Population Studies, **49** (1995), 281–295
Printed in Great Britain

Are Mortality Rates Falling at Extremely High Ages? An Investigation Based on a Model Proposed by Coale and Kisker*

JOHN R. WILMOTH†

If there are limits to the continued reduction in human mortality, how do they operate? Some demographers and gerontologists have expressed the limit in terms of the human life span, which may (or may not) have changed in the past, and which may (or may not) increase to higher ages in the future. Others have argued in terms of life expectancy at birth, which has risen dramatically during the last century or more, but which may (or may not) attain an upper limit in coming years. Other investigations about limits to mortality reduction have analysed measures such as 'life endurancy' (the age to which some small proportion of a life-table population would survive), or the age pattern of mortality as a whole.[1]

A critical aspect of all these discussions is the magnitude of mortality decline that has been or will be observed at very high ages. It has now been well established that mortality rates for ages at least as high as 100 years have declined in industrialized countries during the past 40 years or more.[2] In general, the magnitude of these reductions has been proportionately smaller at older than at younger ages. If the proportionate reduction in mortality rates at ages above 100 years has been smaller still, then it is possible that there has been no (or very little) reduction in mortality rates at extremely high ages. No matter what measure of mortality is employed, knowledge of the trend in death rates at these ages is important in any discussion of possible limits to further mortality reduction.

The direct calculation of mortality rates at very high ages, however, poses substantial problems to demographers and actuaries. The first difficulty is excessive random fluctuation, due to the small number of individuals who survive to advanced ages. In direct calculations of age-specific death rates by single years of age, the relative error that can be attributed to random fluctuations alone will exceed an acceptable level above a

* An earlier version of this article was presented at the Research Workshop on Oldest–Old Mortality, held at Duke University on 4–6 March 1993, and later at the Annual Meeting of the Population Association of America, held in Cincinnati on 1–3 April 1993. The author acknowledges, in particular, the critical comments of James Vaupel, Roger Thatcher, and Ansley Coale; all errors and omission, however, are his responsibility alone. This research was funded, in part, by a research grant (R01 AG11552A-01) awarded by the National Institute on Aging.

† Department of Demography, University of California, Berkeley.

[1] James F. Fries, 'Aging, natural death, and the compression of morbidity', *The New England Journal of Medicine*, **303**, 3 (1980), pp. 130–135. Leonid A. Gavrilov and Natalia S. Gavrilova, *The Biology of Life Span: A Quantitative Approach* (New York, 1991). Kenneth G. Manton, Eric Stallard, and H. Dennis Tolley, 'Limits to human life expectancy: evidence, prospects, and implications', *Population and Development Review*, **17**, 4 (1991), pp. 603–637.

[2] Väinö Kannisto, *Development of Oldest–Old Mortality, 1950–1990: Evidence from 28 Developed Countries* (Odense, 1994).

certain age, which is generally between 95 and 105 years for national populations (depending on the size of the cohorts involved).[3]

The second difficulty is systematic bias due to inaccuracies in the reporting of age in census and vital registration data. At these ages, the primary inaccuracies can be attributed to heaping on ages which are multiples of five or ten, or to overstatement (or exaggeration) of age. Both age heaping and age exaggeration are known to have occurred in census and death registration data.[4]

A tendency toward age heaping can be fairly easily identified, and simple adjustments can be made to minimize its effect on calculated mortality rates.[5] It is more difficult, however, to identify and correct for age exaggeration. In some situations it may be possible for an experienced analyst to recognize that some proportion of the reported ages must have been overstated, but it is all but impossible to correct the raw data with confidence.

Mortality rates based on data which suffer from age overstatement will consistently underestimate the true level of mortality at very high ages.[6] The problem of age exaggeration is quite severe in data for the United States, especially for blacks. This reality has led some analysts to seek indirect means of estimating mortality rates at advanced ages, say, above 85 years. In this article I examine one method proposed for this purpose in 1990 by Ansley Coale and Ellen Kisker; I analyse the underlying assumptions of the model in light of an extensive body of relatively high-quality mortality data; I also fit the model to time series of Swedish and Japanese data using a different procedure from that proposed by Coale and Kisker.[7]

In general, this analysis supports the empirical validity of Coale and Kisker's model (hereafter, the CK model). The two assumptions on which it is based appear to be a close approximation of reality, at least insofar as can be verified by using directly calculated mortality rates for Sweden and Japan. First, the assumption that the rate of increase in age-specific mortality rates above 85 years declines as a linear function of age is verified by this analysis. The second assumption, namely, that the level of mortality at age 110 years is constant across time and for various populations, is more questionable, but it may be an adequate assumption for many purposes.

[3] The coefficient of variation for directly calculated mortality rates approximately equals $1/\sqrt{D_x}$, where D_x is the number of deaths in the numerator of the rate. Taking Japanese women as an example, the annual number of deaths at age 100 increased from 20 in the early 1950s to 350 in the late 1980s (all figures are approximate). Thus, the coefficient of variation of m_{100} decreased from 22 to 5 per cent. The age at which this coefficient of variation crossed a five per cent threshold rose during this period from 94 to 100 years. By combining data in five-year time periods, as I have done here, we encounter this threshold at slightly higher ages: between ages 97 and 103 for Japanese women. For Swedish men, the smallest of the populations considered here, the five-per cent threshold is reached at age 95 in 1951–55 and age 98 in 1986–90. A coefficient of variation of five per cent would be rather large and thus 'unacceptable' for many purposes, since in this situation an observed value of m_x would deviate from its true value by more than five per cent in one out of every three cases.

[4] Ansley J. Coale and Ellen E. Kisker, 'Defects in data on old-age mortality in the United States: new procedures for calculating mortality schedules and life tables at the highest ages', *Asian and Pacific Population Forum*, **4**, 1 (1990), pp. 1–31. Väinö Kannisto, 'On the survival of centenarians and the span of life', *Population Studies*, **42** (1988), pp. 389–406. François Depoid, 'La mortalité des grands vieillards', *Population*, 28, 4–5 (1973), pp. 755–792. Paul Vincent, 'La mortalité des vieillards', *Population*, **6**, 2 (1951), pp. 181–204. A. R. Thatcher, 'Centenarians', *Population Trends*, **25** (1981), pp. 11–14. A. R. Thatcher, 'Trends in numbers and mortality at high ages in England and Wales', *Population Studies*, **46** (1992), pp. 411–426.

[5] Henry S. Shryock and Jacob S. Siegel, *The Methods and Materials of Demography* (Washington, 1973). Jacques Vallin, *La Mortalité par Génération en France Depuis 1899* (Paris, 1973).

[6] The downward bias is due to two factors, both of which operate in the same direction: (1) since individuals inflate their true ages, mortality levels at younger ages are in effect transferred to older ages; (2) the magnitude of age exaggeration tends to be greater in census data (which provide the denominators of calculated mortality rates) than in vital registration data (numerators).

[7] Coale and Kisker, *op. cit.* in fn. 4.

This article begins by reviewing the logic that led to the derivation of the CK model. I then show that recent mortality data from Sweden and Japan are consistent with the first assumption regarding the shape of the age curve of mortality above 85 years. Next, the model is fitted using an alternative procedure to mortality data for these same two countries in order to test the assumption regarding the level of mortality at age 110. I then propose this alternative fitting procedure as a means of estimating, indirectly, the implied level of mortality at extremely high ages. Based on these investigations, I offer some tentative conclusions regarding trends in mortality rates around age 110.

DERIVATION OF COALE AND KISKER'S MODEL OF MORTALITY AT ADVANCED AGES

Coale and Kisker developed their new procedure 'for calculating mortality schedules and life tables at the highest ages'. The procedure predicts the level of mortality above 85 years based on observed mortality rates around age 85. The procedure relies upon a function called the *age-specific rate of mortality change with age*, or k_x. By definition,

$$k_x = \ln(m_x) - \ln(m_{x-1}). \tag{1}$$

If m_{65} is known, for example, it follows immediately that

$$m_x = m_{65} \exp\left[\sum_{y=66}^{x} k_y\right], \quad \text{for } x = 66, 67, \ldots. \tag{2}$$

If k_x is a constant function of age (i.e. $k_x = k$), then

$$m_x = m_{65} \exp[(x-65)k], \tag{3}$$

which is equivalent to Gompertz's familiar formula. In any case, note that if m_{65} is known and k_x is known as well for $x = 66, 67, \ldots$, then m_x is completely defined for ages 65 and above.

The first important assumption of the CK model is that the form of k_x above 85 years should be linear:

$$k_x = k_{85} + (x-85)s, \tag{4}$$

where s denotes the slope of the change in k_x over this age range. Thus, the model implies that

$$m_x = m_{84} \exp\left[\sum_{y=85}^{x} (k_{85} + (y-85)s)\right]. \tag{5}$$

Letting $x = 110$ and solving for the slope, s, we obtain

$$s = -\left[\ln\left(\frac{m_{84}}{m_{110}}\right) + 26k_{85}\right] \Big/ 325. \tag{6}$$

The slope of the k_x curve is now expressed as a function of mortality rates at ages 84 and 85 years (i.e. m_{84} and k_{85}) and the level of mortality at age 110 years (m_{110}). The second major assumption of the CK procedure is that m_{110} is fixed at 1.0 for males, and 0.8 for females. With this assumption, the slope, s, and thus the entire k_x curve above age 85, is completely defined given the level of mortality around age 85. To minimize the effects of random fluctuations, Coale and Kisker used the arithmetic average of mortality rates from ages 82 to 86 in place of m_{84} in the above equations; likewise, k_{85} is replaced by $[\ln(m_{88}/m_{81})]/7$. In this article, when I refer to the 'original CK procedure', it means that I have fitted the model using exactly this method.

Regarding the assumption of a linear change in k_x above age 85, Coale and Kisker justified this choice based on data presented in an earlier article.[8] For a number of countries with exceptionally good mortality data, Coale and his colleagues have shown that calculated values of k_x decline almost linearly with age above 85 years until at least age 100. Nonetheless, their analyses did not demonstrate that this pattern of linear decline continues above age 100. In the following section, I conclude that this pattern extends until at least age 105, and perhaps as high as 110 years.

The assumption that the level of m_{110} is fixed at 1.0 for males and 0.8 for females can have no direct verification. As noted earlier, it is nearly impossible to calculate values of m_{110} directly because of the small number of individuals who survive at that age. Coale and Kisker state that 'the absence of deaths above 110 implies that the assumed m_{110} of 1.0 [for males] is reasonable'. For females, the m_{110} of 0.8 is chosen to 'avoid imposing a crossover of male and female mortality at 110'.

In any case, Coale and Kisker note that the exact choice of a value for m_{110} does not have a large effect on the predicted level of mortality around ages 90 or 95 years. The values of m_{110} chosen for the original CK model, therefore, seem to be justified primarily by the fact that they produce reasonably accurate fits when applied to mortality data for a country like Sweden, where directly calculated rates above age 85 are thought to be quite accurate. Nonetheless, Coale and Kisker's comparison of their model to reliable, directly calculated mortality rates was apparently limited to the case of Swedish males and females during the years 1978–82.

Later in this article, it is shown that the model fits well in another country, Japan, and for other time periods. Using an alternative fitting procedure I also explore the assumption regarding the fixed level for m_{110}. The evidence in this latter case is mixed: for men in Sweden and Japan, it does not appear that the level of m_{110} has changed over time, whereas for women in both countries it seems to have fallen.

EMPIRICAL VERIFICATION OF A LINEAR PATTERN IN k_x

Empirical estimates of k_x are quite unstable due to random fluctuations. Therefore, define

$$_5k_x = \frac{\ln(_5m_x) - \ln(_5m_{x-5})}{5}. \tag{7}$$

It is possible to verify that a $_5k_x$ schedule provides a smoothed version of the corresponding k_x schedule. Age patterns of the $_5k_x$ function above 55 years are shown in Figures 1 and 2 for men and women in Sweden and Japan during the period 1971–90. We can be reasonably certain that the quality of age reporting in these two countries was quite high during this period.[9] A 20-year time interval was chosen to minimize random variation, although the (approximate) 95-per cent confidence bands shown in these figures demonstrate that aleatory fluctuations still preclude precise measurement at very high ages.[10]

These graphs support the conclusion that the decline in k_x is approximately linear

[8] Shiro Horiuchi and Ansley J. Coale, 'Age patterns of mortality for older women: an analysis using the age-specific rate of mortality change with age', *Mathematical Population Studies*, **2**, 4 (1990), pp. 245–267.

[9] John R. Wilmoth and Hans Lundström, 'Extreme longevity in five countries: presentation of trends with special attention to issues of data quality', manuscript (1994).

[10] Confidence bands are plus or minus two times the estimated standard error of $_5k_x$, which equals, approximately, $1/5\sqrt{[(1/_5D_x)+(1/_5D_{x-5})]}$, where $_5D_x$ is the number of deaths between ages x and $x+5$ recorded in the time period. This formula is based on the fact that the variance of $\ln(_5m_x)$ is approximately equal to $1/_5D_x$. See John R. Wilmoth, 'Variation in vital rates by age, period, and cohort', in C. C. Clogg (ed.), *Sociological Methodology*, **20** (1990), pp. 295–335. See also David R. Brillinger, 'The natural variability of vital rates and associated statistics', *Biometrics*, **42**, 4 (1986), pp. 693–734.

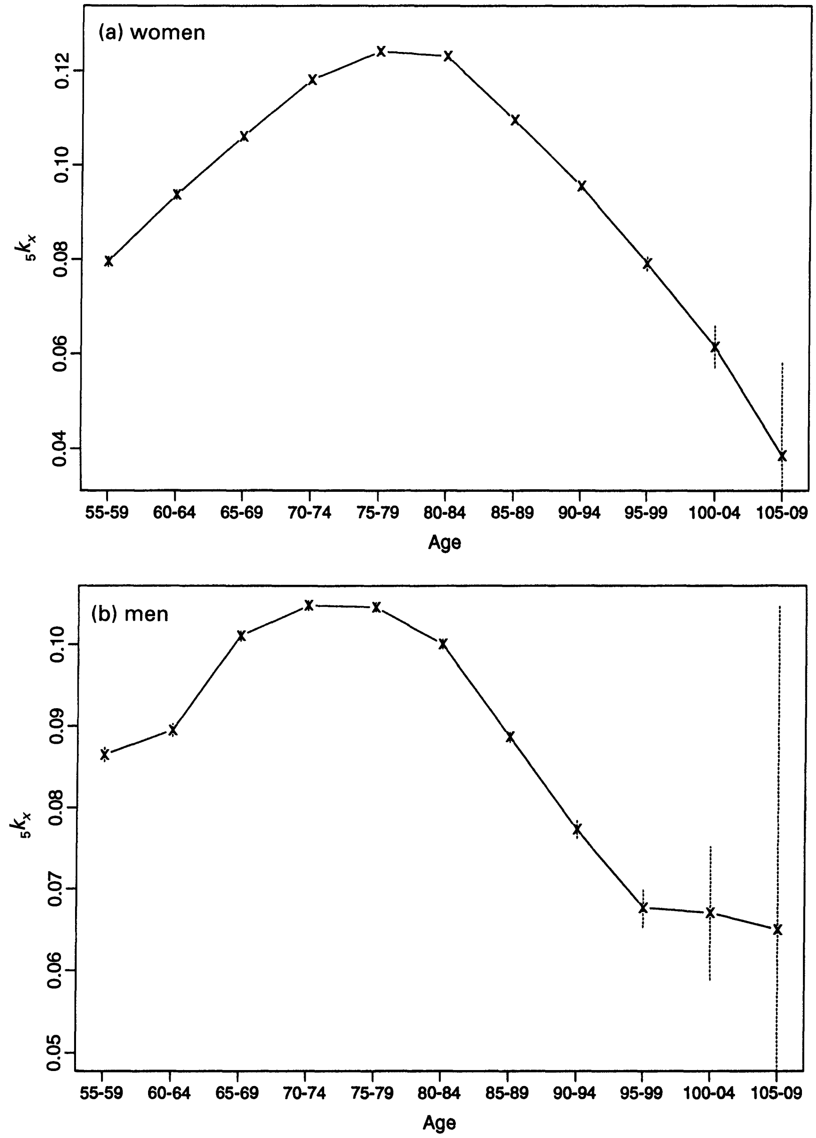

Figure 1. Age-specific rate of mortality change with age, $_5k_x$, Japan, 1971–90. (a) women, (b) men. Note: All confidence intervals shown here are symmetric. The scale of Figures 1 and 2 is restricted for presentation purposes, causing some estimates or intervals to lie outside the visible range.

from age 85 until at least age 105. Only the pattern for Japanese men deviates substantially from a linear pattern over this age range, but the width of the confidence intervals shows that this deviation could be the result of random variation. It is notable, furthermore, that in the largest of these three populations, Japanese women, the near-linear pattern appears to continue into the 105–109 age group.

Thus, the shape of the mortality curve, as measured by the k_x function, appears to be consistent with the model proposed by Coale and Kisker, at least on the basis of Swedish and Japanese mortality during 1971–90. I will provide further evidence, in a later section of this article, that the shape of the mortality curve implied by this model is reasonably consistent with Swedish data from 1896 to the present, and with Japanese data from 1951 to the present.

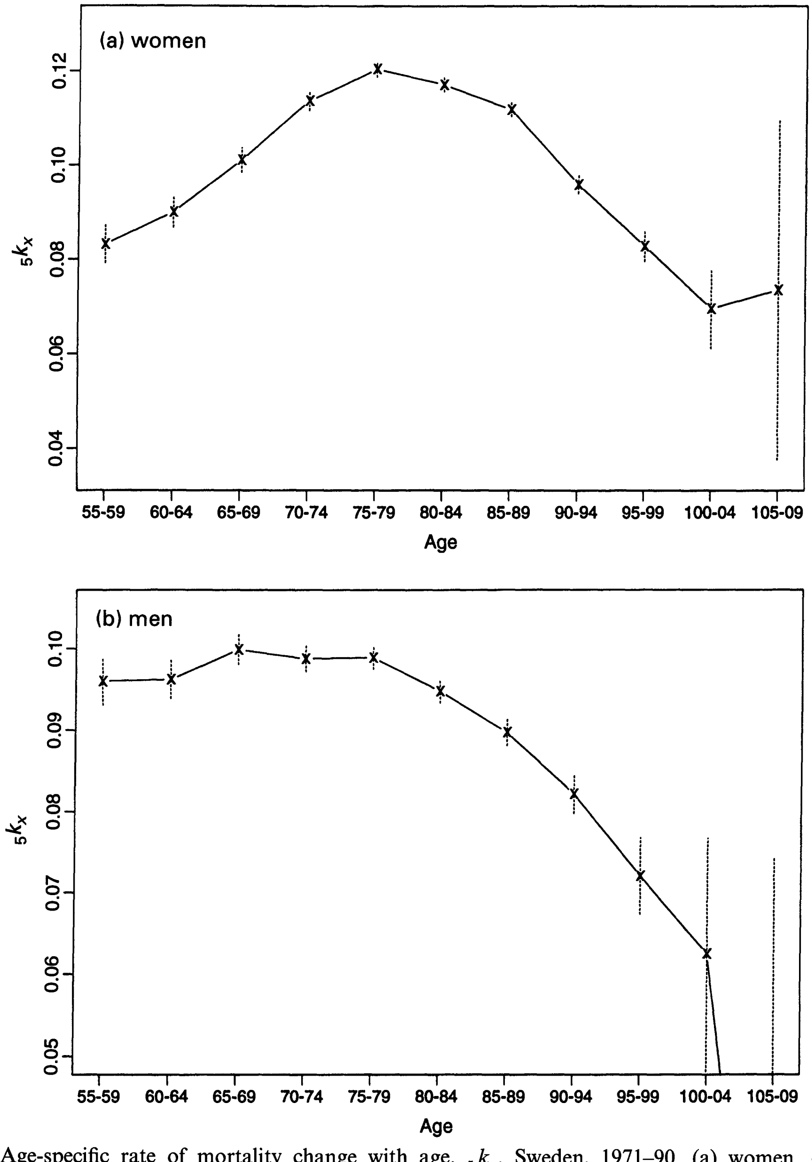

Figure 2. Age-specific rate of mortality change with age, $_5k_x$, Sweden, 1971–90. (a) women, (b) men. Note: See note for Figure 1.

The second assumption of the CK procedure regarding the level of mortality at age 110 is somewhat less certain and requires a more thorough investigation. In the following section, I propose an alternative means of fitting the model that relaxes the assumption about a fixed level for m_{110} but retains the assumption of a linear pattern for k_x.

ALTERNATIVE FITTING PROCEDURE

Note that the CK model, as contained in Equation (5) above, can be re-written as

$$m_x = m_{84} \exp\left\{(x-84)\,k_{85} + \frac{(x-84)(x-85)}{2}s\right\}. \tag{8}$$

Taking logarithms on both sides of this equation and writing it as a regression model, we obtain

$$\ln(m_x) = \ln(m_{84}) + (x-84)k_{85} + \frac{(x-84)(x-85)}{2}s + \varepsilon_x. \tag{9}$$

Coale and Kisker proposed that their model could be fitted based merely on observed values of m_{84} and k_{85}. If mortality data at high ages are thought to be accurate for a given population, however, a better fit can be obtained by using the information contained in the age-specific mortality rates above age 85, rather than basing the fit on measures of mortality around that age alone. Since Coale and Kisker developed their procedure for use in situations where mortality data above age 85 are thought to be unreliable (for example, the United States), they naturally did not consider the method proposed here. The advantage of this alternative fitting procedure, clearly, is that it does not require any assumption about the level of mortality at some very high age. Indeed, it provides a means of estimating mortality rates above age 100 through extrapolation, as I discuss in the next section of this paper.

It is apparent from Equation (9) that the CK model is essentially a quadratic function fitted to the logarithm of the age-specific death rates. In the original procedure, the model is fitted by finding the (unique) quadratic curve that passes through three points, namely the logarithms of m_{84}, m_{85}, and m_{110}.[11] In the alternative procedure proposed here, the quadratic model is fitted to the entire schedule of age-specific death rates above age 85, or to death rates between ages 85 and 100.

Using mortality data for Sweden and Japan, the regression model in Equation (9) was fitted by a variety of techniques.[12] If the error term in this equation follows a Gaussian distribution (though with unequal variances), weighted least squares (WLS) yields the optimum parameter estimates according to standard statistical theory (in the sense that they will have minimum variance within the class of all unbiased estimators). The weights used in this situation should be the reciprocal of the variance of $\ln(m_x)$, which can be estimated by the predicted number of deaths, $\hat{D}_x = E_x \hat{m}_x$, where E_x is the observed exposure to risk at age x. Since $\hat{D}_x$ is a product of the fitting procedure, it is necessary to employ an iterative fitting technique, whereby the weights employed at step n are the predicted deaths at step $n-1$. (It is advisable to use the observed number of deaths as the initial set of weights.)

There were at least two motivations for considering a variety of fitting procedures. First, from a theoretical standpoint, there is no reason to believe that the errors in Equation (9) truly follow a Gaussian distribution, and thus WLS may not be optimal in any formal sense. Nevertheless, the parameter estimates yielded by two robust fitting techniques, least absolute deviations and a bi-weight version of WLS, were very close to the WLS estimates. Therefore, the accuracy of the Gaussian assumption does not seem particularly important, and we can comfortably focus our attention on the WLS results.

[11] Coale and Kisker fit the model using m_{84} and k_{85}, but this procedure is equivalent to using m_{84} and m_{85}. Furthermore, they propose that robust estimates of m_{84} and k_{85} should be used in fitting this curve in order to minimize the influence of random fluctuations. Thus, their procedure amounts to fitting a quadratic curve through (the logarithms of) robust estimates of m_{84} and m_{85}, and an assumed value of m_{110}.

[12] In all, I have fitted the model using seven different procedures: (1) the original method proposed by Coale and Kisker; (2) weighted least squares for ages 85–ω (ω is defined as the maximum age, for some population and time period, such that there are non-zero death and exposure counts for all ages less than or equal to ω); (3) weighted least squares for ages 85–100; (4) least absolute deviations for ages 85–ω; (5) least absolute deviations for ages 85–100; (6) a modified version of weighted least squares in which the usual weights are multiplied by a bi-weight of the residual (thus minimizing the impact of unusually large residuals; see Frederick Mosteller and John W. Tukey, *Data Analysis and Regression* (Reading, 1977)); and (7) ordinary least squares. All these methods produce largely similar results, except for ordinary least squares, which is heavily influenced by random fluctuations at high ages.

Rather than statistical technique itself, a more important issue was the age range employed in fitting the model. The model was fitted using data both above and below 100 years, and also using data for ages 85–100 only. In the following sections, I present results based on these two methods. It is apparent from this analysis that the restriction to ages below 100 years usually does not change the results significantly in the case of Sweden, although the difference between the two methods is sometimes important in analysing the Japanese data. Results from the other fitting procedures are not presented (with the exception of the original CK procedure), since they are quite similar to those obtained by weighted least squares.

INDIRECT ESTIMATION OF m_{110}

As noted earlier, Coale and Kisker proposed the model considered here for the very specific purpose of estimating mortality rates above age 85 in populations where directly calculated rates are thought to be unreliable in this age range. For this reason, it was necessary for them to make an assumption regarding the level of mortality at some very high age. They chose, somewhat arbitrarily, to set m_{110} equal to 1.0 for males and 0.8 for females. The implication of this choice is that mortality rates at this age are constant across space and time (though not sex).

If mortality rates at age 110 (and presumably at higher ages as well) are not declining with time, then progress against mortality would be limited to increasing the chance of survival up to this age. It is possible to show that, if mortality rates are declining below age 110, but constant at that age and above, the survival curve will become progressively more 'rectangular' over time, and gains in life expectancy will decrease as e_0 approaches a limit (which still could be as high as 110 plus e_{110}, which is presumed to be constant).

Coale and Kisker did not mention any of these implications when they proposed their model. Indeed, I do not believe it was their intention to comment upon these issues, either implicitly or explicitly. Nonetheless, their model provides, in part because of its simple form, an excellent vehicle for exploring many of these topics. In this article, I consider only the question whether mortality rates above age 100, and specifically at age 110, have been declining over time. It should also be possible to use the CK model to explore the issue of rectangularization of the survival curve and other topics as well.

If the CK model is fitted by the alternative procedure proposed in this paper, an estimate of m_{110} can be found by simple extrapolation of the predicted mortality curve. In many cases these are purely theoretical values, since persons of such advanced age are extremely rare even today, and were probably much less common in the past.[13] For this reason, I refer to these estimates as 'implied values' of m_{110}, in the sense that they are the values of the death rate at this age that are implied by the fitted model. Presumably, they are also reasonable estimates of the value of m_{110} that would have been observed if a large number of individuals had actually been alive in the population at these ages. The purpose of making these rather theoretical calculations is to assess the plausibility of a model of mortality change where death rates at some very high age (presumably in the neighbourhood of 110 years) are not declining with time.

The results of this exercise are mixed. For the two national populations considered here, the implied level of m_{110} appears to have declined for women over the observation periods, but not for men. Figure 3 presents the different implied values of m_{110} for Swedish women and men from 1896–1900 to 1986–90, and Figure 4, for Japanese women and men from 1951–55 to 1986–90. These estimates are based on three fitting techniques: the original CK procedure (which holds the level of m_{110} constant at 0.8 for

[13] Wilmoth and Lundström, *op. cit.* in fn. 9.

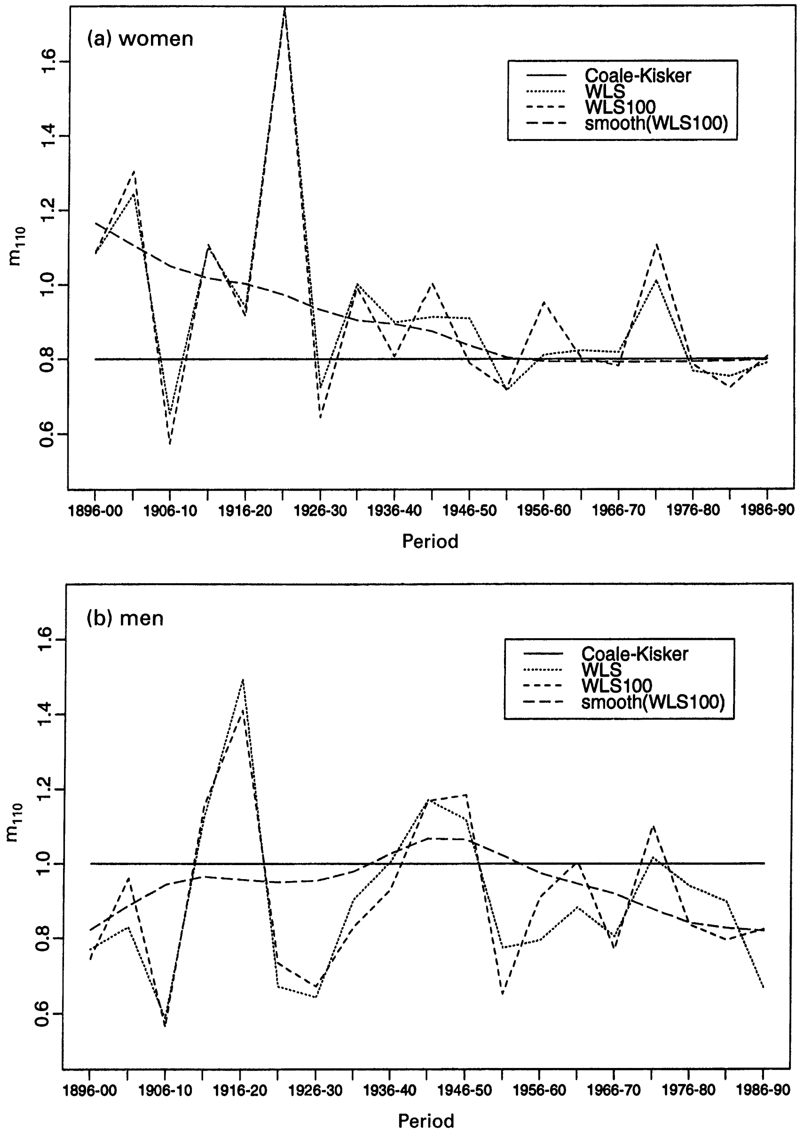

Figure 3. Implied (or extrapolated) value of m_{110}, Sweden, 1896–1900 to 1986–90. (a) women, (b) men. Comparison of three fitting procedures: (1) Coale–Kisker, (2) WLS, all ages 85–ω, (3) WLS, ages 85–100 only.

women and 1.0 for men), and two fitting procedures based on weighted least squares (which differ only in the age range employed, as explained in fn. 12). The values of 0.8 and 1.0 for m_{110} are presented for purposes of comparison only. Thus, the following discussion is focused on the values of m_{110} derived from the two weighted least squares procedures, which are similar in most cases. (Since the Swedish tends are strongly affected by random variations, a smoothed version of the second WLS fitting procedure is also shown in Figure 3.)

In Sweden, the implied value of m_{110} from two WLS fits shows no long-term trend for men. Its level is slightly below the value suggested by Coale and Kisker (its average level for the 19 five-year time periods considered here is about 0.94). For Swedish women, on the other hand, a clear downward trend is apparent over the first half of this period:

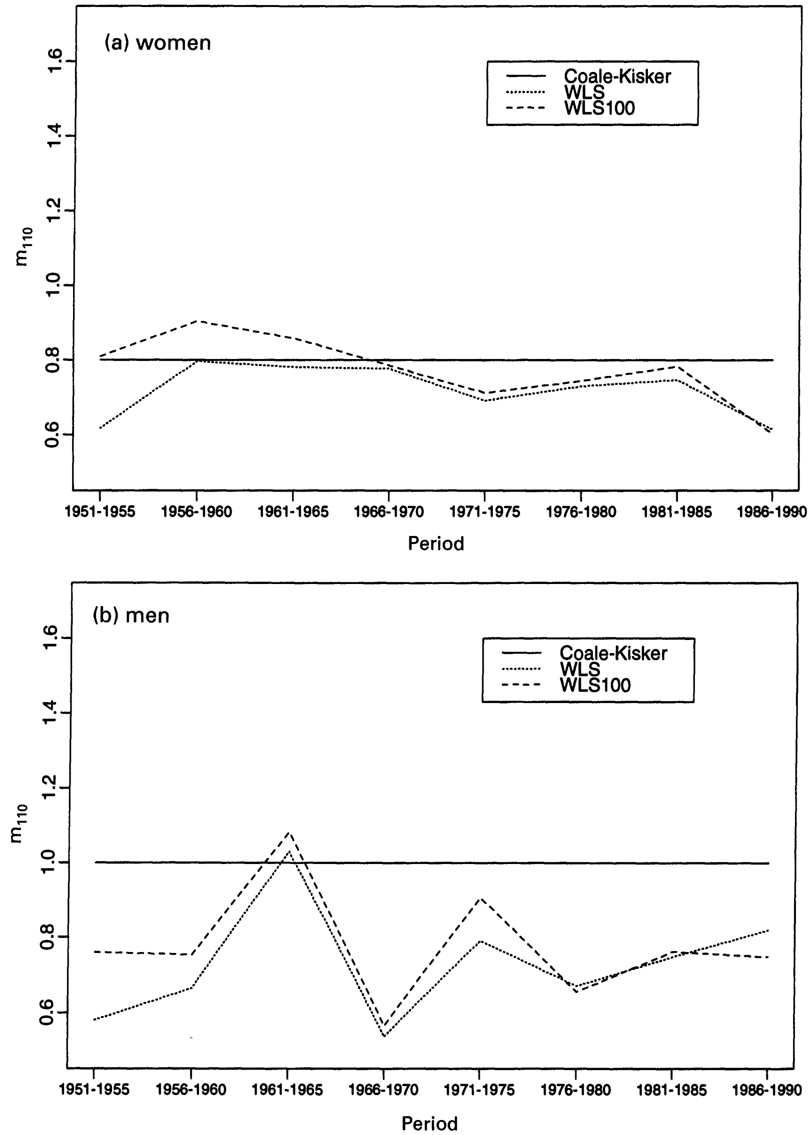

Figure 4. Implied (or extrapolated) value of m_{110}, Japan, 1951–55 to 1986–90. (a) women, (b) men. Comparison of three-fitting procedures: (1) Coale–Kisker, (2) WLS, all ages 85–ω, (3) WLS, ages 85–100 only.

from an initial level in the neighbourhood of 1.2, to an ultimate value around 0.8. This change represents a decline in age-specific mortality rates equal to approximately one-third of their initial level. Over the entire 95-year observation period, this magnitude of mortality improvement compares favourably with changes observed at younger ages (for example, directly calculated mortality rates for Swedish women in age groups 80–84, 85–89, 90–94, and 95–99, declined by 50, 44, 36, and 28 per cent, respectively, between 1896–1900 and 1986–90).

In terms of an annualized rate of decline, the decrease in m_{110} for Swedish women is approximately 0.4 per cent per year over the entire period (obviously, the rate of decrease was faster in the early part of this period). This rate is still well below the annual pace of improvement assumed in some recent mortality forecasts. For example, the most optimistic of three 'alternative projections' of U.S. mortality assumes a constant rate of decline of two per cent for mortality rates at all ages (including those above age 100). Even the middle scenario assumes an invariant annual rate of improvement of one per cent.[14]

Compared with the Swedish findings, the results based on Japanese data show both similarities and differences. In Figure 4, it is immediately noticeable that the estimates of m_{110} derived from the two WLS procedures are less closely in agreement for Japan than was the case for Sweden. For both procedures, however, the pattern for Japanese men is consistent with the conclusion that the implied level of m_{110} has not changed substantially over the 40-year observation period. For Japanese women, on the other hand, a conclusion about the presence or absence of a secular decline clearly depends on the fitting procedure employed: the WLS procedure based on all available observations suggests no overall trend in m_{110}, whereas the WLS procedure based only on observations for ages 85–100 suggests a rather certain, albeit slight, downward trend.

It is likely that the divergence between the two fitting procedures is explained by the presence of age exaggeration in the Japanese mortality data. Age exaggeration, if present, would produce a downward bias in mortality rates above age 100 (and perhaps below this age as well). The bias would be stronger at higher ages, and thus the inclusion of ages over 100 years in a WLS procedure could bias the estimates of m_{110} downward, as seen in Figure 4. Therefore, a plausible explanation is that the difference in the implied values of m_{110} for the two WLS procedures is due to inadequacies in the original data. This conclusion is supported by the results of a separate investigation which shows that age exaggeration is readily apparent in the Japanese data during the period before about 1970, but not afterwards.[15] The greater agreement between the two WLS procedures after 1970, combined with knowledge that data quality in Japan improved around this time, is consistent with the explanation that the difference in the two WLS fits can be attributed to poor data quality above age 100 before 1970.

Thus, we may safely conclude that the pattern of a secular decline in the implied level of m_{110} observed for women in Sweden holds for women in Japan as well. At the same time, the absence of a trend in the implied m_{110} is observed for men in both countries. It is quite notable, however, that the level of m_{110} tends to be somewhat lower for both men and women in Japan than in Sweden. If we compare its level during the overlapping periods (1951–55 to 1986–90), we find that in Japan m_{110} is lower, on average, than in Sweden by 0.06 for women and 0.08 for men. There are at least three possible explanations for this difference. First, it is possible that age exaggeration in the Japanese data has produced a downward bias in mortality rates even below age 100, and thus the lower implied values of m_{110} are the result of extrapolating from faulty data. Secondly, it is possible that genetic and/or behavioural differences between the two populations are responsible for a real difference in the implied levels of mortality at these ages. Thirdly and finally, higher mortality early in life among the currently elderly Japanese cohorts may have selected for greater longevity among survivors. It seems impossible to speculate further about this pattern on the basis of these findings alone, although they certainly merit further investigation.

GOODNESS OF FIT

Although the results presented in the previous sections of this article are provocative, they may be unconvincing unless it can be demonstrated that the model on which they

[14] Dennis Ahlburg and James W. Vaupel, 'Alternative projections of the U.S. population', *Demography*, **27**, 4 (1990), pp. 639–652.

[15] Wilmoth and Lundström, *op. cit.* in fn. 9.

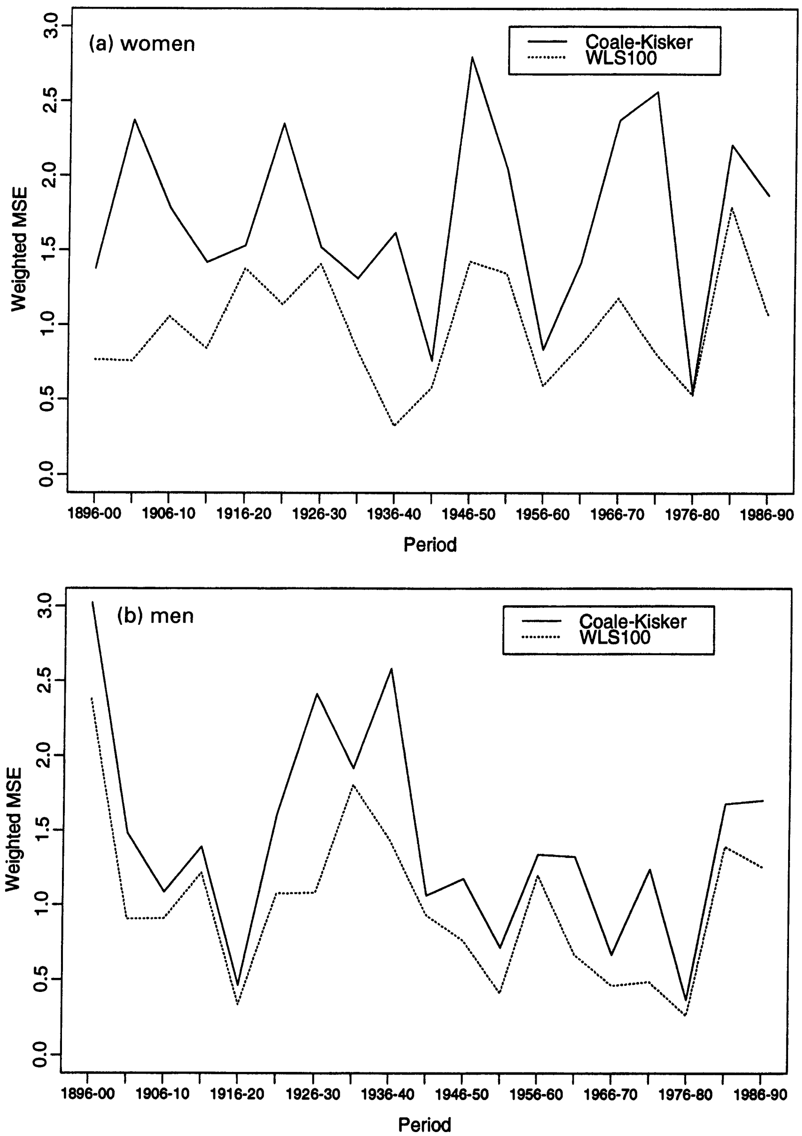

Figure 5. Goodness of fit measure (weighted mean square error for each time period) for Sweden, ages 85–100, years 1896–1900 to 1986–90. (a) women, (b) men. Comparison of two fitting procedures: (1) Coale–Kisker, (2) WLS, ages 85–100 only. Lower values indicate a closer fit.

are based fits observed mortality schedules both well and consistently. For this purpose, I have produced numerous graphical displays. Calculated mortality rates plotted against predicted values, as well as residual plots, demonstrate that the CK model provides an excellent approximation to reality, especially when the fit is obtained using weighted least squares.[16] For the countries and time periods considered in this paper, the only systematic deviations of the model from directly calculated rates occur for Japan during the 1950s; in this case, calculated mortality rates above age 100 tend to be lower than predicted by the model, no matter what fitting procedure is employed. Substantial

[16] These graphs are available from the author upon request.

differences between predicted values for the three fitting procedures considered here are quite rare below age 95, which is consistent with Coale and Kisker's assertion that the exact choice of m_{110} does not have a large effect on fitted values of m_x at these ages.

Nonetheless, it may seem curious that a single model could fit well *consistently* over such long time periods. Indeed, the overall level of mortality changed drastically during this time: between 1896–1900 and 1986–90, life expectancy at birth rose from 55 to 80 years for Swedish women and from 52 to 74 years for Swedish men; and between 1951–55 and 1986–90, life expectancy rose from 65 to 82 years for Japanese women and from 62 to 76 years for Japanese men. Nonetheless, a global measure of goodness of fit, as displayed in Figures 5 and 6, demonstrates that the model, when fitted by weighted

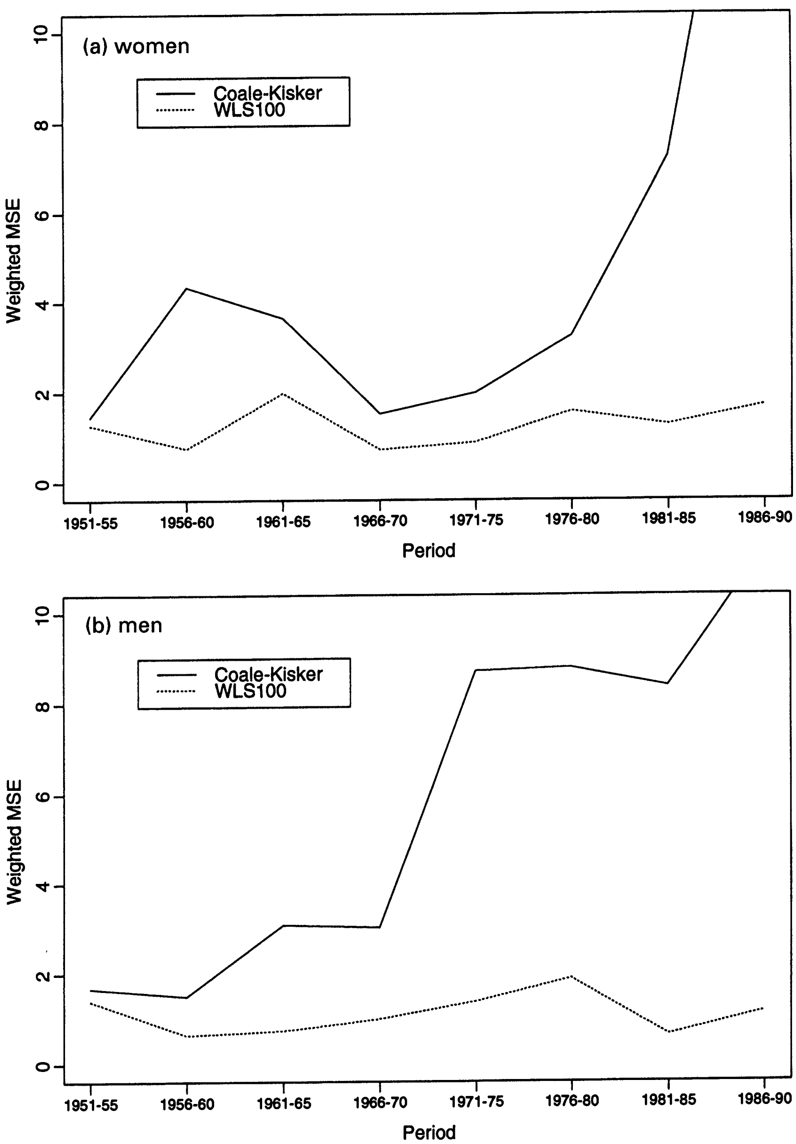

Figure 6. Goodness of fit measure (weighted mean square error for each time period) for Japan, ages 85–100, years 1951–55 to 1986–90. (a) women, (b) men. Comparison of two fitting procedures: (1) Coale–Kisker, (2) WLS, ages 85–100 only. Lower values indicate a closer fit.

least squares, captures reality with similar levels of precision over the entire observation period for men and women in both countries. These figures also demonstrate that the WLS procedure yields consistently closer fits than the original CK procedure. (Note that this measure of goodness of fit is presented only for the preferred WLS fit, which is based on data for ages 85–100 only. In this analysis, goodness of fit for a given time period is measured by the weighted mean square error of the model. Note that a closer fit is indicated by smaller values of this measure, which might more properly be termed 'badness of fit'.)[17]

CONCLUSION

In this article, I have not attempted to compare the CK model to its more well-known competitors (for example, the formulae put forward by Gompertz, Makeham, Beard, or Perks). It seems likely that either Beard's or Perks's models would fit the observed experience above age 85 about as well as the new procedures described here. Fitting procedures for these models, however, are complicated by their extreme sensitivity to random fluctuations in the raw data at very old ages, and thus a full empirical comparison is beyond the scope of this paper. An advantage of these other models is that they provide a somewhat plausible theoretical interpretation for an otherwise mysterious empirical pattern.[18] An advantage of the CK model, on the other hand, is its simplicity and ease of calculation, which would be particularly useful in simulation studies or other theoretical exercises that require large numbers of calculations. It is known that the k_x patterns of the CK and Perks's models are quite similar.[19] The similarity of the empirically based CK model to the more theoretically based model by Perks lends further justification to its use in a variety of situations. We should note, however, that the theoretical interpretations ascribed to Beard's and Perks's models apply to cohort mortality schedules, not to the period patterns analysed here. Therefore, it may be useful to duplicate the analyses presented in this article using cohort data.

In the meantime, we may summarize the conclusions of this article as follows:

(1) The form of the simple model proposed by Coale and Kisker to represent the age pattern of mortality above age 85 appears to fit both men's and women's period mortality schedules remarkably well, at least when we consider the experience of Sweden since 1895 and Japan since 1950. The form of the model has now been verified up to age 105, and it is reasonable to assume that it continues to apply up to age 110 (although no assumption is made here about what may happen at higher ages still).

[17] Let the subscript ij refer to the ith age in the jth time period. Then, the goodness-of-fit measure for period j is defined to equal $\Sigma_{i=85}^{100} \hat{D}_{ij}(\ln(m_{ij}) - \ln(\hat{m}_{ij}))^2/13$. Since the variance of $\ln(m_{ij})$ is inversely proportional to the expected number of deaths in cell (i,j), weights are chosen to equal the predicted number of deaths, $\hat{D}_{ij}$. The measure is divided by the degrees of freedom, which equals the number of observations included (16 in this case) minus the number of parameters (3). I conjecture that the expected value of this measure should be 1 in the case where the underlying mortality curve is indeed quadratic, and where there is no extra-Poisson variation in the observed death rates. In the case of Sweden, for example, the mean value of this measure across 19 time periods was 0.985 for women and 0.999 for men. For Japan, the values were higher (1.27 for women and 1.10 for men). Since the model shows no systematic deviations from the observed death rates in Japan, it is likely that the higher average values of the weighted mean square error reflect the presence of extra-Poisson variation.

[18] It is well known that Beard's formula can be derived by assuming an underlying Gompertz pattern of mortality in a heterogeneous population where individual frailty follows a gamma distribution. Thus, Beard's formula is known as the 'gamma-Gompertz' model. Similarly, Perks's formula is equivalent to a 'gamma-Makeham' model of mortality. See Horiuchi and Coale, *op. cit.* in fn. 8. See also Anatoli Yashin, James Vaupel and Ivan Iachine, 'A duality in aging: the equivalence of mortality models based on radically different concepts', *Population Studies of Aging* 5, Centre for Health and Social Policy, Odense University, 1993.

[19] Horiuchi and Coale, *op. cit.* in fn. 8.

(2) When accurate mortality data are available above age 85, it is not necessary to fit the CK model using an assumed value of m_{110}. More accurate fits can be obtained in this case by means of the weighted least squares procedure described in this paper.

(3) Based on the extrapolation technique proposed here, an analysis of Swedish and Japanese data suggests that (implied) mortality rates at extremely high ages (around 110 years) may have declined in this century for women, but probably not for men.