98.44	101.60	99.21	100.80

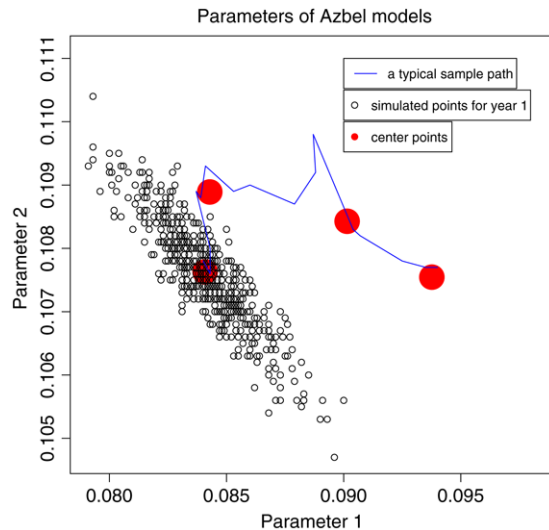


Fig. 6. The fitted parameters of the Azbel model using Canadian male mortality tables for 4 consecutive years (big dots), 500 simulated values for the first year (small circles) and a typical sample path (line).

remaining lifetime (i.e. minimizing (4)) and the simple loglinear regression give very similar results, while the weighted loglinear regression is slightly worse than the above ones.

Just to see the significance of the deviations, Table 6 shows some critical values of the test statistics corresponding to the 95% confidence level in case of various population sizes. The simulations show that the best fitting model is acceptable even if the population size is 16 000.

5. Forecast of development, simulations

Now we are in a position of estimating the future development of the mortality table under investigation. One possibility is the well-known Lee–Carter method. However, especially in Central Europe, where the substantial recent changes in the society resulted in quick and unpredictable changes in the mortality patterns, this may not be the best solution. Thus, we also present an alternative method, based on the analogy approach of above.

The essence of the procedure is that the mortality table in the future develops similarly to how the analogous mortality table developed in the past. Furthermore, we assume that the parameters of the table projected this way for a given year are only mean values, and the parameters defining realistic scenarios can be written as $\theta_t = \mu_t + \varepsilon_t$, where μ_t is the expected parameter-vector for the time point t (the parameters of the Azbel model that provides the best fit to the forecasted mortality table). ε_t is supposed to be a 2-dimensional normally distributed i.i.d. vector sequence with parameters $(0, \Sigma)$. We only have to estimate its covariance matrix Σ .

In order to do so we had to explore the relation between the parameters for real (and forecasted) examples. First we had to find a large dataset, useful as a training set. The Hungarian population mortality from 1949 and its extrapolation by the Lee–Carter model [8] to 2040 was chosen for this purpose. We used this data set to find out the possible dynamics of the parameters. The estimator of Σ was taken as the empirical covariance matrix for the parameters of this population mortality. This can be considered as an upper estimate of the variance component, as it transforms the possible trend to variance as well. On the other hand, it does not include the uncertainties in the parameter estimators, so we suggest its use as a simple tool.

Another problem to solve was that the fitted annuitant tables were given for every fifth year, but we intended to estimate the tables annually. For the $t + k$ th year ($0 < k < 5$) we simply used $\mu_t + (\mu_{t+5} - \mu_t) * k/5$ as the mean value for that year.

The advantage of our model is that it incorporates the natural relationships between the successive years. For instance, Fig. 6 shows the parameters of the Azbel-models fitted to four consecutive years of Canadian male mortality tables as big

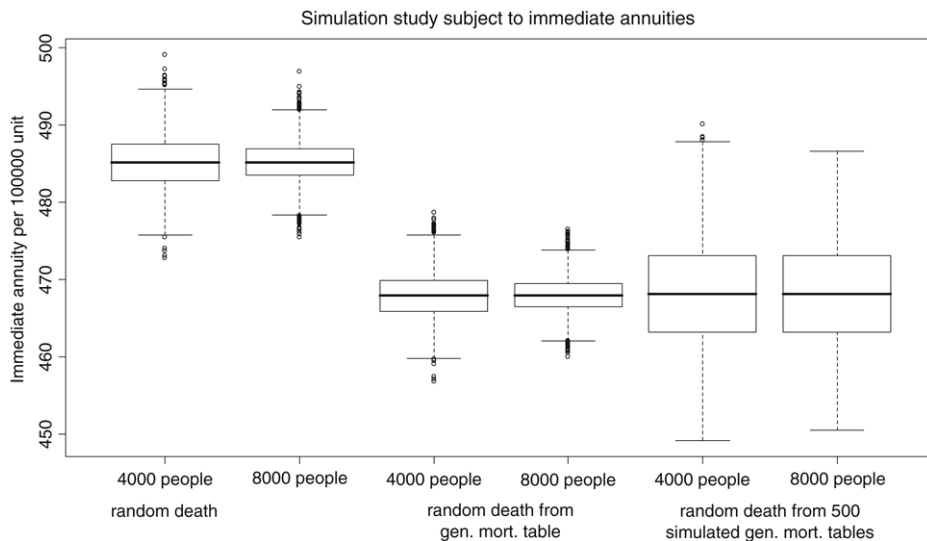


Fig. 7. Monthly annuities for simulated samples: random deaths from a given mortality table (1st and 2nd columns), time-varying tables (3rd and 4th columns) and simulated random development of the tables (5th and 6th columns).

dots, while the small circles represent the 500 simulated values for the first year, and the line displays a typical path of the parameters. The spread of the small circles reflects the strong negative correlation observed between the estimated parameters.

With our approach we can examine how the various sources of uncertainty affect the estimated monthly annuities (in units for 100 000 units). The first two boxes in Fig. 7 display the consequences of the random times of deaths for two different population sizes for the Canadian male mortality table of 1958. The 3rd and 4th boxes show the additional effect of the development of mortality in time (Canadian male mortality from 1958 to 1988). Due to the gradual improvement in health, the expectation of the annuity distribution is decreasing. Also, one can see on the 5th and 6th boxes that the variation of the distribution dramatically increases when the uncertainty of parameter evolution is also taken into account.

References

- [1] L. Carter, R.D. Lee, Modeling and forecasting US mortality: Differentials in life expectancy by sex, *International Journal of Forecasting* 8 (3) (1992) 393–412.
- [2] J.E. Walsh, Some probability results for mortality rates based on insurance data, *Journal of the Institute of Actuaries* 79 (1953) 213–220.
- [3] O.S. Mitchell, D.G. McCarthy, Estimating international adverse selection in annuities, *North American Actuarial Journal* 6 (4) (2002) 38–54.
- [4] B. Benjamin, J.H. Pollard, *The Analysis of Mortality and other Actuarial Statistics*, The Institute of Actuaries and Faculty of Actuaries, Oxford, 1993.
- [5] B. Gompertz, On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies, *Philosophical Transactions of the Royal Society of London* 115 (1825) 513–585.
- [6] W.M. Makeham, On the law of mortality, *Journal of the Institute of Actuaries* 13 (1860) 325–358.
- [7] M.Y. Azbel, Empirical laws of survival and evolution: Their universality and implications, *Proceedings of National Academy of Sciences* 96 (1999) 15368–15373.
- [8] S. Baran, J. Gáll, M. Ispány, G. Pap, Forecasting Hungarian mortality rates using the Lee–Carter method, *Acta Oeconomica* 57 (2007) 21–34.