Projecting Mortality Trends: Recent Developments in the United Kingdom and the United States

Carlos Wong-Fupuy* and Steven Haberman†

Abstract

The sustained reduction in mortality rates and its systematic underestimation has been attracting the significant interest of researchers in recent times because of its potential impact on population size and structure, social security systems, and (from an actuarial perspective) the life insurance and pensions industry worldwide. Despite the number of papers published in recent years, a comprehensive review has not yet been developed.

This paper attempts to be the starting point for that review, highlighting the importance of recently published research—most of the references cited span the last 10 years—and covering the main methodologies that have been applied to the projection of mortality rates in the United Kingdom and the United States. A comparative review of techniques used in official population projections, actuarial applications, and the most influential scientific approaches is provided. In the course of the review an attempt is made to identify common themes and similarities in methods and results.

In both official projections and actuarial applications there is some evidence of systematic overestimation of mortality rates. Models developed by academic researchers seem to reveal a trade-off between the plausibility of the projected age pattern and the ease of measuring the uncertainty involved. The Lee-Carter model is one approach that appears to solve this apparent dilemma.

There is a broad consensus across the resulting projections: (1) an approximately log-linear relationship between mortality rates and time, (2) decreasing improvements according to age, and (3) an increasing trend in the relative rate of mortality change over age. In addition, evidence suggests that excessive reliance on expert opinion—present to some extent in all methods—has led to systematic underestimation of mortality improvements.

1. Introduction

Over most of human existence, mortality has shown a decreasing pattern (Berin, Stolnitz, and Tenebein 1999; Friedland 1998, p. 49). Sharp improvements have been experienced in particular over this century (Charlton 1997; Goss, Wade, and Bell 1998).

Time series of life expectancy estimates at birth seem to suggest that, after having experienced marked improvements during the first half of this century, a law of "diminishing returns" has started to affect improvement trends.[1] Consequently the existence of a biological limit for life expectancy and the impossibility of future significant improvements have been suggested.[2]

However, Lee (1997, p. 1) has pointed out that decreasing increments in life expectancy are the result of the combination of two causes. First,

*Carlos Wong-Fupuy is a Senior Financial Analyst with A.M. Best Europe Ltd., 1 Minister Court, Mincing Lane, London EC3R 7AA, United Kingdom, e-mail: carlos_wong-fupuy@ambest.com.

†Steven Haberman, Professor of Actuarial Science and Deputy Dean, Cass Business School, City University, 106 Bunhill Row, London EC1Y 8TZ, United Kingdom, e-mail: shaberman@city.ac.uk.

[1] Comments from the 1994-based National Population Projections for the United Kingdom, supporting the decision of projecting reductions in mortality rates at a slower progression than in the past (ONS 1996, p. 10).

[2] For example, Olshansky and Hayflick, in the seminar "Impact of Mortality Improvement on Social Security: Canada, Mexico and the U.S." organized by the SOA on 30 October 1997 (Friedland 1998, pp. 52–53).

reductions in mortality rates at advanced ages (i.e., with a remaining lifetime of a few years) contribute less to life expectancy increases than equivalent reductions at lower ages (where the individual has most of their life span remaining). Second, developed countries have reached a stage where mortality improvements are concentrated at advanced ages (Charlton 1997, pp. 22–23). Therefore, even if constant reductions for mortality rates at advanced ages were to be observed, increments of life expectancy would follow a decreasing progression.[3] This corresponds to the so-called "rectangularization" of the life table.

So far there does not seem to be a consensus regarding the possibility or impossibility of observed patterns continuing to hold in the future (Thatcher 1999, pp. 6–7; Tuljapurkar and Boe 1998, pp. 24–26). U.S. experience shows that the maximum age at which a very low fraction of a population, say, 0.001%, stays alive has been increasing steadily over this century (Bell, Wade, and Goss 1992, pp. 14–15). Similar conclusions can be found from U.K. data (Thatcher 1999). In addition, U.S. population data for this century show that age-specific mortality rates follow a decreasing log-linear pattern, with no evidence of significant departures from this trend (Lee and Carter 1992).

A number of broad approaches to forecasting mortality rates exist: using models based on the underlying biomedical processes, causal models involving econometric relationships, and trend models that are extrapolative in nature. We will consider only the last category in this review.

The aims of extrapolative models have been

- To express an age-specific (real-valued) measure of mortality (e.g., expectation of life, mortality rate, force of mortality) as a function of calendar time with parameters estimated from past data.
- To develop a model for forecasting future values of the mortality measure, ideally accompanied by measures of uncertainty.

The types of extrapolative model that have been proposed and analyzed are the following:

a. Models based on the independent projection of age-specific mortality or hazard rates (Alho and Spencer 1985), including mortality reduction factor models (CMIB 1990, 1999; Willets 1999; Renshaw and Haberman 2000). These models can incorporate measures of uncertainty but may generate implausible future age patterns.
b. Relational models that associate life table measures with those from a standard life table (Brass 1971).
c. Models based on graduating mortality measures with respect to age for specific time periods, involving either two stages so that the parameters need to be projected (Congdon 1993; Forfar and Smith 1988; McNown and Rogers 1989) or one stage (Renshaw, Haberman, and Hatzopolous 1996; Sithole, Haberman, and Verrall 2000). These models may generate implausible projected mortality trends but lead to straightforward estimation of prediction errors.
d. The Lee-Carter method (Lee and Carter 1992; Lee 2000; Renshaw and Haberman 2003a), which seems to solve this apparent trade-off between plausibility of the projected age pattern and ease of measuring the uncertainty involved. As in many actuarial applications, a log-linear trend for age-specific mortality rates is often assumed for the time-dependent component. Despite criticisms based on its exclusively extrapolative nature (Gutterman and Vanderhoof 1998), this model has been used for demographic and social security applications (Lee 2000). There are also suggestions for its adoption for U.S. Social Security projections (Lee and Tuljapurkar 1997; Frees 1999).

Official projections, based on deterministic scenarios and a mixture of analysis by cause of death and expert opinion (Gallop 1998, 2002; Goss, Wade, and Bell 1998; ONS 1996; Shaw 1998), have underestimated systematically mortality trends during the last decades (Murphy 1995; Shaw 1994). Actuarial applications in the life insurance and pensions industry tend to follow type (a) and model future mortality as the product of mortality rates of a base year multiplied by a reduction factor, which follows an exponential decay. These models do not commonly allow for explicit measures of uncertainty, and

[3] In apparent contradiction to the view in note 2, Olshansky suggests a numerical exercise supporting Lee's argument (Friedland 1998, pp. 52–53).

parameters are based on a mixture of recently observed trends and expert opinion. Systematic overestimation of mortality rates has also been noted for this type of application (CMIB 1990, 1999; Willets 1999; Renshaw and Haberman 2000).

This article summarizes and compares the most important approaches proposed and used to date to project mortality rates in the United Kingdom and the United States covering the principal contributions to the literature. We have attempted to summarize the main features and conclusions of published research material. Our work could be expanded either by testing the different approaches referred to here on data from a range of countries or by extending the scope of the literature covered by including material published in other languages whose significance we may have missed.

Important common features of the U.K. and U.S. projections are (1) a log-linear relationship between mortality rates and time, (2) decreasing improvements according to age and time, and (3) an increasing trend in the relative rate of mortality change over age.

2. OFFICIAL PROJECTIONS

2.1 General Methodology

Methodologies used by official governmental agencies in the United Kingdom and the United States for national population projections have tended to include the following features in common: (1) analysis by cause of death, (2) use of expert opinion, (3) consideration of cohort effects, and (4) deterministic scenarios, usually involving a central projection along with high and low variants.[4]

In the United Kingdom, the Government Actuary's Department produces official projections every other year based on information provided by the Registrar General for each constituent country. Their use generally relates to government planning, particularly long-term financial aspects of the Social Security System. Because of this, not only the size of the population but also its structure by age and sex are primary concerns.

The projection model begins with a base population, which is then adjusted by projected trends of three demographic processes: fertility, mortality, and migration. A projection period of 40 years into the future is usually considered.

One-year age-specific central mortality rates are available for the base period, then future annual reductions for each attained age (or cohort) are assumed (see below). The assumed annual reductions as well as long-term patterns result from a mixture of data analysis, expert opinion, and analysis by cause of death.

For each age (or cohort), reduction rates—defined as the complement of the ratio between two consecutive mortality rates—during the first projected year are obtained by extrapolation from the most recently observed data.[5] For the remaining years, reduction rates decrease at a decelerating pace until reaching a long-term pattern for all ages.[6] In particular, since the 1992-based projection, such a long-term pattern is equivalent to 0.5% from 2032, halving every further 10 years. A cohort approach is applied for generations born before 1947 because of their particular pattern of mortality.

The Bureau of the Census produces official population projections in the United States. The main purpose is to analyze long-term financial aspects of the Old-Age and Survivors Insurance and Disability Insurance program. As in the case of the United Kingdom, the size of the population and its structure by age and sex are of special concern. Thus, fertility, mortality, and migration are essential components of the projection model. Trends in marriage and divorce rates are also included. However, despite the fact that mortality differentials by marital status are significant (F. Bell 1997, p. 31), projections are carried out only on an age- and sex-specific basis.

As part of these population projections, mortality receives special treatment through the construction of projected period and cohort life

[4] This section is based in the last two official population projections available at the time of writing for the United Kingdom (Gallop 1998; ONS 1996, 1999; Shaw 1998) and the United States (Bell, Wade, and Goss 1992; F. Bell 1997).

[5] For example, the 1994-based projection uses the extrapolated values of 1961–93 data from England and Wales.

[6] Willets (1999, p. 49) states that such a decreasing trend follows an exponential decay. However, the official reports consulted (ONS 1996, p. 10; Gallop 1998, p. 5) do *not* make any assertion regarding that fact.

tables. However, both versions result from a common process where only period effects are considered.

2.2 Analysis by Cause of Death

An important component of the U.S. method is an analysis by cause of death. Reductions in mortality rates for each possible combination of sex, age group, and cause are averaged from the most recent past experience. Based on these historical data and expert opinion, ultimate reduction rates are postulated for the projection period. For the first projected year, extrapolated values from previous years are applied. From that point an exponential pattern is assumed until reaching the ultimate values at a fixed calendar year for all ages.

U.S. data are generally available in the form of central mortality rates for five-year age groups. Because of this, additional interpolation techniques are applied to obtain one-year initial mortality rates. The projection for the base year consists of mortality rates by age group, sex, cause, and calendar year. From these values, cause-adjusted mortality rates for each age group and sex are calculated as weighted averages according to the structure by cause of death for a standard population.

In the 1995-based projection the causes of death are grouped in 10 categories. The period from 1968 to 1994 is used as the base for extrapolation purposes. Ultimate mortality rates are assumed to be reached in 2020 after having followed an exponential trajectory. Expressed as percentages, the reductions are classified by five-year age group, sex, and cause of death.

Notice that, when estimating actual average values, the range of ages is divided into 19 groups, each containing a range of five years (i.e., from age 0 to 94 years). For ultimate values, however, the range of ages is divided into only four groups. Table 1 shows assumed ultimate reductions by sex, age group, and cause of death for the 1995-based Social Security Administration population projections.

Causes of death normally relate to specific ranges of age during the life of each individual. Thus, having a structure including cause of death for each age group and sex means that it is possible to calculate age-sex–specific mortality rates after adjustment by cause. In an intuitive explanation, the assumption of minor advances against infectious diseases and violent causes implies small reductions in mortality for groups younger than 65. Likewise, improvements for advanced ages are expected to continue at a relatively accelerated pace, since significant future declines in degenerative diseases are assumed (F. Bell 1997).

This reasoning fits into the predominant view of epidemiological transitions, which suggests that the trend of mortality in the long run is dominated by (1) pestilence and famine, (2) pandemics and infectious diseases, and (3) chronic, degenerative, and manmade diseases (Tuljapurkar and Boe 1998, p. 21). Thus, in line with this perspective, until the first decades of the twentieth century those countries now industri-

Table 1

Ultimate Reduction Rates to Be Reached by U.S. 1995-Based SSA Population Projections

Age Group	Heart Disease	Cancer	Vascular Disease	Violence	Respiratory Disease
<15	0.6	2	0.6	0.9	2.4
15–64	1.5	0.3	1.8	0.5	0.3
65–84	1.2	0.2	1.7	0.6	0.2
85+	1.1	0.2	1.7	0.6	0.2

Age Group	Infancy	Digestive Disease	Diabetes Mellitus	Cirrhosis (Liver)	Other
<15	2.2	1.6	1.8	1.4	0.5
15–64	1.1	1.6	0.3	1.5	0.3
65–84	1.1	0.4	0.6	0.2	0.2
85+	1.1	0.4	0.6	0.2	0.2

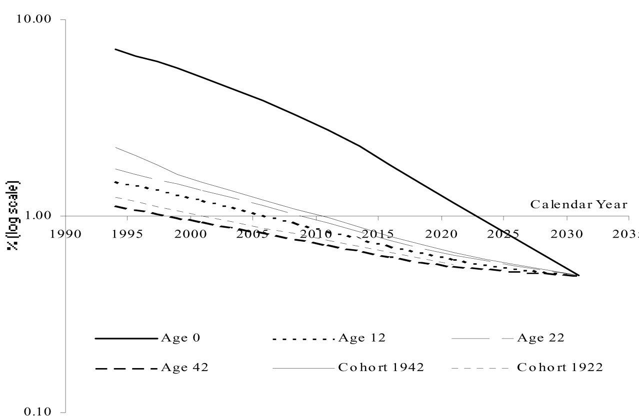
Note: For females same rates apply, except for Violence–Older than 65, where 0.6 is replaced by 0.8.
Source: F. Bell (1997, p. 24).

alized still showed some predominance of infectious diseases. Reductions in the prevalence of these causes of death have led to mortality improvements mainly focused at the earlier ages. More recent times have seen an increase in the prevalence of chronic and degenerative diseases, particularly concentrated at the advanced ages (Bell, Wade, and Goss 1992; Charlton 1997; Charlton and Murphy 1997; Goss, Wade, and Bell 1998).

Regarding an analysis by cause of death, U.K. official projections are less detailed than their U.S. counterparts. Projected reduction rates by age and sex are presented without numerical support. However, age differentials in the trends are attributable to diverse causes of death (Gallop 1998, p. 4; ONS 1996, pp. 9–10). As in the U.S. projections, reduction rates for the first year are extrapolated from most recent trends, following a nonlinear trajectory toward ultimate values (see Figure 1).

Mortality projections disaggregated by cause of death have been found in practice to be more pessimistic than those without disaggregation (Wilmoth 1995). The reason is straightforward: over time the overall trend becomes dominated by the trend for those causes with the slowest decline (or the most rapid increase). This feature is explored by Alho and Spencer (1990) and Wilmoth (1995); the latter investigates how common this pessimism is and how it might be quantified.

Figure 1
Assumed Reduction in Central Death Rates U.K. 1994-Based Population Projections (Males) Table 2.4 Values (Log Scale)



A fundamental assumption underlying the use of cause of death analyses and projections is the independence between the cause of death, as in the conventional multiple decrement model. However, the concept of dependence may be incorporated along with the partial and complete elimination of the effects of key causes of death, as investigated by Dimitrova et al. (2003).

2.3 Expert Opinion

As noted above, together with a detailed analysis of the mortality experience subdivided by age, sex, and cause of death, official projections rely also on expert opinion. Neither ultimate reduction rates nor the period after which they are reached result from any mechanical extrapolative process. Even the assumption of an exponential behavior is partly a mixture of beliefs and expertise. U.S. 1995-based official projections support the ultimate reduction values that are proposed as follows:

> Future reductions in mortality will depend upon such factors as the development and application of new diagnostic, surgical, and life-sustaining techniques, the presence of environmental pollutants, improvements in exercise and nutrition, the incidence of violence, the isolation and treatment of causes of disease, the emergence of new forms of disease, improvements in prenatal care, the prevalence of cigarette smoking, the misuse of drugs (including alcohol), the extent to which people assume responsibility for their own health, and changes in our conception of the value of life. After considering how these and other factors might affect mortality, we postulated three alternative sets of ultimate annual percentage reductions in death rates by sex, age group, and cause of deaths for the years after 2020. (F. Bell 1997, p. 24)

Similarly, U.K. 1994-based projections justify the assumed future reductions in mortality:

> It is unlikely that the trends of high reductions in mortality observed in the past at some ages will continue at the same rate indefinitely; instead it is expected that a law of diminishing returns will set in and death rates will fall less steeply in the future. . . . This transition is not assumed to take place linearly, but more rapidly at first. (ONS 1996, p. 10)

2.4 Cohort Effects

A cohort effect not yet fully explained has been observed in the United Kingdom for generations born between 1925 and 1945. This group has persistently experienced lower death rates throughout their adult lives than earlier and later generations. Being the first favored by medical advances and by the expansion of the welfare state, having a low incidence of smoking, having a tendency to have fewer children than previous generations, and experiencing a relatively low impact from the 1930s depression have been suggested as possible explanations (Charlton 1997, pp. 27–28; Willets 1999, p. 38).

A similar phenomenon is perceived for cohorts born after 1900 once they reach age 65. For this reason U.K. official projections since that based on 1992 data have incorporated a cohort approach for generations born before 1947 (Gallop 1998, pp. 4–5; 2002).

Regarding U.S. population data, McNown and Rogers (1989, p. 652) and Lee and Carter (1992, p. 660) have noted that male cohorts born between 1940 and 1960 have experienced unusually high rates of mortality at young adult ages, falling to normal levels around 35 years old. The influence of this phenomenon over life expectancy does not seem to be strong enough to alter its increasing trend.

2.5 Treatment of Advanced Ages

Inaccuracies in recording ages in official statistics and high variability in the estimates due to small exposures to risk are common problems when estimating mortality rates for the oldest age groups. Usually advanced ages are grouped into an open age interval. For projection purposes, therefore, it is necessary to apply special techniques to calculate exact rates at each age. Tuljapurkar and Boe (1998, pp. 31–32) give a brief review of recent attempts to obtain more detailed estimates from U.S. population data.

In U.S. official projections, mortality rates are available only until age 94. Therefore, some extrapolation technique is necessary for older ages. For these purposes mortality rates are assumed to increase at a constant rate of 5% from age 100. For intermediate ages between 94 and 100, factors representing the increase from q_{x-1} to q_x are assumed to follow a linear segment starting at (q_{94}/q_{93}) and ending at 1.05. For females, the ultimate increase factor is 1.06 instead of 1.05. Both values are approximations based on an analysis of the U.S. Social Security experience (Bell, Wade, and Goss 1992, p. 11).

Regarding reduction rates for the aged, U.S. projections expect these to continue at relatively high levels because medical advances against degenerative diseases are assumed to occur in the future.

Official projections in the United Kingdom use a variant of the method of extinct generations for ages 85 and over. According to this method, accumulated deaths are used to calculate mortality rates and reduction factors retrospectively once a generation is extinct. Such values can then be applied to estimate rates for nearly extinct generations (Gallop 1998, p. 5).

Mortality reductions for the oldest ages in the United Kingdom are treated on a cohort basis. Apart from some discussion attempting to explain the cohort effect for generations born before 1947, projections of mortality rates for the oldest ages are presented without numerical support. The 1994-based population projection report states that "making projections of death rates is *speculative* and users of projections of numbers of the elderly must bear in mind that the range of possibilities is wide" (ONS 1996, p. 10; emphasis added).

2.6 Uncertainty and Alternative Deterministic Scenarios

Uncertainty in official projections is usually covered through the use of alternative scenarios. Like the main projection, these follow a deterministic model and provide an idea about the sensitivity of the forecast in case the main assumptions are not realized. A high- and a low-mortality variant are usually considered in U.K. projections. Given the overall nature of recent U.K. projections, the variants simply reflect a departure from the central projection in an upward or downward direction. All scenarios converge to the same number of deaths in the long term, despite the differences in population size and structure (ONS 1996, p. 34, Fig. 9.3.b).

Scenarios in U.S. projections are a little more complicated because of the use of a detailed analysis by cause of death in the central projection.

The criterion used when defining an alternative scenario is not the level of mortality, but the magnitude of the financial cost borne by the Social Security System. This cost depends on the age structure of the population and the benefits provided under the scheme. Generally the low-cost alternative is associated with higher mortality rates. Two exceptions are considered, however: deaths under the category "congenital malformations and diseases of early infancy" and AIDS. This is because mortality related to infant diseases is concentrated in the range of ages under 5, and higher death rates for AIDS mean greater costs for the Social Security System (F. Bell 1997, p. 25).

U.K. population forecasts tend to consist of three separated projections: fertility, mortality, and migration. Each of these variables has its own alternative scenarios, in such a way that, in theory, 27 combinations would be possible. The sensitivity analysis for the U.K. projections may be described as marginal in character in that it shows the effects of changing only one variable at a time while the others are kept unchanged.

2.7 Accuracy

A comparison of a prediction with the actual outcome can, of course, be used to assess the performance of forecasts. The performance of official forecasts in the United Kingdom and the United States during the last decades demonstrates that reductions in mortality rates have been systematically understated.

As Murphy (1995, p. 337) says with respect to U.K. projections, "the values for expectation of life at birth for 2004 projected in 1964 were achieved around 1990 . . . whereas the values projected for 2016 in 1976 were actually achieved in 1981 for men and in 1983 for women." Persisting errors have not been so evident because the basic results from official projections are population size and structure subdivided by age and sex. A mutual offsetting of errors in the assumptions for fertility and migration has reduced the final effect (Shaw 1994).

With respect to the U.S. experience, Olshansky (1988, p. 496)[7] states that:

> in spite of rapid and near monotonic gains in life expectancy observed from age 65 from 1940 to 1986, the actuaries making each forecast simply could not believe that these gains would continue at that pace beyond the projection year. As a result, the gains in life expectancy that were forecasted to occur by the year 2000 were actually achieved within just a few years following the publication of the forecasts, because the trend towards declining mortality in advanced ages continued.

Official projections rely not only on an analysis by cause of death and allowances for cohort effects—elements often not considered explicitly by other models—but also on so-called "expert opinion"—for example, a figure reflecting expert views on the long-term reduction in mortality rates. Expert opinion, however, particularly in respect of mortality improvements, is not an exact science. This is exemplified by the view in U.K. projections that established trends cannot be extrapolated into the future and a mitigation of these trends is needed, leading to the imposition of a "diminishing returns" effect (ONS 1995, p. 10; 1996, p. 10; 1999, p. 29). Difficulties in making quantitative predictions about the prospect of mortality are illustrated in the results of a survey taken among expert attendants of a recent seminar organized by the Society of Actuaries (Rosenberg and Luckner 1998). Table 2 shows a summary of the responses from the 59 attendants at this seminar, in terms of the average annual percentage reduction in long-term mortality rates that these experts were expecting. A high variability in the responses and a significant percentage of nonresponses give a measure of the difficulties implicit in solving such a complex problem.

3. APPLICATIONS IN LIFE INSURANCE AND PENSIONS: PROJECTIONS PREPARED BY THE ACTUARIAL PROFESSION

The persistent decreasing trends in mortality constitute one of the main concerns of annuities and pensions providers. This concern becomes more acute when interest rates are low or reducing. Systematic underestimation of mortality rates for pricing and reserving, particularly in respect to guaranteed annuity and pension benefits,

[7] Quoted by Murphy (1995, p. 347).

Table 2

Summary of Results of Survey of Seminar Attendants (59 Surveys) Estimated Average Annual Percentage Reductions in Long-Term Mortality Rates

Country		Males			Females		
		High	Mean	Low	High	Mean	Low
Canada	Respondents		40			40	
	Average	1.05	0.58	0.27	1.10	0.64	0.29
	Std Dev.	0.32	0.28	0.38	0.34	0.26	0.39
Mexico	Respondents		39			39	
	Average	1.14	0.76	0.38	1.25	0.83	0.42
	Std Dev.	0.38	0.29	0.43	0.48	0.39	0.51
U.S.	Respondents		41			41	
	Average	1.06	0.67	0.26	1.08	0.70	0.25
	Std Dev.	0.43	0.32	0.48	0.44	0.36	0.48

Source: Rosenberg and Luckner (1998, p. 75).

might carry serious financial consequences in the long term (Willets 1999, pp. 3–4; Ballotta and Haberman 2003). Such concerns have been reflected, for many years, in the efforts of actuarial bodies in the United Kingdom and the United States to incorporate future mortality trends when constructing life tables.

3.1 Methods Developed in the United Kingdom

3.1.1 The Continuous Mortality Investigations Bureau: The "80" and "92" Series

The first life tables for annuitants to make allowance (admittedly limited) for the downward future trends in mortality were based on insurance company data for 1900–1920 (Evans 1998, p. 40). More recently the CMIB has periodically been considering future improvements in mortality for annuitants and pensioners. For example, the PA(90) and a(90) tables, constructed from the 1967–70 pensioners' and annuitants' experience, assumed uniform future mortality reductions, whose cumulative effect after 20 years was equivalent to one-year age reduction in the base table. The dependence of the implicit reduction factors on age, however, was almost imperceptible. Subsequent experience showed a significant correlation between mortality reductions and age, and so more recent sets of published tables (CMIB 1990, 1999) have included explicit formulae that allow for this feature in the derivation of the projected mortality rates.

The last two sets of CMIB tables (from the 1979–82 and 1991–94 experiences, called, respectively, "80" and "92" Series because of their base year) include explicit formulae to project mortality rates for annuitants and pensioners. The method adopted by the CMIB for estimating a projected value is to multiply the corresponding mortality rate $q_{x,0}$ from the base table by a *reduction factor,* $RF(x, t)$, where x denotes age and t time starting from the base year. equation (3.1) shows the way that reduction factor values are modeled:

$$RF(x, t) = \alpha(x) + [1 - \alpha(x)] \cdot [1 - f_n(x)]^{t/n},$$
$$q_{x,t} = q_{x,0} \cdot RF(x, t). \qquad (3.1)$$

As shown, $RF(x, t)$ follows an exponential decay characterized by two age-dependent parameters. The parameter α denotes the value to be asymptotically approached when t tends to infinite, while f_n is the percentage of the total fall $(1 - \alpha)$ assumed to occur in n years.

For the "80" Series, n was fixed at a value of 20 and f_{20} at 0.6 for all ages. $\alpha(x)$ was assumed to be linearly increasing in respect to x for ages between 60 and 110, with fixed values of 0.5 at ages under 60, and of 1 above 110. The function that describes $\alpha(x)$ is

$$\alpha(x) = \begin{Bmatrix} 0.5, & x < 60 \\ \dfrac{x-10}{100}, & 60 \le x \le 110 \\ 1, & x > 110 \end{Bmatrix}. \qquad (3.2)$$

The "92" Series shows important differences in the parameters assumed. The value for n remains fixed at 20, but f_{20} varies linearly from 0.45 to 0.71 between ages 60 and 110. Below age 60 and above 110, constant functions with the values already mentioned apply. In addition, $\alpha(x)$ takes the value of 0.13 instead of 0.5 for ages under 60. equations (3.3) and (3.4) show the functions corresponding to $\alpha(x)$ and $f(x)$, respectively:

$$\alpha(x) = \begin{Bmatrix} 0.13, & x < 60 \\ 1 + 0.87 \cdot \dfrac{(x - 110)}{50}, & 60 \le x \le 110 \\ 1, & x > 110 \end{Bmatrix}, \tag{3.3}$$

$$f_{20}(x) = \begin{Bmatrix} 0.55, & x < 60 \\ \dfrac{(110 - x) \cdot 0.55 + (x - 60) \cdot 0.29}{50}, & 60 \le x \le 110 \\ 0.29, & x > 110 \end{Bmatrix}. \tag{3.4}$$

Tables 3 and 4 show the limiting value $\alpha(x)$ when t tends to infinity, the percentage of the total reduction to be achieved in the first 20 years (f_{20}), and reduction factors calculated for selected ages x and time t ahead of the base year. Clearly, higher reductions are allowed under the new basis: that is, all the new $\alpha(x)$ values are lower than those from the "80" Series. Further, percentages of the total reduction to be achieved at the end of the first 20 years are much lower than the assumptions under the old basis. The limiting minimum values for the reduction factors increase with age, and the rate of change decreases with age. These assumptions implicitly reflect the operation of some "law of diminishing returns" over time and age, as has been suggested in U.K. official projections (ONS 1995, p. 10; 1996, p. 10; 1999, p. 29; and see the comment in Section 2.7).

The values assumed for the parameters have tended to result from the analysis of past trends of male pensioners and expert opinion. This is because the data for female pensioners show some irregularities, which make it difficult to identify clear patterns, and the available data for annuitants are relatively scarce in comparison to those for pensioners. This feature applies to the "80" Series as well as the "92" Series.

Average observed reductions over the periods 1967–70 to 1983–86 and 1971–74 to 1983–86 were calculated and studied for setting the corresponding parameter values for the "80" Series. An analysis of amounts and lives data by age group for pensioners shows a clear pattern for males: reduction factors increase with age and decrease with time. Female experience, however, does not show such a regular pattern. Reduction factors do not necessarily decrease at all ages or monotonically over time. In addition, at some ages female reduction factors are much lower than the corresponding male values.

Despite this, the CMIB (1990, pp. 52–53) states:

> The Committee is of the opinion that it would be excessively cautious to incorporate into the projected tables for female pensioners the rapid rates of improvement in mortality [observed]. . . . To do so would be to widen differences between male and female mortality in a manner which might be difficult to justify—particularly as the latest UK population projections assume a narrowing of the sex differential in mortality, on the basis of a consideration of trends in mortality from specific

Table 3
Calculated Values for Projected Tables: Base Table "80" Series

Age	Limiting Value	Percentage of Total Reduction Achieved in First 20 Years	Reduction Factors: Values of t (t = 0 for the Base Year)			
			10	20	40	60
≤60	50.0%	60.0%	81.62%	70.00%	58.00%	53.20%
70	60.0	60.0	85.30	76.00	66.40	62.56
80	70.0	60.0	88.97	82.00	74.80	71.92
90	80.0	60.0	92.65	88.00	83.20	81.28
100	90.0	60.0	96.32	94.00	91.60	90.64
≥110	100.0	60.0	100.00	100.00	100.00	100.00

Table 4

Calculated Values for Projected Tables: Base Table "92" Series

Age	Limiting Value	Percentage of Total Reduction Achieved in First 20 Years	Reduction Factors: Values of t (t = 0 for the Base Year)			
			10	20	40	60
≤60	13.0%	55.0%	71.36%	52.15%	30.62%	20.93%
70	30.4	49.8	79.71	65.34	47.94	39.20
80	47.8	44.6	86.65	76.72	63.82	56.68
90	65.2	39.4	92.29	86.29	77.98	72.94
100	82.6	34.2	96.71	94.05	90.13	87.56
≥110	100.0	29.0	100.00	100.00	100.00	100.00

causes. . . For practical purposes it is reasonable to adopt the same reduction factors for each projected table.

It is important to notice that for practical purposes there is no differentiated treatment between amounts and lives data, although mortality improvements for amounts data are usually higher than for lives. This reflects the difficulty with modeling (and hence projecting) the trends in amounts-based data.

Parameters for the "92" Series were calculated fitting the model to the experience over the period 1975–94.[8] As in the "80" Series, an analysis of the recent male pensioners experience identified a pattern in improvement rates that was decreasing with age and increasing with time. For females, irregularities persisted. However, as for the former basis, the CMIB recommended a single set for all groups. As a result the same reduction factors are applied for pensioners, widows, immediate annuitants and retirement annuitants; males or females; lives or amounts.

Table 5 shows estimates of observed reduction factors for the period 1980–90 from male pensioners' experience in comparison to the corresponding projected values applying the "80" Series methodology. The observed reduction factors are estimated (without fitting a model) from the ratio between crude mortality rates for the period 1991–94 to those of 1975–78, for lives and amounts. Since this interval is 16 years long, the values were standardized by means of a geometric adjustment (i.e., raising to the power of 10/16). The projected values were calculated using the formulae from the "80" Series with the midpoint of each age group

[8] No details are provided. See CMIB (1999), third page of the sixth chapter.

Table 5

Reduction Factors for the Period 1980–90, Male Pensioners

Observed (Approximate) vs. Projected ("80" Series)

Age Group	Ratios of Crude Mortality Rates 1991–94/1975–78 (16 Years): Lives (1)	Amounts (2)	Reduction Factors: Observed (10 Years Adjusted): Lives (3)	Amounts (4)	Reduction Factors: Projected "80" Series (5)
61–65	0.65	0.49	76.40%	64.03%	82.73%
66–70	0.65	0.56	76.40	69.60	84.56
71–75	0.69	0.63	79.30	74.92	86.40
76–80	0.75	0.69	83.54	79.30	88.24
81–85	0.83	0.74	89.01	82.85	90.08

Notes: (3) = (1) raised to the power of (10/16).
(4) = (2) raised to the power of (10/16).
(5) calculated with the central age of each age group.
Source: (1) and (2) from CMIB (1999), Table 6.1.

and time equal to 10. It is clear that, in all cases, the projected values underestimated mortality improvements, with the differences being much wider at the younger ages.

3.1.2 Other Approaches

Willets (1999) warns against a possible underestimation of improvement mortality in the CMIB methodology[9] even before the official publication of the "92" Series. Adopting similar assumptions to those of official projections in the United Kingdom (Section 2), he proposes a model allowing for both period and cohort effects. Improvement rates are assumed to move exponentially toward long-term values, with initial values calculated by extrapolation from the most recent experience. Finally, a low and a high improvement scenario are formulated. Willets's main conclusions can be summarized as follows:

- It would be preferable to measure improvement rates from population data, in order to take advantage of greater exposures to risk, avoiding unexplained fluctuations.[10]
- The CMIB methodology is even less cautious than his low improvement (optimistic) alternative.

Willets's analysis is mainly descriptive, with no statistical modeling. By ignoring CMIB data, he assumes that national experience after a simple adjustment can be applied to specific groups (in this case, annuitants and pensioners). The proposed projection basis is derived from the most recent past data, although the importance of a good run of data is acknowledged.

3.1.3 A Generalized Linear Model Approach

Since the preparation of the "80" Series, the CMIB has been using graduation techniques based on the following mathematical formula (CMIB 1990; Forfar, McCutcheon, and Wilkie 1988):

$$GM_{\alpha}^{r,s}(x) = \sum_{i=1}^{r} \alpha_i \cdot x^{i-1} + \exp\left\{\sum_{i=r+1}^{r+s} \alpha_i \cdot x^{i-j-1}\right\}. \tag{3.5.a}$$

With $r = 0$ and $s = 2$, this expression equates with the Gompertz law, while when $r = 1$ and $s = 2$, a Makeham equation is obtained. For this reason it has been called the Gompertz-Makeham (GM) formula.

In addition, the Logit Gompertz-Makeham (LGM) formula with parameters r, s is defined as

$$LGM_{\alpha}^{r,s} = \frac{GM_{\alpha}^{r,s}(x)}{1 + GM_{\alpha}^{r,s}(x)}. \tag{3.5.b}$$

Initial and central mortality rates (q_x and m_x) as well as the force of mortality μ_x can be graduated using either a GM or an LGM formula. In some cases it may be preferable to use an LGM formula for initial mortality rates, since values may be automatically restricted to the range 0–1. When possible the use of low values for r and s is preferred, ensuring an acceptable degree of smoothness.

An extension of this approach has been suggested for the modeling of secular mortality changes by including time as an explanatory variable. Renshaw, Haberman, and Hatzopolous (1996) have proposed the following equation for the force of mortality dependent on age and time:

$$\mu_{x,t} = \exp\left\{\beta_0 + \sum_{j=1}^{s} \beta_j \cdot L_j(x')\right\} \times \exp\left[\sum_{i=1}^{r} \left\{\alpha_i + \sum_{j=1}^{s} \gamma_{i,j} \cdot L_j(x')\right\} \cdot t'^i\right], \tag{3.6}$$

where L_j terms are Legendre polynomials and x', t' denote age and time after having been rescaled onto the range $[-1, +1]$. The first multiplicative term takes the form of a GM(0, s) formula. The second term can be seen as the product of r expressions very similar to a GM(0, s), with the difference that now each exponent is multiplied by a power of t'.

In an exploratory application of this model to male assured lives data from the United Kingdom over the period 1958–90, Renshaw, Haberman, and Hatzopolous (1996) found that the dependence on time is expressed by polynomials of low

[9] Willets refers to the proposal presented by the CMI Bureau at a seminar in London in December 1998. The "92" Series were officially published in June 1999. The only relevant difference is that in the initial version different parameters were assumed for females.

[10] This assertion could particularly apply to female experience, though the author does not specify this.

order ($r = 2$). The model is fitted to the data without further attempts to predict future trends. In addition, the authors warn against the indiscriminate use of the model for extrapolative purposes. Additional constraints beyond the observed period are suggested in order to avoid the danger of atypical turning points.

Sithole, Haberman, and Verrall (2000) have applied the same model to various data sets from the U.K. experience for immediate annuitants and life office pensioners over the period 1958–94, with a view to projecting mortality rates forward, and recognizing the constraints mentioned by Renshaw, Haberman, and Hatzopolous (1996). Two main conclusions can be taken from this work:

1. The optimum values of r and s when fitting the model do not necessarily generate plausible trends for extrapolating purposes. Lower-order polynomials have to be tested, sacrificing some degree of goodness of fit. The well-known trade-off between goodness of fit and smoothness experienced when constructing base life tables seems to apply to two-dimensional graduations for forecasting purposes.
2. For all data sets the most satisfactory results were obtained using the following formula:

$$\mu_{x,t} = \exp\left\{\beta_0 + \sum_{j=1}^{3} \beta_j \cdot L_j(x')\right\} \times \exp\{[\alpha_1 + \gamma_{11} \cdot x'] \cdot t'\}, \quad (3.7)$$

which is based on equation (3.6), with the difference that the value of the parameter s does not apply to both terms. The second term, involving the trend adjustment, is linear in age and time. As the authors remark, "including γ_{ij} terms involving higher terms in x resulted in mortality rates that did not progress smoothly at ages outside the range of ages graduated. On the other hand, introducing higher order terms in time t resulted in an unrealistically rapid improvement in mortality" (Sithole, Haberman, and Verrall 2000).

The structure of equation (3.7) is particularly interesting. The logarithm of the force of mortality is a function of two additive terms: the first depends exclusively on age, while the second is a product of a term depending on age and linear in time. As noted later in Section 5, this empirical conclusion coincides with the implementation of other approaches, including the Lee-Carter model.

Renshaw and Haberman (2000) show how the generalized linear model approach can be used to reformulate models involving mortality reduction factors (as described in Section 3.1.1.) so that a coherent overarching approach (covering Sections 3.1.1 and 3.1.3) can be seen to emerge.

Renshaw and Haberman (2003a, 2003b) advocate a variant of the above methodology by modeling the reduction factor, defined in terms of the force of mortality, viz.,

$$\mu_{xt} = \mu_{x0} RF(x, t), \quad (3.8)$$

using generalized linear models. The functional form successfully used for $\log RF(x, t)$ is

$$\log F(x, t) = \beta_x(t - t_0), \quad (3.9)$$

where the β_x term is modeled by a linear predictor. This is very similar in structure to equation (3.7). Versions involving break point terms that allow for a change of the slope term β_x in the historical data are found to be particularly successful in explaining the trends for males in England and Wales over the period 1950–98 (Renshaw and Haberman 2003b).

3.2 Methods Developed in the United States: The Society of Actuaries Group Annuity Valuation Table Task Force GAR94 Tables

The consideration of mortality improvements when constructing life tables and the inclusion of an explicit allowance for future improvements for annuitants in the United States can be found in the seminal work by Jenkins and Lew (1949). The Society of Actuaries recommended a projection basis for the first time with the publication of the GAR94 tables for statutory reserving purposes for annuitants (SOA 1995).

The model relies on a geometric progression of age-dependent improvement factors, following the equation

$$q_x^{1994+n} = q_x^{1994} \cdot (1 - AA_x)^n, \quad (3.10)$$

where, as is customary, q values denote initial mortality rates, x age, n time in years beyond the base year 1994, and AA_x is an improvement factor depending on age. This type of approach is similar to the reduction factor framework of Equation

(3.1) that has been used in the United Kingdom but with $\alpha(x) = 0$: see Benjamin and Pollard (1993, p. 408) and CMIB (1990, p. 50).

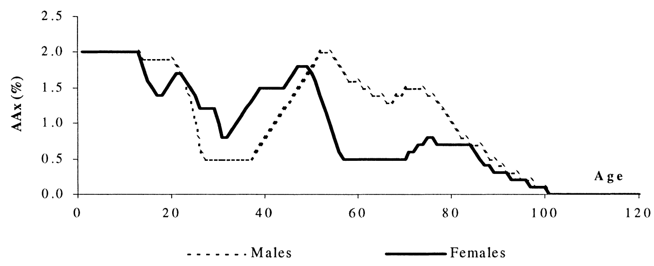
The life table is constructed from observed experience centered on 1988. For ages 66–95, group annuity data based on amounts are taken from the insurance industry. Younger and older ages receive different treatments because of the reduced exposures to risk. For ages 25–65, data based on lives are obtained from the Civil Service Retirement System. Mortality rates for ages 1–24 and 96–120 after some adjustments are extracted from the life tables prepared by the Social Security Administration (Bell, Wade, and Goss 1992, Section 2.2).

It must be noted that mortality rates at the most advanced ages are set at a maximum of 0.5, instead of 1.0, when reaching some theoretical limit to life span in the table. After having stated that a number of studies show mortality rates peaking at values below 0.5, it is noted that "Studies of mortality at the very old ages have shown that the mortality rate has a second bendpoint in the 80s or 90s, which reflects a deceleration in the rate of increase. The rate then proceeds to an approximate level ultimate rate after age 100" (SOA 1995, pp. 875–76).[11]

A life table for 1994 was derived from the previously mentioned 1988 values. Corresponding improvement factors were obtained after averaging age-specific log-linear trends observed from the CSRS experience over the period 1987–93. For this purpose life tables for each intermediate year were graduated using the Whittaker method. Improvement rates obtained were adjusted using the same approach. The resulting mortality rates for 1994 were graduated using the Karup-King four-point formula,[12] a smooth-junction interpolation technique, and later revised to allow for uncertainty margins.

The core information used to estimate improvement factors (AA_x) is statistical data obtained from the CSRS and the SSA life tables mentioned above. Average trends were calculated from the 1977–93 experience by fitting a linear regression through the logarithms of central rates

Figure 2
Mortality Improvement Factors by Age SOA Table GAR94



for five-year age groups. Interpolation techniques for calculating age-specific values and some adjustments allowing for expert opinion were also considered. Figure 2 shows the resulting AA_x values calculated for each age and for both sexes.

Remarkable irregularities are visible in both curves, and these are not explained in SOA (1995). A lack of smoothness combined with potential inconsistencies in the projected values is a drawback of the nonparametric techniques that have been used, and suggests that the methodology needs also to include smoothing, for example, by splines.

4. OTHER SCIENTIFIC METHODS

4.1 Introduction

The main criticisms of official forecasts have focused on the systematic overestimation of mortality and the lack of measures of sensitivity and uncertainty (Alho and Spencer 1985; Murphy 1995). Some of the alternative approaches proposed by academic statisticians and demographers are based on separated age-specific projections of mortality rates. Although they permit easy estimation of prediction intervals, potential inconsistencies in the projected age pattern may arise (Section 4.2). Other methodologies aim to achieve plausible age patterns by projecting the parameters from the mathematical formulae previously used for the static graduation of mortality rates for a set of calendar periods (Section 4.3).

The Lee-Carter model (Section 4.4) seems to be a solution to the trade-off mentioned above. Although this method has its critics, it represents U.S. mortality experience over this cen-

[11] This assertion is along the same lines as Thatcher (1999, Section 4.1.1).

[12] For details on these techniques, see London (1985).

tury very well, having been tested for use in official projections.

4.2 Independent Projections of Age-Specific Mortality Rates (Alho and Spencer's Approximate Linear Model)

Alho and Spencer (1985) have proposed a stochastic model for the purpose of making population projections and applied it to U.S. data. Each vital rate series (i.e., fertility, mortality, and migration) is considered to follow a so-called "approximately" linear model based on a stochastic process $\{y(t)|t \in Z\}$, where

$$y(t) = f(t) + \varepsilon(t),$$

$$f(t) = \sum_{j=1}^{k} \beta_j \cdot f_j(t) + \eta(t). \qquad (4.1)$$

Hence, $f(t)$ is a linear combination of known functions $f_j(t)$ with unknown parameters β_j. The expected value of $\varepsilon(t)$ is assumed to be zero, and $\eta(t)$ is *the bias in respect of the ideal mean*—the mean when the model reduces to the usual linear regression model, that is, when $\eta(t)$ is zero.

The value of the bias $\eta(t)$ increases in relation to the distance $|t - n|$, where n represents the initial year in the projection. This component has three effects in the model:

1. Flexibility is increased in the sense that it is not restricted to a linear expression
2. The specification of $\eta(t)$ involves assigning greater relative weight to the more recent observed values in the estimation of the coefficients β_j and
3. Prediction intervals whose width increases according to the distance $|t - n|$ are allowed for.

Thus, the narrower (wider) the range within which $\eta(t)$ can fluctuate—determined by the criteria of the researcher—the more (less) restricted the functional form, the shorter (longer) the relevant data period, and the narrower (wider) the resulting prediction intervals. Generally $\eta(t)$ is bounded at a minimum level considered suitable for generating "adequately" wide intervals. In addition, applying the technique of mixed estimation (proposed by Alho 1983), the authors are able to incorporate expert opinion about the future trends of $y(t)$, through a convex combination of suggested future values and results obtained from the model.

Finally, a linear growth model is applied to an estimated starting population using the predicted future vital rates. The purpose of this projection is to generate point forecasts of the population as well as to measure the propagation of errors resulting from the interaction of variability in the vital rates. This enables estimation of the corresponding prediction intervals.

The example of a population projection for the United States from 1980 to 1994 is given by Alho and Spencer, using data from 1957 to 1979. The results are compared with the projections of the U.S. Bureau of the Census, which uses a deterministic approach with several scenarios (and which has underestimated future experience). Although the model is applicable to any of the three vital rates, its use is exemplified only with the fertility rates series.

For developed countries, mortality is often regarded as the least uncertain of the three vital processes, although it is acknowledged that variability at advanced ages is significant (W. Bell 1997; Gallop 1998).[13] Hence, for each age Alho and Spencer take the official projections as the true means and estimate the variability around the means from past data. In addition, given the age-specific nature of mortality rates, the forecasting process would involve using the approximately linear model separately as many times as the number of age-specific mortality rates that the researcher wants to estimate (91 times, for single-year age groups from 0 to 89 and with a final 90+ group).

Despite the fact that the example referred to does not represent a rigorous application of the method, but rather a simplified illustration of the proposal, the following results from Alho and Spencer should be noted:

- In the only case where the model is used to estimate future vital rates (fertility), it is showed that even the resulting 67% (one standard deviation) prediction intervals are much

[13] Murphy (1995), measuring the performance of official forecasts in the United States and England and Wales for the period 1962–89, argues that, against the dominant opinion in this respect, mortality rates results are much poorer than those of fertility rates.

wider than the corresponding high-low bands obtained from the U.S. Bureau of the Census official variant projections.
- When forecasting population numbers, deterministic scenarios implicitly assume a perfect correlation among the errors from the different vital rates over time.[14] Hence, while prediction intervals during the first years are generally wider than high-low bands—because of the stochastic starting population assumed—when the propagation of errors is high enough, the high-low deterministic bands become wider.
- Because of the unacceptable width of the resulting prediction intervals, it would be risky to apply a stochastic population forecasting for time horizons longer than 15 years.

The main advantage of the Alho and Spencer model is to show how to apply a stochastic analysis to the process of forecasting vital rates including mortality, and how to allow for interactions over time while projecting populations. The proposed model allows for an analysis of the propagation of errors and facilitates the calculation of the corresponding prediction intervals. It also provides a formal way to incorporate expert opinion and results in the forecast being approximately independent of the functional form chosen over reasonably short projection periods. Furthermore, Alho and Spencer demonstrate that it is inappropriate to interpret the results of a deterministic methodology based on high-medium-low scenarios as if they were equivalent to a stochastic approach.

On the other hand, the approximately linear model proposed for the estimation of future rates is a one-dimensional model. In order to project mortality rates $q_{x,t}$ or forces of mortality $\mu_{x,t}$, we would need to apply the model separately for each age. This implies a heavy computational workload, and the model lacks any theoretical way of dealing with potential correlations between mortality rates, which correspond to different ages for a fixed year (as might be anticipated, given the theory of graduation of cross-sectional mortality rates). With regard to actuarial applications for mortality projections, this might generate point-estimated curves that lack an acceptable degree of smoothness or that follow unacceptable trajectories with respect to age.

4.3 Independent Projections of Life Table Parameters

As mentioned at the end of Section 4.2, one of the main disadvantages of the independent age-specific projections of mortality rates is the risk of obtaining irregular shapes for the forecasted life table at some point not far beyond the base year. This problem arises from the fact that models such as that described in the previous section do not include any relationship between the different age-specific mortality rates for a fixed period in time.

In order to preserve a logical age pattern over the projection period, some authors have suggested forecasting the parameters of some mathematical curve that fits observed static experiences for the whole range of ages well. Because of its capability of representing different patterns of mortality specific to each stage of the human life span, the Heligman-Pollard curve has been used for these purposes (McNown and Rogers 1989; Forfar and Smith 1988; Congdon 1993).

4.3.1 Static Mortality Models for the Whole Range of Ages: The Heligman-Pollard Curve and Alternatives

Heligman and Pollard (1980) have proposed a mathematical curve to represent mortality over the whole range of ages, showing that it fits very well to postwar Australian experience. The proposed model is based on the equation

$$\frac{q_x}{p_x} = A^{(x+B)^C} + D \cdot e^{-E(\ln[x]-\ln[F])^2} + GH^x, \quad (4.2)$$

where q_x and p_x are probabilities of mortality and survival, respectively, and the letters A to H denote the eight parameters to be estimated from each experience.

As seen, the model is made up of three terms, each of which represents mortality behavior for a specific stage of the life span: infancy, young adulthood, and maturity. The meaning of each parameter is explained as follows:

A is approximately the initial mortality rate at age 1, q_1.

[14] Tuljapurkar and Boe (1998, p. 40) point out the same problem, defining it as a "difficulty of consistency" with respect to the criteria that the researcher has to use to combine the high, low, and medium values from separate forecasts to generate an aggregate estimate.

B indicates the location of infant mortality, q_0, in the interval $[q_1, 0.5]$. When $B = 0$, $q_0 = 0.5$, irrespective of the values of A and C.

C measures the speed of infant mortality decline.

D, E, and F represent severity, spread, and location of the so-called accident hump.

G and H, components of the term representing mortality at old ages (a Gompertz curve), denote the initial level of senescent mortality and its rate of increase.

Despite its apparently complicated formulation and the number of parameters involved, the H-P curve has the important advantage in that it represents specific mortality patterns for different stages of life span with a single equation. Examples of empirical investigations include McNown and Rogers (1989) and Congdon (1993), who have investigated the suitability of the H-P model for representing the mortality experience of the United States and of greater London, respectively.

Two more recent developments might be seen as alternatives to the H-P curve. Thatcher (1999) has pointed out that the Gompertz law would not be a good representation of mortality at very old ages. A logistic model is proposed based on the hypothesis that mortality rates reach a plateau at the highest ages.[15] This formulation is tested with several data sets over history, from Hungarian medieval to English recent experience, providing a reasonable fit to the data. It should be noted that, although this model seems to represent senescent mortality as a special feature not included in previous approaches, it considers neither infant mortality nor the accident hump. Thatcher has used a Perks model (Perks 1932), which also arises from the frailty model of Vaupel, Manton, and Stallard (1979) with a Gompertz/gamma combination (see Butt and Haberman [2004] for further discussion and applications to life insurance and annuitant data sets). Thus, Thatcher suggests[16]

$$\mu_x = \gamma + \frac{\alpha \cdot e^{\beta \cdot x}}{1 + \alpha \cdot e^{\beta \cdot x}}. \qquad (4.2a)$$

Carrière (1992) has pointed out that the shape of the curve represented by the H-P equation is not universally applicable. U.S. experience shows that female patterns in particular seem to be rather unstable over time. In general, Carrière postulates that there is no single typical shape, but that most age patterns can be represented as combinations of different stage-specific "laws of mortality." Hence the following equation is proposed:

$$s(x) = \sum_{k=1}^{m} \psi_k \cdot s_k(x), \qquad (4.3)$$

where $s(x)$ denotes the survival function,[17] ψ_k the probability of death by cause (during stage) k so that $\sum_{k=1}^{m} \psi_k = 1$, and $s_k(x)$ the probability of survival at age x given that a life will die from cause k. It is assumed that the causes of death can be classified as childhood, teenage, and adult causes so that $m = 3$. Therefore, each function $s_k(x)$ is replaced by a typical model for each stage. For example, given its decreasing force of mortality, a Weibull model is proposed for the childhood ages, whereas to represent the teenage accident hump, an Inverse-Weibull or Inverse-Gompertz model is postulated according to the specific experience. The Gompertz model, as usual, represents adult mortality.

Carrière has applied the model to several recent U.S. population experiences, giving better measures of goodness of fit than those obtained when the H-P formula is used. Carrière suggests only Gompertz, Weibull, Inverse-Gompertz, and Inverse-Weibull functions for $s_k(x)$, but equation (4.3) is flexible enough to incorporate other possible laws of mortality. Carrière has reparameterized these in terms of their mode and standard deviation, in such a way that the fitting process involves estimation of $(3 \cdot m - 1)$ informative parameters where m is the number of stages considered by the researcher. For each different value of k, apart from the parameters representing mode and standard deviation, the probability ψ_k must also be estimated: hence the factor 3. One parameter is deducted since the sum of the probabilities is always equal to unity.

It must be remarked that, even though Carrière does not mention senescence among the categorized causes of mortality, the model is flexible

[15] This view is shared by SOA (1995, pp. 875–76).

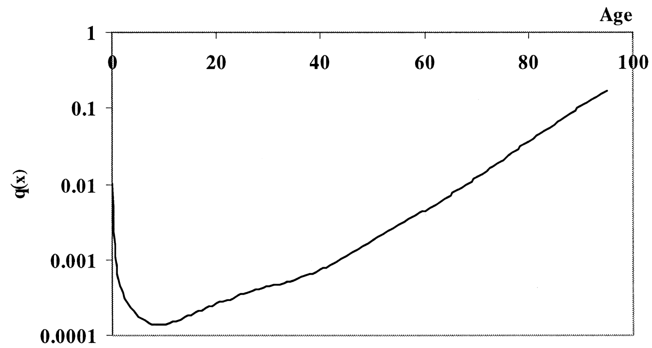
[16] This is effectively very similar to the final term of a variant of Equation (4.2), proposed by Heligman and Pollard (1980).

[17] By definition, $s(x) = \exp(-M(x))$, where $M(x) = \int_0^x \mu_{(s)} \cdot ds$.

enough to incorporate any feature specific to any stage of life span. There is no restriction on either the maximum value for m, or for the number of times a particular type of distribution can be reused in the same model.

To illustrate, Table 6 and Figure 3 show the results of fitting equation (4.3) to U.S. female population data from 1980. In this case four causes are postulated (i.e., 11 parameters are estimated). The first two components on the right-hand side of equation (4.3) are Weibull distributions, while the other two are Gompertz. Table 6 shows that 97% of overall mortality is explained by the Gompertz model, centered at age 90 and approximately affecting the range from 72 to 108 (i.e., mean $\pm$ 2 standard deviations).

Figure 3
U.S. 1980 Female Experience Using Carrière (1992) Model (Log Scale)



Source: Recalculated from data taken from Carrière (1992).

4.3.2 Series of Calendar Year–Specific Graduations and Projection of Parameters

Forfar and Smith (1988) have applied the H-P curve to represent the published graduated mortality rates from the first 13 English life tables from the 1841 to the 1970–72 population experience for males and females. Applying a common formula for the whole range of ages over a long period has the advantage that trends in the parameters can be identified. After finding that the model fits very well the graduated data from all the periods considered, forecast values are suggested for the 1981-based life tables (ELT 14). In a postscript a comparison with the ELT 14 tables, then recently published, is provided, showing the closeness of the projection.

In an exploratory exercise McNown and Rogers (1989) have fitted the H-P curve to U.S. population data from 1900 to 1985. However, only the period after 1941 is considered as experience for projection purposes because of the high variability in the observed parameters, which results from considering the complete sample. This is implemented despite a recognition that the reduced observational sample might lead to large standard errors and a lack of statistical significance for the estimated parameters. Individual univariate ARIMA models have been used to fit the trend over time of each parameter and to project forward up to 2000. Unlike Alho and Spencer (1985) (Section 4.2), an implicit validation of the results is made by pointing out the closeness with SSA forecasts.

Table 6
Parameters of the Carrière Model Applied to U.S. 1980 Female Experience

Distribution	Weibull (K)		Gompertz (K)	
	1	2	3	4
Prob(k)	0.01473	0.006268	0.008959	0.970043
Mode	0.3388	33.3	55.76	90.46
Standard deviation	1.904	10.52	6.67	9.128

Source: Carrière (1992, p. 95).

In a similar vein, Sithole (2003) uses the Gompertz-Makeham model of equation (3.5a) but adapts it so that the parameters are allowed to be time dependent. Using CMI Bureau data for female annuitants over the period 1958–94 (allowing for missing observations for 1968, 1971, and 1975), Sithole fits a model cross-sectionally of the form

$$\mu_{xt} = \exp[\beta_{0t} + \beta_{1t}x' + \beta_{2t}\tfrac{1}{2}(3x'^2 - 1)], \quad (4.4)$$

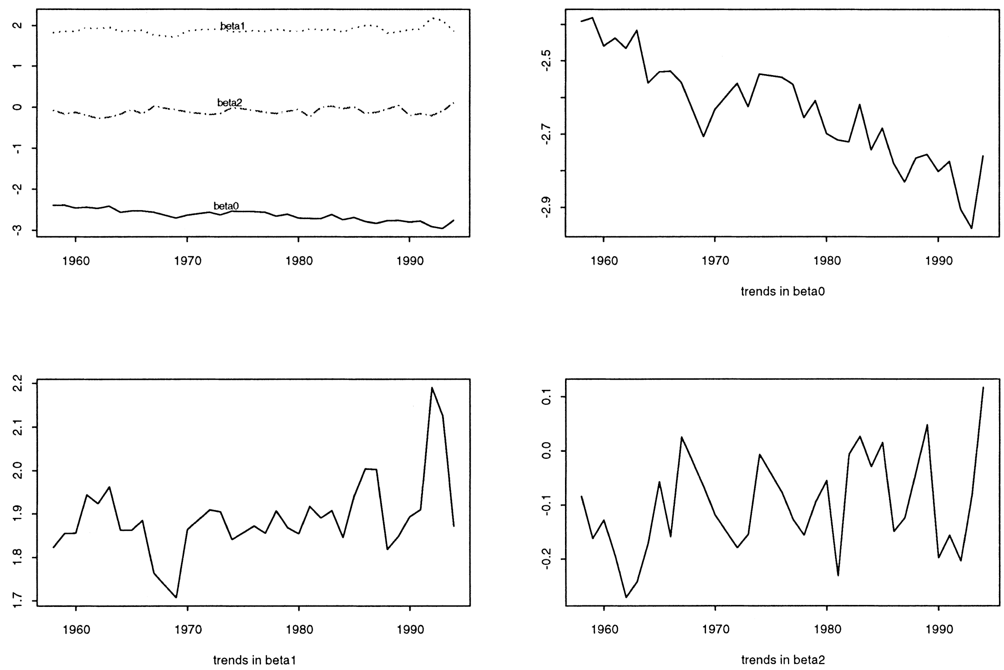
where x' is a rescaled age parameter as in equation (3.6).

Figure 4 shows the plots of the parameters over time, and we observe a steady decline in the level of mortality as represented by the decrease in the estimates over time of β_{0t}. The estimated values of β_{1t} and β_{2t} fluctuate around central values in the ranges 1.7–2.0 and -0.3 to 0, respectively. Using standard univariate time series methods,

Figure 4

Female Immediate Annuities (U.K., CMI Bureau)

Plots of Estimated (1958–1994) Parameter Values for beta0, beta1, and beta2 over Time *t*



Sithole (2003) then fits ARIMA (0, 1, 0), (2, 0, 0), and (1, 0, 0) to the three-parameter series and uses these to generate forecasts.

It is clear that these time-series-based methods are computationally demanding in that they require (a) an age-dependent model that provides a good fit to the historical data for each cross-sectional period, (b) the derivation of forecast models for each of the parameters in the age-dependent model, and (c) using the forecast time series models then to derive projected mortality rates (Sithole 2003). Further, two problems arise when projecting life table parameters independently: the complexity of the procedures involved in estimating prediction intervals, and the exclusion of any potential interrelations among the parameters for each static life table.

4.3.3 Independent Projection of Parameters Using the Relational Approach

In cases where there are not enough detailed data to construct life tables, a relational approach can be useful. It consists of a set of life tables each expressed as a mathematical function (usually with a few parameters) of a common standard table available in a tabular form. Relational approaches have been applied mainly for the graduation, adjustment, and extension of the limited and defective data from developing countries (Brass 1971).

This approach can be summarized by means of the next equation (Congdon 1993):

$$\log it(l_t(x)) = \alpha_t + \beta_t \cdot \log it(l_s(x)), \qquad (4.5)$$

where $l_t(x)$ denotes the proportion of people alive at age x in the life table for period t, while $l_s(x)$ has similar interpretation for the standard table; α_t and β_t are parameters estimated for each table t. Empirical evidence supporting this relationship is available from the analysis of United Nations mortality schedules (Brass 1971).

As an example, Congdon (1993) has applied the relational approach for projection purposes by using male life tables for greater London over 1971–90. Equation (4.5) is fitted for each life table, where t represents calendar year and the standard corresponds to the base year. The series of values for the parameters α_t and β_t are then modeled by time series ARIMA methods (as in the applications described in Section 4.3.2 and for k_t in the Lee-Carter method discussed in Section 4.4).

4.4 The Lee-Carter Model

Independent projections of age-specific mortality rates (Section 4.2) may face a lack of consistency since this approach does not ensure the plausibility of the shape of the forecasted age patterns. However, it is relatively straightforward to estimate prediction intervals given the use of such individual univariate models.

On the other hand, the projection of the parameters of life table models (Section 4.3) leaves less scope for generating illogical age patterns in the future,[18] but has a consequent cost in terms of complexity when obtaining measures of uncertainty. Analysis of these approaches suggests a trade-off between the plausibility of the projected age patterns of mortality and the simplicity of construction of the prediction intervals.

With a relatively simple formulation, Lee and Carter (1992) postulate a one-parameter family of life tables, attempting to ensure plausibility of the projected age pattern and to measure uncertainty without complexity. The model suggested is based on the equation

$$\ln(m_{x,t}) = a_x + b_x \cdot k_t + \varepsilon_{x,t}, \qquad (4.6)$$

where $m_{x,t}$ is the central mortality rate for age x for year t; a_x and b_x are parameters dependent only on age; k_t is a factor to be modeled as a time series; and $\varepsilon_{x,t}$, the error term, is assumed to have mean zero and standard deviation σ_ε.

Equation (4.6) shows that there are as many values of a_x and b_x that need to be estimated as there are age groups. Along with k_0, these values make up a base life table, which is deformed for future t over the range of values taken by k_t.

The proposed model cannot be fitted by simple regression methods, because there is no observed variable on the right-hand side of the equation (Lee 1997). Further, it allows for several solutions. To deal with this, it is proposed that b_x and k_t are normalized to sum to unity and to zero, respectively. With this convention, each a_x takes the value of the averages of the sequence of $\ln(m_{x,t})$ over time for each x. Subject to these restrictions, a unique least-squares solution for b_x and k_t can be obtained (using the method of singular value decomposition).[19]

Taking the estimated values for a_x and b_x as fixed ($\hat{a}_x$, $\hat{b}_x$), Lee and Carter then propose a second stage of estimation of k_t, whereby the k_t's are recalculated from the equation

$$D(t) = \sum [N(x, t) \cdot \exp(\hat{a}_x + k_t \cdot \hat{b}_x)], \qquad (4.7)$$

where $D(t)$ is the total number of observed deaths and $N(x, t)$ the age structure of the corresponding population for year t. This second step has two justifications: first, the values obtained from the first stage minimize errors in the logs of the death rates instead of the death rates themselves, and hence, differences can occur between actual and estimated numbers of deaths. So the life tables derived for the years in the sample will be consistent with the total number of deaths, $D(t)$, and the population age structure, $N(x, t)$. Second, equation (4.7) allows the model to be fitted and values of k_t to be estimated even if only the total number of deaths and the population age structure are known but age-specific mortality rates are not available (because of data inaccuracies or time lags in publishing the rates, for example).

Lee and Carter obtain a linear decreasing pattern for the estimated k_t from equation (4.7) over the period 1900–1989 for U.S. population experience. Hence, using the ARIMA methodology, a random walk with drift is found to provide a good fit, that is,

$$k_t = c + k_{t-1} + u_t. \qquad (4.8)$$

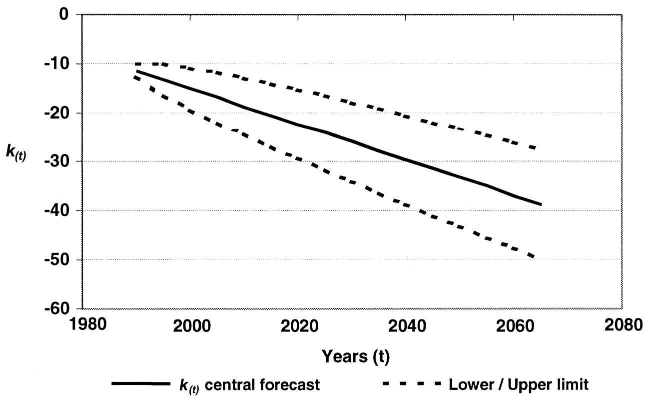
Figures 5 and 6 represent the projected values of k_t and central death rates. The data considered both sexes combined. Thus, when applied to U.S. mortality, a log-linear relationship between death rates and time is assumed, with the intercept and slope depending solely on age.

Clearly the Lee-Carter approach belongs to the class of extrapolative models, which represent past trends and assume continuity of inertia into the future, without allowing explicitly for expert opinion.

[18] There is nothing that guarantees "reasonable" values for projected parameters, since each has an interpretation and therefore is implicitly restricted to some interval (Lee and Carter 1992, Rejoinder).

[19] Full details can be found in Lee and Carter (1992).

Figure 5
$k_{(t)}$ Projected Values under the Lee-Carter Model Prediction



Source: Lee and Carter (1992), Table 2.

Relating to measures of uncertainty, Lee and Carter show that, for not too short projection periods, the error resulting from the estimation of k_t dominates the *other* sources of variability: the error associated with the estimation of the parameters a_x and b_x, and the error of the forecasting equation $\varepsilon_{x,t}$. Provided this is true for other data sets, the problem of estimating the variance of forecasted values is reduced to estimating the variance of the k_t estimates and then multiplying by $\hat{b}_x^2$. Lee (2000) and Lee and Miller (2001) provide further consideration of the model, its implementation, and its application.

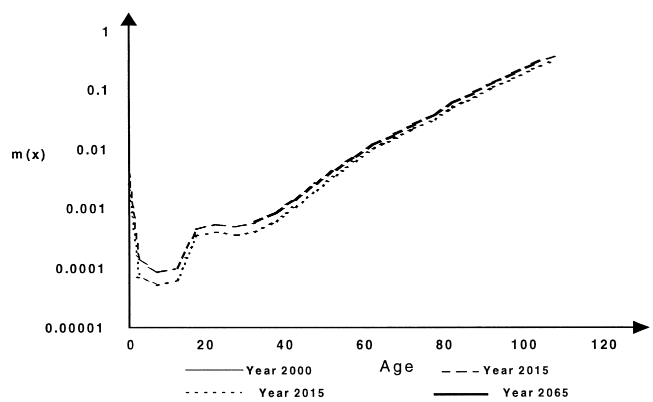
Lee (2000) has suggested that one of the problems with the method is the fact that the rates of decrease at different ages ($b_x \cdot (dk_t/dt)$) always maintain the same ratios to one another over time. It implies that the method does not take into account changes in age patterns over time, contradicting experiences like that of Sweden, where mortality rates at old ages are declining more quickly than previously.

W. Bell (1997) compares the methods of Section 4.3.2 with the Lee-Carter method by evaluating their out-of-sample performance in forecasting age-specific mortality rates for the United States over an 11-year time horizon. The forecasts produced by these methods and variants are found to be less accurate than from a simple random walk with drift model applied to the age-specific mortality rates without any age-dependent modeling.

There have been a number of recent developments to the Lee-Carter methodology that have been proposed and implemented. These include formulating the methodology in a manner comparable to the mortality reduction approach described in Section 3.1.1 (Renshaw and Haberman 2003a); interpreting the model structure underpinning the Lee-Carter approach to make it more comparable to a generalized linear model–based methodology (Renshaw and Haberman 2003b); introducing Poisson error structures (Brouhns, Denuit, and Vermont 2002; Renshaw and Haberman 2003c); and constructing the forecasts on the basis of the first two sets of single-value decomposition vectors (Bell 1997; Booth, Denuit, and Vermont 2002; Renshaw and Haberman 2003c). This last point begins to address Lee's criticism above by incorporating a second age-time interaction term.

As an illustration of the results that emerge from different approaches, we show in Table 7 some forecast mortality reduction factors for England and Wales for males for the year 2020, based on empirical data for the period 1950–98 using the models of Renshaw and Haberman (2003b). The comparison is of two approaches: one based on Lee-Carter and the other using the Generalized Linear Modeling approach of the type in Equation (3.9) but with a change of slope in 1975, marking an acceleration in the downward trend of mortality rates after this point. These results show some notable differences in level and profile against age (e.g., for age group

Figure 6
Central Death Rates Projected with the Lee-Carter Model (Log Scale)



Source: Lee and Carter (1992), Tables 1 and 2.

Table 7
Projected Reduction Factors for 2020: England and Wales, Males

Age Group	Lee-Carter Method	GLM Method with Hinge at 1975
0–1	0.44198	0.34048
1–4	0.48510	0.44384
5–9	0.48695	0.48838
10–14	0.61410	0.54040
15–19	0.78005	0.65959
20–24	0.81164	0.91066
25–29	0.82886	1.04488
30–34	0.81505	0.99917
35–39	0.76590	0.89455
40–44	0.74045	0.65160
45–49	0.70469	0.54706
50–54	0.69253	0.53693
55–59	0.69949	0.54370
60–64	0.72274	0.58991
65–69	0.75646	0.63794
70–74	0.78836	0.62988
75–79	0.80220	0.68044
80–84	0.81636	0.73441
85 and over	0.84679	0.76108

25–29, where the Renshaw-Haberman method leads to a reduction factor greater than 1 reflecting recent upward trends at these ages, which have been attributed to HIV and AIDS: Daykin 2000).

We note that the Lee-Carter method is based on a multiplicative decomposition of time effects that allows for their collective treatment and modeling across the whole age group. The Renshaw-Haberman method, based on GLMs, is designed to identify and project, through the incorporation in this instance of a change in slope in 1975, the most recent age-specific trends in mortality. This difference of approach leads to the different results obtained.

5. DISCUSSION

As we have seen from the earlier sections, important similarities exist between the different models, despite some apparent differences in structure or methodology. We review and compare the different models under the following important headings.

5.1 Log-Linearity in Time of Mortality Rates or Forces of Mortality

A log-linear relationship between mortality rates and time is present in most actuarial applications and in the Lee-Carter model.

Projected mortality rates under the CMIB model (equation 3.1) can be expressed as

$$q(x, t) = q_x \cdot \alpha(x) + q_x \cdot [1 - \alpha(x)] \cdot e^{-\beta_x \cdot t}. \quad (5.1)$$

Under the Renshaw-Sithole alternative, the projected force of mortality is expressed as equation (3.7):

$$\mu_{x,t} = \exp\left\{\beta_0 + \sum_{j=1}^{3} \beta_j \cdot L_j(x')\right\} \times \exp\{[\alpha_1 + \gamma_{11} \cdot x'] \cdot t'\}. \quad (5.2)$$

Equation (3.8), on which the SOA methodology is based, can be rewritten as

$$q_x^{1994+n} = q_x^{1994} \cdot e^{n \cdot \log(1-AA_x)}. \quad (5.3)$$

Equation (4.6) from the Lee-Carter model is expressed as

$$\ln(m_{x,t}) = a_x + b_x \cdot k_t + \varepsilon_{x,t}, \quad (5.4)$$

with k_t modeled by a random walk. With the exception of the CMIB approach, each model assumes a log-linear relation between mortality rates or forces of mortality in respect to time.

Because of the inclusion of the first additive term, $\alpha(x)$, expressing a limiting value for the force of mortality in the very long term, this common feature is not observed in equation (5.1). If this term were omitted, log-linearity would also apply to this approach.

5.2 Decreasing Improvements by Age

A decreasing pattern of mortality improvements over the range of ages is present in official projections, actuarial applications, and the Lee-Carter model, reflecting a generalized view about a supposed law of diminishing returns. Observed experience does not necessarily support this assumption, especially in the case of females, where severe irregularities have been observed in a number of cases.

From equations (5.1), (5.2), and (5.3) and the estimated parameter values, it can be seen that the corresponding exponents take always negative values, ensuring projected decreasing trends. For the CMIB methodology (equation 5.1), β_x is always positive, since $f_n(x)$ can take only values between 0 and 1 (equation 3.4). Hence the time factor in equation (5.1) corresponds to exponential decay with a parameter β_x that depends on age.

From the analysis of the empirical results obtained by Sithole, Haberman, and Verrall (2000) and equation (5.2), α_1 is always negative and much greater in absolute value than γ_{11}, which is generally positive. Since age and time are rescaled onto the range $[-1, +1]$, any projection would imply values for t' greater than $+1$. Consequently the exponent of the term trend is always negative and constant for each age. As above, we have exponential decay at ages such that $\alpha_1 + \gamma_{11} \cdot x' < 0$.

In equation (5.3), if we assume that improvement rates can only be positive, $\log(1 - AA_x)$ is always a negative value, since AA_x is contained in the range [0, 1]. Thus, again, we have exponential decay with parameter $\log(1 - AA_x)$, dependent on age.

In addition to this, except in the SOA methodology—where severe irregularities occur over the range of ages resulting from the use of nonparametric techniques—all the models assume a decreasing improvement in relation to age.

5.3 Increasing Trends in the Relative Rate of Increase in Mortality According to Age

Analyzing relative mortality progression with age corresponding to diverse English data sets over history, Thatcher (1999, p. 21) suggests that:

> This is a remarkable stability over such a very long period, which calls for explanation. It would be obviously desirable for this finding to be checked, if further historical data can be analysed or collected, but at first sight these results appear to support the views of those who believe that there is an underlying pattern of aging, genetically determined.

The results obtained by Thatcher are summarized in Table 8. The relative rate of increase of mortality is defined as the differential $(d\mu_x/dx)/\mu_x$, where μ_x is the force of mortality at age x. This ratio is assumed to be approximately constant in Thatcher's model where, from equation (4.2.a), μ_x is approximated by $\alpha \cdot \exp(\beta \cdot x)$.

Forfar and Smith (1988) note a similar pattern from their analysis of a sequence of graduated English life tables. Table 9 shows that the estimated values for the relative rate of increase in the initial mortality rates q_x vary with age. If, in the Heligman-Pollard curve (equation 4.2), only the term for adult mortality is taken into account (i.e., the Gompertz term, GH^x), and the probability of survival, p_x, is assumed to be close to 1 for adult ages, hence $(dq_x/dx)/q_x = \ln(H)$.

Table 8

Rate at Which the Force of Mortality Increases (Thatcher's Logistic Results)

Observed	Relative Increase Rate	Standard Error
England 1640–1689	0.086	0.0043
England 1687–1691	0.089	0.0049
England and Wales 1841, males	0.101	0.0008
England and Wales 1841, females	0.105	0.0008
England and Wales 1980–1982, males	0.099	0.0001
England and Wales 1980–1982, females	0.109	0.0001

Source: Thatcher (1999).

Close inspection of Tables 8 and 9 shows that the relative change of mortality rates according to age seems to increase slightly over time. Table 10 shows estimated average values for $(dq_x/dx)/q_x$ obtained by Heligman and Pollard (1980) from Australian postwar experience. Since Table 10 is based on the Heligman-Pollard curve, the relative rate of mortality increase according to age is approximately measured by $\ln(H)$.

We have constructed Tables 11 and 12 by fitting a Gompertz model to historical and projected U.S. data. The Gompertz parameters G and H are estimated by fitting a linear regression based on $\log(\mu_x)$ using the minimum least-squares criterion. Forces of mortality have been derived from published mortality rates under the assumption of a uniform distribution of deaths over the life year.

From the analysis of these tables, it is possible to identify for males an increasing pattern over time in the relative rate of mortality change according to age, and remarkable similarities in the absolute values from diverse experiences. This result is consistent with the evidence that young ages have benefited most from mortality reductions over this century (Charlton 1997, p. 17; Goss, Wade, and Bell 1998, pp. 112–13). Table 10 shows that the trend is not so clear for females. Some commentators have suggested a changing role for women in modern society as an explanation for this different feature (F. Bell 1997, p. 25; Berin, Stolnitz, and Tenebein 1989, pp. 23–24).

Thus, an increasing pattern over time of the relative rate of increase in mortality according to age seems to be a recurrent characteristic, at least for males. Such a consistent trend could support the approach of McNown and Rogers (1989) discussed in Section 4.3.2, based on the modeling of key parameters.

5.4 Smoothness and Adherence to Data

A clear feature of the set of mortality improvement rates estimated by the SOA (Fig. 2) is preference for goodness of fit over smoothness. A problem of consistency might then arise when projecting mortality rates and constructing forecast life tables. Despite using a smooth base table with increasing values of q_x over the range of adult ages, it is possible to generate crossovers (i.e., age ranges where mortality rates decrease by age). This anomaly could be the result of the accumulated effect of improvement rates increasing over some age ranges.[20]

Table 9

Values for the Parameter *H* of the Heligman-Pollard and Estimated Relative Rates at Which *q*(*x*) Varies with Age Obtained by Forfar and Smith with English Results

Observed	Estimated Relative Increase (ln(*H*))	Values for *H*
ELT1	0.089	1.093
ELT2	0.084	1.087
ELT3	0.081	1.085
ELT4	0.084	1.088
ELT5	0.085	1.088
ELT6	0.082	1.085
ELT7	0.080	1.083
ELT8	0.081	1.085
ELT9	0.090	1.094
ELT10	0.088	1.093
ELT11	0.100	1.105
ELT12	0.103	1.109
ELT13	0.103	1.109

Source: Forfar and Smith (1988), Table 1.

Table 10

Values for the Parameter *H* of the Heligman-Pollard and Estimated Relative Rates at Which *q*(*x*) Varies Obtained by Heligman and Pollard (1980) with Australian Results

	Males		Females	
Observed	Estimated Relative Increase Rate	Value for *H*	Estimated Relative Increase Rate	Value for *H*
1946–48	0.092	1.0970	0.107	1.1136
1960–62	0.094	1.0992	0.097	1.1022
1970–72	0.096	1.1013	0.095	1.1007

Source: Heligman and Pollard (1980), Table 1.

The same reduction factors used in the CMIB "80" and "92" Series are applied to males and females for both annuities and pensions. This contradicts the more recent findings of Sithole, Haberman, and Verrall (2000), who report noticeable differences between these groups differentiated by type, and the analysis in CMIB Reports Nos. 10 and 17, which is unable to identify a clear time pattern for females.

The Renshaw-Sithole approach is a clear example of a trade-off between smoothness (plausibility of the projected trend) and goodness of fit. In order to avoid any crossover in the projected rates, the degree of the polynomials involved in the best-fitted model has to be lowered.

There are suggestions regarding the use of mathematical formulae instead of tabular methods in U.S. actuarial practice. Thus, in the discussion accompanying the publication of the SOA GAR94 tables, Carrière (1995) states that "In my opinion, parametric formulas for mortality tables are always preferable to tabular rates." In one of the main textbooks used by actuarial students in the United States, London (1985, Chap. 1) gives an excellent account of the advantages of and justification for smoothing.

However, there is no guarantee that the projected parameter values will necessarily take

[20] Several values from the GAR94 table were tested. It would seem that these anomalies only arise in very long projection periods. The improvement projection model proposed by the SOA is considered to be valid for no longer than 15 years (SOA 1995, p. 869).

Table 11

Values for the Base of the Exponential Term Assuming a Gompertz Progression, and Estimated Relative Rates at Which the Force of Mortality Varies with Age, U.S. Data, Males Aged 30–65

Observed Period	Estimated Relative Increase Rate (ln(*H*))	Estimated Exponential Base (*H*)	*G*	Sum of Squared Errors
1900	0.0557	1.0573	0.00108344	0.00008623
1930	0.0677	1.0700	0.00048605	0.00001271
1960	0.0858	1.0896	0.00013793	0.00000695
1980	0.0885	1.0926	0.00009233	0.00000163

Note: Force of mortality assumed to be constant over each one-year life.
Source: Calculated from initial mortality rates taken from Bell, Wade, and Goss (1992) using the least-squares criterion.
Force of mortality assumed to follow the function $G \cdot \exp(x \cdot \ln(H))$ for the range of ages 30–65.

"reasonable" values. For example, the values of A in the Heligman-Pollard curve, or of any ψ_k under the Carrière model, should be restricted to the interval [0, 1], since they represent probabilities. Straightforward application of time series models could lead to projections that stray beyond this range, unless a method of incorporating such additional constraints is considered.

For example, McNown (1992) notes that the Lee-Carter model gives implausible profiles for 2030 and 2065 for teenagers and young adults. Similarly Lee (1997, p. 6) acknowledges having obtained unreasonable trends in the mortality differentials between the two sexes when applying this model to Canadian data.

5.5 Difficulties in Modeling Mortality Improvements at the Advanced Ages

Inaccuracies in the data available and variability due to small exposures to risk are often mentioned as the main difficulties in modeling most mortality improvements at the advanced ages. As a result, the methodologies reviewed usually restrict the range of ages under study. The only approach giving special treatment to advanced

Table 12

Values for the Base of the Exponential Term Assuming a Gompertz Progression and Estimated Relative Rates at Which the Force of Mortality Varies with Age, Lee-Carter Projections for the U.S. (Both Sexes Aggregated)

Observed Period	Estimated Relative Increase Rate (ln(*H*))	Estimated Exponential Base (*H*)	*G*	Sum of Squared Errors
1990	0.0874	1.0913	0.00006490713	0.00000062010
1995	0.0888	1.0929	0.00005608246	0.00000065998
2000	0.0902	1.0944	0.00004844479	0.00000069380
2005	0.0913	1.0955	0.00004296593	0.00000071035
2010	0.0920	1.0963	0.00003880330	0.00000071681
2015	0.0929	1.0973	0.00003470414	0.00000071989
2020	0.0938	1.0984	0.00003091813	0.00000071708
2025	0.0951	1.0997	0.00002701901	0.00000070881
2030	0.0961	1.1008	0.00002399331	0.00000069556
2035	0.0967	1.1015	0.00002181772	0.00000067902
2040	0.0980	1.1029	0.00001901224	0.00000065816
2045	0.0987	1.1037	0.00001722527	0.00000063575
2050	0.0991	1.1042	0.00001580330	0.00000061484
2055	0.1017	1.1071	0.00001267867	0.00000059040
2060	0.1027	1.1082	0.00001128283	0.00000056332
2065	0.1013	1.1066	0.00001165725	0.00000053918

Note: Force of mortality assumed to be constant over each one-year life period.
Source: Calculated with central mortality rates estimated with parameters from Lee and Carter (1992).
Force of mortality assumed to follow the function $G \cdot \exp(x \cdot \ln(H))$ for the range of ages 32–67 (MLS criteria).

ages is that of Thatcher (1999), based on the Perks formula. Alho and Spencer (1985) point out that with an open interval for age 85 and over in official statistics, the variances of mortality rates increase very steeply.

The plateau at advanced ages suggested by SOA (1995, pp. 875–76) and by Thatcher (1999) is not captured by the widely used Gompertz term GH^x, which appears, for example, in the Heligman-Pollard curve. Heligman and Pollard (1980, p. 60) have suggested variants of that term, which have been tested by Forfar and Smith (1988, p. 128) for English life tables, but this leads to a sacrificing of a reasonable interpretation of the parameters involved. The formula proposed by Perks and then by Thatcher (which also arises from the frailty model of Vaupel, Manton, and Stallard 1979) combines a Gompertz progression for middle-aged groups and a level curve for oldest ages in a logistic equation.

Coale and Guo (1989) describe a widely accepted procedure in demography for "closing out" mortality rates at the oldest ages, in which the observed mortality rates at the extreme ages (e.g., for the open-ended age group 85+) are replaced by a sequence of extrapolated mortality rates (e.g., at ages 85–89, 90–94, . . . , 105–9: Lee and Carter 1992; Lee 2000). The procedure uses the assumption of a steady decrease in the rate of increase in mortality rates with ages 75 and above. Renshaw and Haberman (2003b) investigate the suitability of this approach for the United Kingdom.

Carrière's model is flexible enough to incorporate any specific pattern for the elderly as what he calls a "cause of mortality." Unreliability in the data available, however, has restricted their application to ages under 85.

5.6 Sensitivity Analyses and Measures of Uncertainty for the Projections

Sensitivity analyses and measures of uncertainty for projections are not discussed in the published actuarial applications, despite the sophisticated methods traditionally used for static graduations. In official, published projections, scenario testing is normally used, although the use of scenarios for measuring uncertainty has been subject to two main criticisms:

1. Construction of scenarios generally involves defining a combination of separated forecasts without necessarily considering their possible interrelations, as discussed in Section 4.2 (Alho and Spencer 1985; Tuljapurkar and Boe 1998).
2. Empirical research in the United States shows that intervals bounded by high and low scenarios in official projections are much narrower than one-standard-deviation prediction intervals (Alho and Spencer 1985).

For the Lee-Carter model, the relative simplicity with which prediction intervals are estimated relies on the crucial assumption that the variability of the trend variable k_t dominates any other source of error, namely, the error associated with parameter estimation and the error in the fit of the equation. For relatively short projections, the authors recognize that this does not necessarily apply (Lee and Carter 1992, p. 670).

5.7 Accuracy

The performance of official forecasts in the United Kingdom and the United States and of pensions mortality forecasts during the last decades demonstrates that reductions in mortality rates have been systematically understated (see Sections 2.7 and 3.1).

5.8 Reasonable Length of Forecasting Periods

Except for Alho and Spencer (1985) and Booth, Maindonald, and Smith (2002), there is a lack of discussion about the reasonable length of forecasting periods. Some guidance can be obtained from the prediction intervals under the Lee-Carter model. Usually the length of projection horizons in official projections and actuarial applications is set without explicit justification.

5.9 Irregularities from Female Data

Female data usually behave less smoothly than male experience. There are exceptions, for example, the U.K. annuitants data set collected by the CMI Bureau where the female data set is larger than that for males. Despite this, male patterns are often used to make generalizations for the whole population. The Lee-Carter model has tended to be applied only to aggregated data, without allowing for sex differentials, but see Renshaw and Haberman (2003a, 2003b, 2003c) for applications to the two genders separately.

5.10 Cohort Effects

Although in many cases cohort effects influence improvement trends, techniques are based primarily on period considerations. Recent official projections are the exception (see Section 2.4). As Benjamin and Pollard (1993, p. 32) point out:

> The normal life table, based on the deaths over a limited period, is not likely to be reproduced in the future; it mixes the experience of different generations—the lives who contribute to the death rates at advanced ages were born many decades before those who contribute to the rates at young ages.

5.11 Pure Extrapolative Methods and Expert Opinion

As with most methods used in official forecasts, actuarial applications and other scientific methods also involve a mixture of extrapolation and expert opinion. From the discussions so far, we can conclude the following:

1. Models relying mainly on expert opinion (official projections and CMIB current methodology) have incurred systematic underestimation of mortality improvements to date. On the other hand, the Lee-Carter model has overestimated gains in life expectancy, although errors lie within the estimated prediction intervals (Lee 1997, p. 6).
2. Extrapolative models implicitly include expert opinion because of the process that leads to the selection of a mathematical formulation. Smoothness and the shape of future trends are subjective aspects depending on the analysis of past experience and personal beliefs. As London (1985, p. 5) notes, by definition, any graduation process contains some degree of subjectivity: "Undoubtedly, the element of prior opinion most frequently used by actuaries in graduating mortality (or other decrement) rates has been the belief that the true rates form a smooth sequence in some sense. This originally intuitive belief is certainly supported by empirical evidence." For example, the adjustment of the degree of the polynomial on which the Renshaw-Sithole model is based in order to balance adherence and smoothness relies is in itself a subjective exercise and could be considered as involving "expert opinion."

There seems to be no evidence as yet available that inclusion of expert opinion in a more direct way might result in better projections. As Alho (1990, p. 668) remarks, "[it has been] found that the use of experts . . . hindered rather than helped the forecasts in the past, in the sense that statistical time series models would have performed better."

6. CONCLUDING COMMENT

Clearly there is a need for regular studies of mortality trends at all ages. For actuarial applications, the trends for adult ages are particularly important. For pensions and annuity applications, the modeling of trends at postretirement ages cannot be ignored. Studies should focus on both the modeling and the forecasting of trends and need to be based on relevant data of good quality. As concluded by Oeppen and Vaupel (2002), there is strong research-based evidence that record life expectancy (among countries) has increased almost linearly over a 160-year history since 1840 and that mortality rates are continuing to decline at a steady rate so that "the belief that the expectation of life cannot rise much further" and that the underlying trend in the rates will flatten out are both misconceived.

ACKNOWLEDGMENT

We would like to thank Dr. Terry Sithole for permission to use Figure 4 from her doctoral thesis.

REFERENCES

ALHO, J. M. 1983. "Uncertain Population Forecasting." Ph.D. thesis, Northwestern University.

———. 1990. "Stochastic Methods in Population Forecasting." *International Journal of Forecasting* 6: 521–30.

ALHO, J. M., AND B. SPENCER. 1985. "Uncertain Population Forecasting." *Journal of the American Statistical Association* 80: 306–14.

———. 1990. "Effects of Targets and Aggregation on the Propagation of Error in Mortality Forecasts." *Mathematical Population Studies* 2: 209–27.

BALLOTTA, L., AND S. HABERMAN. 2003. “Valuation of Guaranteed Annuity Conversion Options.” *Insurance: Mathematics and Economics* 33: 87–108.

BELL, F. 1997. *Social Security Area Population Projections: 1997.* Actuarial Study No. 112. Washington, DC: Social Security Administration.

BELL, F., A. WADE, AND S. GOSS. 1992. *Life Tables for the United States Social Security Area 1900–2080.* Actuarial Study No. 107. Washington, DC: Social Security Administration.

BELL, W. R. 1997. “Comparing and Assessing Time Series Methods for Forecasting Age-Specific Fertility and Mortality Rates.” *Journal of Official Statistics* 13: 270–303.

BENJAMIN, B., AND J. H. POLLARD. 1993. *The Analysis of Mortality and Other Actuarial Statistics.* London: Institute and Faculty of Actuaries.

BERIN, B., G. STOLNITZ, AND A. TENENBEIN. 1989. “Mortality Trends of Males and Females over the Ages.” *Transactions of the Society of Actuaries* 41: 9–27.

BOOTH, H. J., J. MAINDONALD, AND L. SMITH. 2002. “Applying Lee-Carter under Conditions of Variable Mortality Decline.” *Population Studies* 56: 325–36.

BRASS, W. 1971. “On the Scale of Mortality.” In *Biological Aspects of Demography,* ed. W. Brass, pp. 69–110. London: Taylor and Francis.

BROUHNS, N., M. DENUIT, AND J. K. VERMONT. 2002. “A Poisson Log-Bilinear Regression Approach to the Construction of Projected Life-Tables.” *Insurance: Mathematics and Economics* 31: 373–93.

BUTT, Z., AND S. HABERMAN. 2004. “Applications of Frailty-Based Mortality Models using Generalized Linear Models.” *ASTIN Bulletin,* 34: Forthcoming.

CARRIÈRE, J. F. 1992. “Parametric Models for Life Tables.” *Transactions of the Society of Actuaries* 44: 77–99.

———. 1995. “Discussion of the 1994 Uninsured Pensioner Mortality Table.” *Transactions of the Society of Actuaries* 47: 915–18.

CHARLTON, J. 1997. *Trends in All-Cause Mortality 1841–1994. The Health of Adult Britain, 1841–1994.* Vol. 1, Decennial Supplement No. 12, pp. 17–29. London: Office for National Statistics.

CHARLTON, J., AND M. MURPHY. 1997. *Trends in Causes of Mortality: 1841–1994, An Overview. The Health of Adult Britain, 1841–1994.* Vol. 1, Decennial Supplement No. 12, pp. 30–57. London: Office for National Statistics.

COALE, A., AND G. GUO. 1989. “Revised Regional Model Life Tables at Very Low Levels of Mortality.” *Population Index* 55: 613–43.

CONGDON, P. 1993. “Statistical Graduation in Local Demographic and Analysis Projection.” *Journal of the Royal Statistical Society, Series A* 156: 237–70.

CONTINUOUS MORTALITY INVESTIGATION BUREAU. 1990. *Standard Tables of Mortality Based on the 1979–82 Experiences.* CMI Report no. 10. London: Institute of Actuaries and Faculty of Actuaries.

———. 1998. *Proposed New Tables for Life Office Pensioners, Normal, Male and Female, Based on the 1991–94 Experience.* CMI Report No. 16, pp. 113–41. London: Institute of Actuaries and Faculty of Actuaries.

———. 1999. *Standard Tables of Mortality Based on the 1991–94 Experiences.* CMI Report No. 17. London: Institute of Actuaries and Faculty of Actuaries. Electronic version: http//www.actuaries.org.uk.

DAYKIN, C. D. 2000. “The Recent Trend of Mortality in the United Kingdom.” *British Actuarial Journal* 6: 861–71.

DIMITROVA, D., S. HABERMAN, V. KAISHEV, R. KRACHUNOV, AND J. SPREEUW. 2003. “Quantifying the Effect of Partial and Complete Disease Elimination in the Presence of Dependency.” Working paper, Cass Business School, City University, London.

EVANS, J. 1998. “Mortality, Behold and Fear.” In *Life, Death and Money,* ed. D. F. Renn, pp. 29–42. London: Institute and Faculty of Actuaries.

FORFAR, D., J. MCCUTCHEON, AND A. WILKIE. 1988. “On Graduation by Mathematical Formula.” *Journal of the Institute of Actuaries* 115: 1–149.

FORFAR, D., AND D. SMITH. 1988. “The Changing Shape of English Life Tables.” *Transactions of the Faculty of Actuaries* 40: 98–134.

FREES, E. 1999. Personal communication to S. Haberman, School of Business, University of Wisconsin.

FRIEDLAND, R. 1998. “Life Expectancy in the Future: A Summary of Discussion among Experts.” *North American Actuarial Journal* 2(4): 48–63.

GALLOP, A. 1998. *Projecting Future Mortality: Methodologies Adopted in the Population Projections for the United Kingdom.* London: Government Actuary’s Department.

———. 2002. “Demographic Trends.” Presented to Ageing Population—Burden or Benefit?, Institute and Faculty of Actuaries’ International Conference, Edinburgh.

GOSS, S., A. WADE, AND F. BELL. 1998. “Historical and Projected Mortality for Mexico, Canada, and the United States.” *North American Actuarial Journal* 2(4): 108–26.

GUTTERMAN, S., AND I. VANDERHOOF. 1998. “Forecasting Changes in Mortality: A Search for a Law of Causes and Effects.” *North American Actuarial Journal* 2(4): 135–38.

HELLIGMAN, L., AND J. POLLARD. 1980. “The Age Pattern of Mortality.” *Journal of the Institute of Actuaries* 107: 49–75.

JENKINS, W. A., AND E. A. LEW. 1949. “A New Mortality Basis for Annuities.” *Transactions of the Society of Actuaries* 1: 369–466.

LEE, R. 1997. “The Lee-Carter Method for Forecasting Mortality, with Various Extensions and Applications.” University of California.

———. 2000. “The Lee-Carter Method for Forecasting Mortality, with Various Extensions and Applications.” *North American Actuarial Journal* 4: 80–93.

LEE, R., AND L. CARTER. 1992. “Modelling and Forecasting U.S. Mortality.” *Journal of the American Statistical Association* 87: 659–71.

LEE, R., AND T. MILLER. 2001. “Evaluating the Performance of the Lee-Carter Approach to Modelling and Forecasting.” *Demography* 38: 537–49.

LEE, R., AND S. TULJAPURKAR. 1997. “Death and Taxes: Longer Life, Consumption and Social Security.” *Demography* 34: 67–81.

LONDON, D. 1985. *Graduation: The Revision of Estimates.* Westport, CT: ACTEX Publications.

MCNOWN, R. 1992. “Comment on Modelling and Forecasting Mortality by Lee and Carter.” *Journal of American Statistical Association* 87: 671–72.

MCNOWN, R., AND A. ROGERS. 1989. “Forecasting Mortality: A Parameterized Time Series Approach.” *Demography* 26: 645–60.

MURPHY, M. 1995. “The Prospect of Mortality: England and Wales and the United States of America, 1962–1989.” *British Actuarial Journal* 1: 331–50.

OEPPEN, J., AND J. W. VAUPEL. 2002. “Broken Limits to Life Expectancy.” *Science* 296: 1029–30.

OLSHANSKY, S. J. 1988. “On Forecasting Mortality.” *Milbank Quarterly* 66: 482–530.

ONS. 1995. *1992-Based National Population Projections.* Series PP2 No. 19, Office of Population Censuses and Surveys. London: Government Statistical Service.

———. 1996. *1994-Based National Population Projections.* Series PP2 No. 20, Office for National Statistics. London: Government Actuary’s Department.

———. 1999. *1996-Based National Population Projections.* Series PP2 No. 21, Office for National Statistics. London: Government Actuary’s Department.

PERKS, W. 1932. “On Some Experiments in the Graduation of Mortality Statistics.” *Journal of the Institute of Actuaries* 63: 12–40.

RENSHAW, A. E., S. HABERMAN, AND P. HATZOPOLOUS. 1996. “Modelling of Recent Mortality Trends in UK Male Assured Lives.” *British Actuarial Journal* 2: 449–77.

RENSHAW, A. E., AND S. HABERMAN. 2000. “Modeling for Mortality Reduction Factors.” Actuarial Research Paper No. 127. Department of Actuarial Science and Statistics, City University, London.

———. 2003a. “On the Forecasting of Mortality Reduction Factors.” *Insurance: Mathematics and Economics* 32: 379–401.

———. 2003b. “Lee-Carter Mortality Forecasting, a Parallel GLM Approach, England and Wales Mortality Projections.” *Journal of Royal Statistical Society Series C (Applied Statistics)* 52: 119–37.

———. 2003c. “Lee-Carter Mortality Forecasting with Age Specific Enhancement.” *Insurance: Mathematics and Economics* 33: 255–72.

ROSENBERG, M., AND W. LUCKNER. 1998. “Summary of Results of Survey of Seminar Attendees.” *North American Actuarial Journal* 2(4): 64–82.

SHAW, C. 1994. “Accuracy and Uncertainty of the National Population Projections for the United Kingdom.” *Population Trends* No. 77, autumn, Office of Population, Censuses and Surveys. London: Government Statistical Service.

———. 1998. “1996-Based National Population Projections for the United Kingdom and Constituent Countries.” *Population Trends* No. 91, spring. London: HMSO.

SITHOLE, T. Z. 2003. “Projection of Mortality Rates with Specific Reference to Immediate Annuitants and Life Office Pensioners.” Ph.D. thesis, City University, London.

SITHOLE, T. Z., S. HABERMAN, AND R. J. VERRALL. 2000. “An Investigation into Parametric Models for Mortality Projections, with Applications to Immediate Annuitants’ and Life Office Pensioners’ Data.” *Insurance: Mathematics and Economics* 27: 285–312.

SOCIETY OF ACTUARIES GROUP ANNUITY VALUATION TABLE TASK FORCE. 1995. “1994 Group Annuity Mortality Table and 1994 Group Annuity Reserving Table.” *Transactions of the Society of Actuaries* 17: 865–913.

THATCHER, A. 1999. “The Long-Term Pattern of Adult Mortality and the Highest Attained Age.” *Journal of the Royal Statistical Society, Series A* No. 162: 5–43.

TULJAPURKAR, S., AND C. BOE. 1998. “Mortality Change and Forecasting: How Much and How Little Do We Know?” *North American Actuarial Journal* 2(4): 13–47.

VAUPEL, J. W., K. G. MANTON, AND STALLARD, E. 1979. “The Impact of Heterogeneity in Individual Frailty on the Dynamics of Mortality.” *Demography* 16: 439–54.

WILLETS, R. 1999. *Mortality Trends: An Analysis of Mortality Improvement in the UK.* London: General & Cologne Re.

WILMOTH, J. R. 1995. “Are Mortality Projections Always More Pessimistic When Disaggregated by Cause of Death?” *Mathematical Population Studies* 5: 293–319.

Discussions on this paper can be submitted until October 1, 2004. The authors reserve the right to reply to any discussion. Please see the Submission Guidelines for Authors on the inside back cover for instructions on the submission of discussions.