MODELLING SOCIO-ECONOMIC DIFFERENCES IN MORTALITY USING A NEW AFFLUENCE INDEX

BY

ANDREW J.G. CAIRNS, MALENE KALLESTRUP-LAMB, CARSTEN ROSENSKJOLD, DAVID BLAKE AND KEVIN DOWD

ABSTRACT

We introduce a new modelling framework to explain socio-economic differences in mortality in terms of an affluence index that combines information on individual wealth and income. The model is illustrated using data on older Danish males over the period 1985–2012 reported in the Statistics Denmark national register database. The model fits the historical mortality data well, captures their key features, generates smoothed death rates that allow us to work with a larger number of sub-groups than has previously been considered feasible, and has plausible projection properties.

KEYWORDS

Affluence index, Danish mortality data, CBD-X model, gravity model, multi-population mortality modelling, forward correlation term structure.

JEL codes: J11, C53, G22.

1. INTRODUCTION

There has been increasing interest in recent years in explaining and projecting mortality in closely related populations, such as those of neighbouring countries or sub-groups of a national population for which there is sufficient heterogeneity in the mortality experience of the sub-groups to justify modelling them separately. Multi-population stochastic mortality models have been developed which help us understand historical differences in the mortality of different populations and ensure that we obtain plausible projections of mortality improvements. Such projections will satisfy the following five criteria:

- Generate a projected trend that is consistent with the historical trend in mortality improvements.

Astin Bulletin 49(3), 555-590. doi:10.1017/asb.2019.14 © 2019 by Astin Bulletin. All rights reserved.

https://doi.org/10.1017/asb.2019.14 Published online by Cambridge University Press

- Produce levels of variability consistent with that observed in the historical data.
- Preserve sub-group rankings, where these seem justified.
- Generate coherent projections across sub-groups, discussed in more detail below.
- Produce plausible forward correlations between sub-groups over different ages and future horizons (see Appendix A for a definition and analysis of forward correlations).

For discussion of these criteria, see Hyndman *et al.* (2013), Villegas *et al.* (2017), and the references therein.

A key problem that is frequently encountered is that some of the populations of interest can be quite small and hence involve significant sampling variation. In some cases, it is possible to circumvent this problem by using the mortality experience of a larger (e.g., the national) population to help inform and improve projections of the smaller (e.g., a sub-) population of interest. To illustrate, the mortality experience in an annuity provider's book of business can be modelled jointly with that of the national population. For examples, see Coughlan *et al.* (2011), Villegas *et al.* (2017), and Cairns and El Boukfaoui (2017). In other cases, we might be interested in subdividing the national population into a large number of smaller populations on the basis of certain criteria. A prominent current example is socio-economic differences in mortality experience. The challenge is to extract underlying death rates from the noisy sub-population data.

Our paper addresses this challenge. Specifically, we wish to:

- exploit information contained in a national population database of historical mortality rates to construct sub-group data that exhibit a strong and consistent degree of heterogeneity in those mortality rates (i.e., maintain a consistent ranking of mortality rates) on the basis of the selected criteria;
- develop a stochastic multi-population mortality model that fits the data well;
- use Bayesian estimation methods to dampen the impact of sampling variation in the sub-group data; and
- use the resulting multi-population mortality model to make mortality projections that link sub-group-specific mortality improvements to the national trend through a gravity effect which prevents the former deviating too far from the latter.

Our specific purpose is to model socio-economic differences in mortality. We have access to a comprehensive national database maintained by Statistics Denmark (SD) which allows us to model and project the socio-economic mortality of older Danish males using the above four-part modelling approach. This database also allows us to construct a new affluence index that can be used to divide up the Danish male population into distinct socio-economic sub-groups. We find that this index is a sufficiently strong predictor of high or low

https://doi.org/10.1017/asb.2019.14 Published online by Cambridge University Press

mortality to allow subdivision of the national population into 10 sub-groups. In addition, the use of Bayesian methods enables us to separate the population into a larger number of sub-groups than would normally be advisable (for a population the size of Denmark) using more conventional likelihood-based techniques common in mortality modelling. The results achieved using the affluence index have greater precision than those using alternative predictive variables (or covariates), such as income or wealth alone.

Our approach:

- Fits the historical data well.
- Dampens sampling variation while still preserving the main features of the data (e.g., the shape of the underlying mortality curves by age in each group, the gaps between groups at all ages and in all years, and the idiosyncrasies of the underlying time trends). In doing so, it also generates a clear ordering (i.e., avoids crossovers) of the death rates for the 10 deciles at all ages from 55 to 94 and in all years.
- Generates coherent (in the sense of Hyndman *et al.* (2013), and discussed later) projections of future death rates that are both biologically and socio-economically reasonable over all ages and time horizons considered. Biological reasonableness was introduced in Cairns *et al.* (2006) and is defined as a method of reasoning used to establish a causal association (or relationship) between two factors that is consistent with existing medical knowledge. Socio-economic reasonableness implies that someone of a given age from a specific socio-economic group has a lower probability of dying than someone of the same age and gender from a lower socio-economic group over the next and subsequent years.
- Produces a plausible forward correlation term structure between the mortality rates of different groups (see Appendix A). Knowledge of this correlation term structure is important in some financial applications where risk management strategies depend on reliable correlation estimates (see Section 6).

It is important to note that our modelling framework does not impose any specific group ordering, since neither the prior distributions nor the likelihood function contains any assumption about the ordering of the groups. It therefore follows that, while multiple crossovers of the different groups are perfectly possible in our framework, this would be a sure indication that the affluence index was a poor covariate for discriminating between the different social classes.

The approach has a number of potentially useful applications:

- It allows government agencies to generate more accurate estimates of the cost and projected increases in the cost of state pension benefits in different population segments.
- It helps corporate pension plans improve estimates of their liabilities, given information about the socio-economic mix of their plan members.

https://doi.org/10.1017/asb.2019.14 Published online by Cambridge University Press

- It allows annuity providers to price annuities more accurately on a socio-economic basis.
- It allows for the improved design of hedges for mortality-linked financial products.

In short, the use of a powerful predictive variable, such as the affluence index, allows life insurers, for example, to identify the different life expectancies (LEs) of different socio-economic groups more accurately, which, in turn, leads to more reliable risk rating of customers, improved premium setting, more accurate reserving calculations, and more effective hedging.

The outline of the paper is as follows. The next section introduces and explains the Danish national mortality data set. It also discusses the approach used to process the data using the affluence index. Following that, we set out the proposed multi-population gravity model that is used to explain both mortality by socio-economic grouping and how the parameters of the model are estimated. We then consider the model fit over a range of years and ages and confirm that our approach fits the data well, smoothing out the effects of sampling variation, while still preserving the essential characteristics of the crude sub-group mortality data (such as the trend change around 1995, the mortality and LE gaps between different groups at different ages, and the changing levels of inequality between groups over time). The section following analyses properties of projected mortality: central trends, uncertainty, and correlations between sub-groups. We finish with some applications, extensions, and conclusions.

2. Data

2.1. The SD database

The analysis in this paper makes use of a high-quality database from SD, derived from administrative records. Since every individual in Denmark is given a central personal register (CPR) number either at birth or when given residence permission in the country, we are able to uniquely identify each individual across all components of the public register system. For each individual, we have information on their date of birth, education, income, wealth, and, ultimately, their date of death. We can also identify the same information for an individual's spouse or partner, thereby enabling us to allocate income and wealth within households.

The database also allows us to identify three financial indicators for each individual and couple, all rebased to year 2000 real values: gross individual annual income, total net household income, and net household wealth. All financial measures are derived from calculations made by the tax authorities which are linked to the CPR. See Appendix B for further details.

Data are available for calendar years 1980–2012. However, as the quality of both the income and the wealth data (especially for married couples) is

https://doi.org/10.1017/asb.2019.14 Published online by Cambridge University Press

regarded as less reliable at the beginning of the period, we disregard the first four years of the available data. In addition, since our methodology for allocating individuals into sub-groups requires data for the previous calendar year, we calculate exposures and deaths data for each sub-group for the 1985–2012 period only.

One of the key objectives is to identify variables derived from information available on the SD database that have a strong predictive power in explaining the mortality of different individuals in the population. More specifically, we wish to subdivide the population at each age and in each year into a number, N, of approximately equal-sized sub-groups, with a clear and unchanging ordering between sub-groups in terms of mortality rates *at all ages*: that is, Group 1 should ideally have the highest mortality at all ages and Group N the lowest. From the outset, we aimed for a larger number of sub-groups than has typically been the case in other studies (Brønnum-Hansen and Baadsgaard (2012), for example, subdivide the Danish population into $N = 4$ income-based quartiles), settling on $N = 10$ deciles, giving us the flexibility to aggregate sub-groups should this be desirable or necessary in the modelling work.

For a subsample of the US population (over 50 times bigger than Denmark), Waldron (2013) is able to achieve statistically meaningful results using deciles based on lifetime earnings covering a more limited range of ages (63–71) than in this paper (55–94). Chetty *et al.* (2016) report similar results for the US over ages 40–76 using taxable income.

For Denmark, we show that separation into deciles also gives meaningful results, but requires greater effort to filter out or smooth the effects of sampling variation. We also found that the use of deciles revealed weaknesses in some previously considered covariates that would not be so readily apparent if we were to use, say, $N = 3$ or 4 sub-groups. Extensive experimentation helped not only to identify these weaknesses, but also led us to identify a new and, in our view, superior predictive variable, namely affluence with lockdown (i.e., fixing sub-group membership) at a certain age (67 in the case of Danish data).

In this study, we aim to identify a *single* composite index derived from personal financial data that is a strong predictor of mortality. Our reasons for choosing to focus on financial variables from the outset are twofold. First, income and wealth are well known to be strong predictors of mortality (e.g., Brønnum-Hansen and Baadsgaard, 2012; Waldron, 2013; Chetty *et al.*, 2016). Second, we seek to develop a modelling framework that could be applied to other countries and sub-populations where some measure of income and/or wealth is available.

The candidate covariates considered were income, individual wealth, household wealth, and, as discussed below, linear combinations of these.

Educational attainment has long been used as a predictive variable (see, e.g., Kitagawa and Hauser, 1968). As a measure of social status, it has recently received considerable attention in the study of the relationship between social groups and LE (for a recent example, see Olshansky *et al.*, 2012). These authors identify significant and often widening differences in LE between education

https://doi.org/10.1017/asb.2019.14 Published online by Cambridge University Press

groups. However, for most developed countries, there has been a sustained trend towards higher levels of educational attainment over many decades. A likely consequence is that the least educated group will contain an increasing concentration of the most deprived in society. Thus, we are comparing a different, shrinking, and increasingly vulnerable segment of the population each year (Bound *et al.*, 2015), making inferences about the impact of education more complex and less convincing than the affluence index that we discuss in detail below. We deliberately, therefore, avoid using educational attainment as an indicator for socio-economic sub-grouping in this study.

2.2. The affluence index: construction and allocation of individuals to sub-groups

In this section, we show how the affluence index is constructed and then how it is used to allocate individuals to sub-groups.

The affluence index, $A(i, t, x)$, for individual i, in year t at age x (at the start of the year) is defined as the individual's wealth plus K times their income in the preceding year, that is:

$$A(i, t, x) = W(i, t-1, x-1) + K \times Y(i, t-1, x-1), \tag{2.1}$$

where $W(i, t-1, x-1)$ is the (physical and financial) net wealth for individual i, at age $x-1$ in year $t-1$ and $Y(i, t-1, x-1)$ is the income of individual i, at age $x-1$ in year $t-1$. K is a constant across the whole population. We interpret K as a capitalisation factor for retirement income. We chose $K = 15$ to reflect the idea that, around the age of retirement, $15 \times$ income is a rough estimate of the present value of an individual's future retirement income. This follows because, in a low interest-rate environment, K is approximately equal to national LE at retirement age (see Figure 2). Although income and wealth are highly correlated, we found that the variables combine to create a more effective index that produces consistent results between all 10 groups across all years and ages.

Representative values for the affluence index are given in Table 1 for ages 55 and 65 in 1990 and 2010. The values shown are the boundaries between adjacent groups. From Table 1 we can see that: in real terms, there has been modest growth in affluence (0.5–1% per annum depending on age and group); and, in relative terms, there has been a slight widening of the gap between Groups 1 and 10 (0.3–0.4% per annum).

Individuals not resident in Denmark for the entirety of the previous year are excluded (a small proportion – about 5% – of the full population). The affluence index is then used to allocate all remaining individuals to specific sub-groups. Based on the value of the index, $A(i, t, x)$, each individual, i, in year t at age x is ranked from $1, 2, \ldots, n$, where n is the number of individuals at age x in year t. The rank of the individuals is given by $R(i, t, x)$. This rank is normalised by $U(i, t, x) = R(i, t, x)/(n+1)$, so it is evenly spread between 0

https://doi.org/10.1017/asb.2019.14 Published online by Cambridge University Press

TABLE 1

BOUNDARIES (IN 1000S OF DANISH KRONE) BETWEEN THE 10 AFFLUENCE GROUPS FOR AGES 55 AND 65 AND IN YEARS 1990 AND 2010; FOR EXAMPLE, INDIVIDUALS AGED 65 ON 1 JANUARY 2010 WILL BE ASSIGNED TO AFFLUENCE GROUP 3 IF THEIR AFFLUENCE INDEX FOR 2009 LIES BETWEEN 2,451,000 DKK AND 2,898,000 DKK. AMOUNTS ARE REBASED TO YEAR 2000 VALUES. 1000 DKK ≈156 USD (31 OCTOBER 2017).

	Year: 1990		Year: 2010	
Group boundary	Age 55	Age 65	Age 55	Age 65
$1 \rightarrow 2$	2324	1803	2521	2026
$2 \rightarrow 3$	3141	2184	3492	2451
$3 \rightarrow 4$	3588	2518	4056	2898
$4 \rightarrow 5$	3970	2842	4542	3354
$5 \rightarrow 6$	4361	3218	5038	3866
$6 \rightarrow 7$	4836	3655	5568	4474
$7 \rightarrow 8$	5428	4219	6245	5224
$8 \rightarrow 9$	6335	5094	7261	6317
$9 \rightarrow 10$	8139	6855	9421	8356

and 1. Then, if $U(i, t, x)$ lies between 0 and 1/10, the individual is allocated to Group 1, and so on. This procedure implies that 10% of the total population is assigned to each sub-group, based on data that are available at the start of the year. Individuals can therefore transition between affluence sub-groups over time (until an individual reaches age 67). These transitions could be modelled formally using a multi-state model (see, e.g., Kwon and Jones, 2008). We comment briefly on the potential impact of these transitions in Section 2.3, but leave a formal analysis of this for future work.

The ranking and allocation is repeated every year and age until an individual reaches their 67th year, the main state pension age for most of the period. (The state pension age was reduced from 67 in 2004 to 65, although it will increase to 67 between 2024 and 2027, with further increases after 2027 linked to increases in LE.) From age 67 onwards, each individual is assumed to remain in the same sub-group that they were allocated to at age 66. We call this procedure *lockdown*.

There are two key benefits arising from this procedure. First, we found that the use of the affluence index is a more effective covariate *across all ages* (from 55 to 94) and in all years than income or wealth on their own. Second, lockdown at age 67 produces much better results (including improved separation between the 10 sub-groups) than not locking down and allowing transitions between groups after age 67. An additional advantage of lockdown is that it restricts the potential for Group 1 to fill up after age 67 with individuals in severely declining health who have used up most of their personal savings on long-term care. Allowing for such transitions would artificially inflate Group 1 mortality at high ages and depress it in the more affluent sub-groups.

https://doi.org/10.1017/asb.2019.14 Published online by Cambridge University Press

Despite these benefits of using lockdown, the following counterpoints are worth considering:

- Figure 4 (top left plot) shows that the differences in mortality are small at older ages, suggesting that lockdown might not be necessary. However, Figure 5 (top left plot) reveals that without lockdown, it is possible for the curves to remain far apart even at very high ages – in defiance of the compensation law of mortality discussed later. In any event, our interest is in the *relative* differences between the 10 mortality curves and, in particular, preserving into higher ages the rankings that we observe at age 67: we are much less interested in the *absolute* differences between groups.
- Since affluence is a good predictor of future mortality, lockdown at 67 disregards potentially valuable predictive information on affluence at higher ages. To illustrate, individuals age 70 who have used up all personal savings could subsequently experience mortality rates closer to those of Group 1 than those of whatever group they were in at age 67. While accepting such possibilities, we provide a couple of cases where lockdown at, say, 67 is necessary. The first is where we are more interested in, say, the LE of individuals at 67 than single-age mortality rates. Lockdown avoids the problem of individuals changing groups and enables LE to be calculated, assuming the individual stays in the same group. There are many variants of this. Second, from an actuarial perspective, we might wish to price an annuity at the time of retirement. Since we have information about the individual at age 67 only, we need to know what is the typical mortality trajectory of someone with those characteristics at this age and this necessitates the use of lockdown. Further, for reserving, once in payment, an insurance company is unlikely to receive additional information about the health or affluence of an annuitant until they die – again all it will know is the annuitant's status at age 67 when the annuity was purchased.
- Lockdown may produce better separation between the sub-groups, but it might be a disadvantage when designing effective hedges for certain mortality-linked products. While it would clearly be useful to be able to recalibrate a hedge in the light of new information, as with the previous point, it depends on whether or not information on those covered by the hedge is updated after they retire.

Finally, we note that the 10 sub-group allocations were not sensitive to the choice of K, the capitalisation factor for retirement income, in the range 10–20. A more sophisticated model would make K a function of age and interest rates. However, the lack of sensitivity of the sub-group allocations to values of K across this wide range suggests that such additional sophistication might not add much – at least in the case of older Danish males – and instead might lead to an unjustified illusion of precision. As a rule of thumb, we suggest that K could be set approximately equal to the LE at the lockdown age.

https://doi.org/10.1017/asb.2019.14 Published online by Cambridge University Press

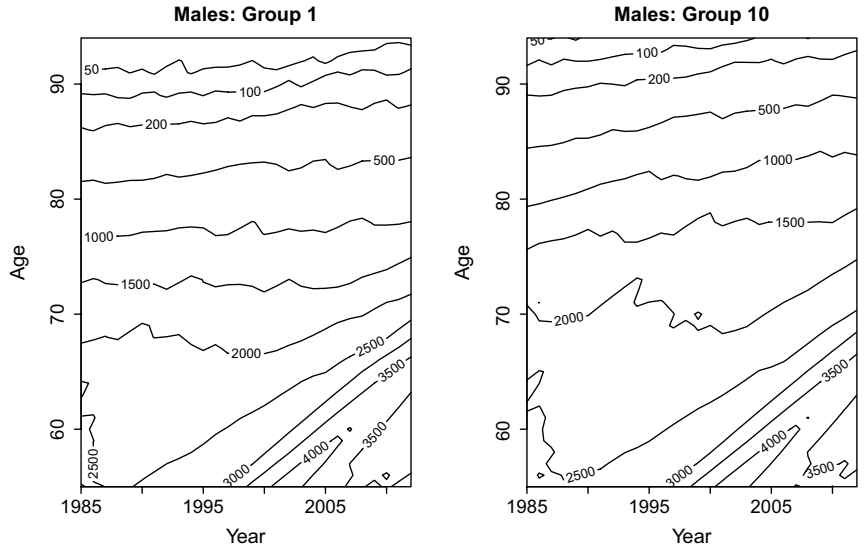


FIGURE 1: Contour plot of the exposures by year and age for Groups 1 and 10. Contours connect points that have equal values for the exposures, $E(i, t, x)$.

2.3. Deaths, exposures, and age-specific death rates

We exploit the detailed nature of the database to calculate deaths and exposures using individual records as follows:

- $D(i, t, x) =$ the number of individuals allocated to sub-group i at the start of the year, born in calendar year $t - 1 - x$ (i.e., who: (a) are age x at the start of calendar year t or (b) attain their $x + 1$th birthday during year t), who die during calendar year t.
- $E_I(i, t, x) =$ the number of individuals allocated to sub-group i, born in calendar year $t - 1 - x$, and counted on the first of January in year t; $E_I(i, t, x)$ is referred to by actuaries as the *initial exposed to risk*.
- $E(i, t, x)$ is the corresponding central exposed to risk and measures the average size of the population during year t in sub-group i born in year $t - 1 - x$. Migration of people into and out of Denmark prevents us from calculating $E(i, t, x)$ exactly, so we have chosen to use the common approximation $E(i, t, x) = E_I(i, t, x) - D(i, t, x)/2$.

The age-specific death rate is then

$$\hat{m}(i, t, x) = D(i, t, x)/E(i, t, x)$$

in sub-group i, calendar year t, and age x at the last birthday *at the beginning of the year*.

Contour plots of the exposures over the period 1985–2012 and ages 55–94 for Groups 1 and 10 are presented in Figure 1. We can see variation in cohort

https://doi.org/10.1017/asb.2019.14 Published online by Cambridge University Press

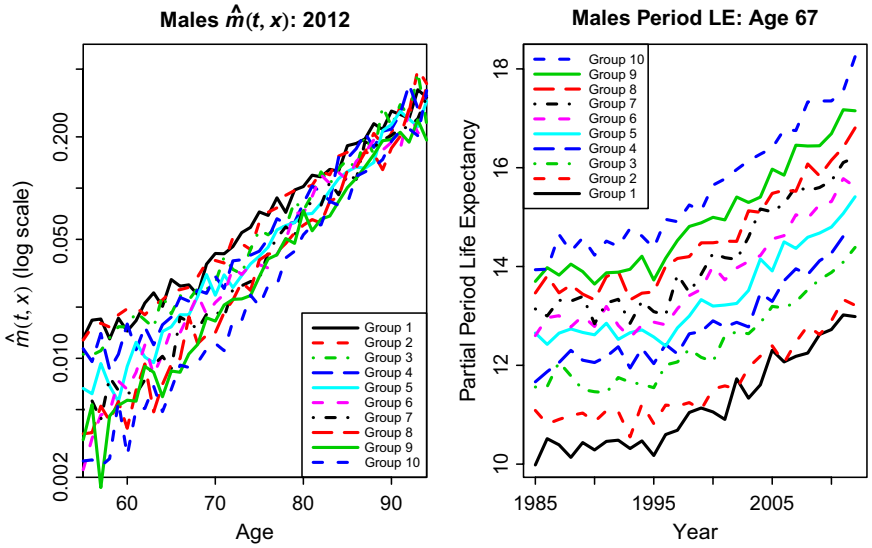


FIGURE 2: Left: age-specific death rates, $\hat{m}(i, t, x) = D(i, t, x)/E(i, t, x)$, between ages 55 and 94 for Danish males Groups 1–10 by age in 2012. Right: partial period LEs for Danish males Groups 1–10 at age 67 (capped at age 95) for years 1985–2012.

sizes over time, peaking at age 55 in 2002. By design, the two groups are approximately equal in size up to age 67. Thereafter, due to lockdown, Group 1 exposures decrease at a faster rate than Group 10: at age 94, Group 10 typically has 3–4 times the number of members remaining than Group 1.

Exposures, $E(i, t, x)$, range (across all years) from about 4250 (at age 55) down to 13 (at age 94). Deaths, $D(i, t, x)$, range from 151 (peak mortality ages) down to 4 (age 94). It is evident, therefore, that subdividing an already small national population produces crude age-specific death rates that will be subject to quite significant sampling variation.

Age-specific death rates for the 10 sub-groups in 2012 are plotted in Figure 2 (left-hand plot). This plot is typical of what we see for all years 1985–2012. Despite the significant levels of sampling variation along each curve, we can still see a clear ranking in the plot: Group 1 has consistently the highest death rates across all ages; Group 2, the next highest; and so on down to Group 10. In particular, the rankings are what could be described as socio-economically reasonable: the more affluent someone is, the less likely they are to die in the next year compared with a less affluent person of the same age. We can also observe that relative differences in death rates between sub-groups are biggest at younger ages and smallest at very high ages.

We experimented with numerous ways of segmenting the population, but always ended up with a similar (or even narrower) spread of curves at higher ages. This finding is in line with the compensation law of mortality (Gavrilov and Gavrilova, 1991) which postulates that mortality rates for different populations tend to converge with age. This 'law' implies that wealth and

https://doi.org/10.1017/asb.2019.14 Published online by Cambridge University Press

lifestyle-related factors have a reduced impact as we age, while other, mainly genetic, factors become more significant. The narrowing gap with age is consistent with the findings of, for example, Cristia (2009) for US males, and Villegas and Haberman (2014) for males in England and Wales. In our data set, the compensation law seems to point to convergence in the high 90s. For reasons discussed earlier, it is desirable to have distinct rankings over as wide a range of ages as possible. However, at high ages, it can be difficult to differentiate between groups. Still, we have not found any evidence that groups cross over even at very high ages in our data set, although the model does not preclude this.

It is also of interest to look at the variation in LEs between groups. To avoid the need to extrapolate beyond the maximum age in our data set, we calculate *partial* period LEs using a maximum age of 95, defined as the expected number of years survived from age x to age $x_u = 95$, conditional on having survived to age x, and assuming that mortality stays at the same levels as in year t. This measure can be approximated as follows:

$$LE(t,x) \approx \frac{1}{2} + \sum_{y=x+1}^{x_u-1} S_p(t,x,y) + \frac{1}{2} S_p(t,x,x_u),$$

where $S_p(t,x,y) = \exp\left[-\sum_{s=x}^{y-1} \hat{m}(t,s)\right]$ defines the observed period survival probabilities. Partial LE up to age 95 is only very slightly less than complete (untruncated) LE: proportionately the difference is less than 5% at all ages up to age 83 for all groups and years. (If LEs above age 83 are required, then it would be recommended to extrapolate the life table first before calculating full, untruncated period LEs.)

Figure 2 (right-hand plot) shows partial period LEs for the 10 groups. At age 67 (and the same holds for other ages), there is a clear separation between observed partial period LEs in all years across all 10 sub-groups (Group 1 has the lowest LE; Group 10 has the highest), even though we are using unsmoothed data. The gap between observed age-67 LEs for Groups 1 and 10 widened between 1985 and 2012 from 4.0 to 5.3 years. Nevertheless, Group 1 improved at a somewhat faster rate than Group 2 (and, to a lesser extent, Groups 3 and 4). An augmented version of the LE plot can be found in Appendix C where confidence bands accounting for sampling variation in death counts have been added. The general absence of overlap of the bands confirms that the differences between the groups and the group rankings are statistically significant.

Similar observations have been found previously with Danish data by Brønnum-Hansen and Baadsgaard (2012) using *disposable income* as a covariate. However, these authors only report period LE at a single age 0 and separate the population into quartiles. By contrast, we have carried out a much more detailed analysis using deciles rather than quartiles, used partial period LE over a range of ages from 55 upwards, and used death rates at individual

https://doi.org/10.1017/asb.2019.14 Published online by Cambridge University Press

ages from 55 to 94. Our conclusion is that the use of *disposable income* as the key metric could be improved upon significantly. Specifically, income, on its own, does not achieve a satisfactory separation of the 10 sub-groups across all ages and all years: a problem that we have resolved through the use of the $A = W + 15Y$ affluence index with lockdown at age 67.

The widening over time of the gap in LE noted above is not unique to Denmark and can be observed elsewhere. The UK Office for National Statistics (2011) considers LEs in England and Wales by occupation group and finds a widening gap between professional/managerial and unskilled manual workers; Cristia (2009) considers US males and females subdivided by lifetime earnings and also finds a widening gap over time; and Mackenbach *et al.* (2003) find increasing gaps in six European countries using other socio-economic measures.

3. Modelling framework and estimation

3.1. Modelling sub-group mortality

Let $m(i, t, x)$ be the *underlying* age-specific death rate for population i in year t for individuals aged x last birthday on the first of January in the year of death. A standard hypothesis in multi-population modelling is that the relationship between the death rates at given ages in two related populations should not diverge in the long run: that is, the ratio $m(i, t, x)/m(j, t, x)$, for $i \neq j$, should remain stable over time. This stability condition is often referred to as *coherence*, and further discussion can be found, for example, in Cairns *et al.* (2011) and Hyndman *et al.* (2013).

We propose the following multi-population gravity model of the CBD-X type for underlying age-specific death rates in the 10 sub-groups:

$$\log m(i, t, x) = \beta_0^{(i)}(x) + \kappa_1^{(i)}(t) + \kappa_2^{(i)}(t)(x - \bar{x}), \tag{3.1}$$

where i is the sub-group, t is the year, and x is the age last birthday *at the start of the year*. This is a variant in the style of Plat (2009) of the CBD model (Cairns *et al.*, 2006) that adds a non-parametric age effect, $\beta_0^{(i)}(x)$, to the basic CBD model, and models the log death rate rather than the logit of the mortality rate. The non-parametric age effect was found to be necessary to preserve the mortality rankings between sub-groups over the full range of ages from 55 to 94. The $\kappa_1^{(i)}(t)$ terms capture changes in the level of mortality, while the $\kappa_2^{(i)}(t)$ terms pick up changes in the slope of the log-mortality curve relative to the baseline $\beta_0^{(i)}(x)$. As detailed below, we model period effects by incorporating a common long-term trend for all sub-groups, combined with a short-term gravity effect.

https://doi.org/10.1017/asb.2019.14 Published online by Cambridge University Press

We specify the following stochastic model for the period effects:

$$\kappa_1^{(i)}(t) = \kappa_1^{(i)}(t-1) - \psi\left(\kappa_1^{(i)}(t-1) - \bar{\kappa}_1(t-1)\right) + \mu_1 + Z_{1i}(t), \quad (3.2)$$

$$\kappa_2^{(i)}(t) = \kappa_2^{(i)}(t-1) - \psi\left(\kappa_2^{(i)}(t-1) - \bar{\kappa}_2(t-1)\right) + \mu_2 + Z_{2i}(t), \quad (3.3)$$

$$\text{where } \bar{\kappa}_1(t) = \frac{1}{n}\sum_{i=1}^{n}\kappa_1^{(i)}(t) \text{ and } \bar{\kappa}_2(t) = \frac{1}{n}\sum_{i=1}^{n}\kappa_2^{(i)}(t). \quad (3.4)$$

$\bar{\kappa}_1(t)$ and $\bar{\kappa}_2(t)$ (Equation 3.4) can be interpreted as processes that govern mortality improvements at the national level, with additional sub-group deviations from $\bar{\kappa}_1(t)$ and $\bar{\kappa}_2(t)$ (Equations (3.2) and (3.3)) that are mean reverting to prevent sub-groups from diverging from each other. The random innovation terms $Z_{fi}(t)$ are multivariate normal with mean 0 and covariances

$$Cov\left(Z_{fi}(t), Z_{gj}(t)\right) = \begin{cases} v_{fg} & \text{for } i = j, \\ \rho v_{fg} & \text{for } i \neq j, \end{cases}$$

for $f, g = 1, 2$, where $-1 < \rho < 1$ and the $Z_{fi}(t)$ are independent from one year to the next.

The $\kappa_1^{(i)}(t)$ share a common drift, μ_1, and the $\kappa_2^{(i)}(t)$ share a common drift, μ_2, where μ_1 and μ_2 need to be estimated. The components of Equations (3.2) and (3.3)

$$-\psi\left(\kappa_1^{(i)}(t-1) - \bar{\kappa}_1(t-1)\right) \text{ and } -\psi\left(\kappa_2^{(i)}(t-1) - \bar{\kappa}_2(t-1)\right)$$

represent gravity effects between sub-groups (similar to Cairns *et al.* (2011) and Dowd *et al.* (2010)) that prevent individual sub-group death rates from drifting away from the trend. Although $0 < \psi < 2$ ensures stationarity, we restrict the range further to $0 < \psi < 1$ to avoid oscillatory behaviour.

From Equations (3.2)–(3.4), it is straightforward to show that $(\bar{\kappa}_1(t), \bar{\kappa}_2(t))$ is a bivariate random walk with drift $(\mu_1, \mu_2)'$ and one-step-ahead covariance matrix

$$\frac{1 + (n-1)\rho}{n}\begin{pmatrix} v_{11} & v_{12} \\ v_{12} & v_{22} \end{pmatrix}.$$

Now define $\Delta_{1i}(t) = \kappa_1^{(i)}(t) - \bar{\kappa}_1(t)$. The $\Delta_{1i}(t)$ are correlated AR(1) processes with AR(1) parameter $1 - \psi$, that revert and add to 0, and are independent of $\bar{\kappa}_1(t)$. Similar properties hold for the $\Delta_{2i}(t) = \kappa_2^{(i)}(t) - \bar{\kappa}_2(t)$.

We have deliberately chosen to have a single short-term contemporaneous correlation parameter, ρ, and a single gravity parameter, ψ, to keep the model simple (while still fitting the data well) and to reduce the computational burden. This parameterisation implies that the one-step-ahead correlations in log death rates between all pairs of sub-groups are equal.

https://doi.org/10.1017/asb.2019.14 Published online by Cambridge University Press

The gravity effect operates through the κ parameters which gravitate towards each other over time. However, the non-parametric age effects, $\beta_0^{(i)}(x)$, do not gravitate towards each other and it is the group-specific values of these parameters that keep the groups separate over the long run and in the same order as we have observed in the past. The key implication of this is that the mortality rates across different groups will not diverge systematically in the long run, but this does not imply that they will converge. The gaps between them can fluctuate without any requirement to converge to some constant value. Further, AR(1) processes (here, $\Delta_{1i}(t)$ and $\Delta_{2i}(t)$) do not converge to zero. Again, the underlying principle is that of coherence, but this does not preclude periods of divergence (as some studies noted above involving educational attainment have found); it just requires that in the long run there is no divergence.

3.2. Parameter estimation

The latent state variables and process parameters were estimated using Bayesian methods, providing a rigorous and coherent framework for deriving point estimates and for assessing parameter uncertainty. This approach allows us to mitigate problems that arise as a result of small population sampling variation, as recently highlighted by Villegas *et al.* (2017) and Chen *et al.* (2017).

We seek to estimate the posterior distribution for β (representing all of the $\beta^{(i)}(x)$), κ (representing all of the $\kappa_j^{(i)}(t)$) and ϕ (representing all of the process parameters governing the dynamics of the $\kappa_j^{(i)}(t)$), given the specific information, D, about deaths by sub-group, year, and age. The posterior distribution is proportional to

$$f_1(D|\beta, \kappa, \phi) f_2(\beta, \kappa|\phi) f_3(\phi),$$

where f_1 is the probability of observing $D(i, t, x)$ deaths, given β, κ and ϕ; f_2 is the density function for β and κ, given ϕ; and f_3 is the prior density for the process parameters, ϕ.

In our formulation, f_1 depends only on D, β and κ, and not on ϕ. For modelling the conditional distribution of the deaths, given κ, we use conditionally independent normal distributions for the $\log D(i, t, x)$ (conditional on κ) with mean $\log \hat{m}(i, t, x)E(i, t, x)$ and variance $1/D_{obs}(i, t, x)$, where $D_{obs}(i, t, x)$ is the observed number of deaths. This choice of distribution is discussed further in Cairns *et al.* (2016).

$f_2(\beta, \kappa|\phi)$ is based on the multivariate time series structure for the $\kappa_j^{(i)}(t)$. In our formulation, f_2 does not depend on β: that is, our prior assumption is that the $\beta_0^{(i)}(x)$ are independent of each other and have an improper uniform prior distribution; and we let the deaths data drive the estimation of the $\beta_0^{(i)}(x)$. For the given combination of Equations (3.1)–(3.3), we only need two identifiability constraints to uniquely identify the posterior density. We choose to fix

https://doi.org/10.1017/asb.2019.14 Published online by Cambridge University Press

$\bar{\kappa}_1(1) = 0$ and $\bar{\kappa}_2(1) = 0$. Beyond these constraints, we also need a distributional assumption for $\Delta(t) = \{\Delta_{ji}(t) : j = 1, \ldots, 10; i = 1, 2\}$ at time $t = 1$. We assume the stationary distribution of the underlying vector autoregressive process for $\Delta(1)$. Since $0 < \psi < 2$ and the $Z_{ki}(t)$ are multivariate normal, the stationary distribution of $\Delta(t)$ required for $\Delta(1)$ exists and is also multivariate normal.

The use of the log-normal for deaths in combination with pre-specified variances and the chosen time series model for κ results in a log-likelihood function that is quadratic in the latent state variables, $\beta_0^{(i)}(x)$, $\kappa_1^{(i)}(t)$, and $\kappa_2^{(i)}(t)$. An advantage of this specification is that when we use Markov chain Monte Carlo (MCMC) to sample from the posterior distribution for the parameters, we can use computationally efficient Gibbs sampling from the conditional posterior distributions (i.e., the multivariate normal) to update the $\beta_0^{(i)}(x)$, $\kappa_1^{(i)}(t)$ and $\kappa_2^{(i)}(t)$.

The log-likelihood does not lead to a simple conditional posterior distribution for ρ, so we use the Metropolis–Hastings (MH) algorithm to update ρ instead of the Gibbs sampler (see, e.g., Gilks *et al.*, 1996). Estimation of the posterior distribution for ψ also uses the MH algorithm for similar reasons.

The parameters v_{11}, v_{22}, and v_{12} are estimated subject to constraints. Specifically, $v_{ij} = \nu \hat{v}_{ij}$, where the prior point estimates for the $\hat{v}_{ij}$ are based on estimated values for the total Danish population, and the scalar parameter, $\nu > 0$, has to be estimated.

Uniform prior distributions are assumed throughout, except for the following:

- A Beta(2, 2) prior for ρ to ensure it remains in the range (0, 1).
- A Beta(2, 2) prior for ψ.
- An inverse gamma prior for ν, centred around 1 (shape parameter 11 and rate parameter 10); since the log-likelihood function is quadratic in relevant latent state variables, ν has an inverse gamma distribution for its conditional posterior distribution.

4. Model fitting results and validation

4.1. Fitted age-specific death rates

Point estimates for fitted age and period effects are plotted in Figure 3. The base mortality tables (left-hand plot), represented by the $\beta_0^{(i)}(x)$, provide the basis for the rankings of the different sub-groups in individual years. The $\kappa_1^{(i)}(t)$ period effects (middle plot) drive changes in the overall level of mortality. The individual series all follow a consistent downward trend (of generally improving mortality) that steepens after around 1995. Changes in the $\kappa_2^{(i)}(t)$ (right-hand plot) indicate how the slopes of the individual mortality curves have been changing. These have tended to rise over time (by varying amounts) indicating that mortality has improved at a faster rate at lower ages than at higher ages.

https://doi.org/10.1017/asb.2019.14 Published online by Cambridge University Press

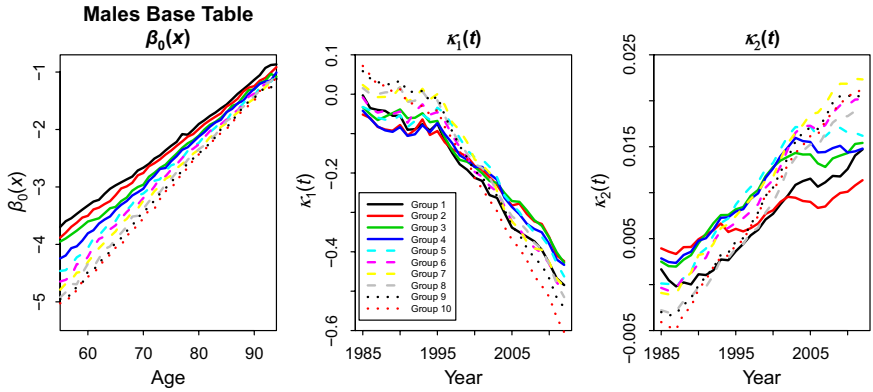


FIGURE 3: Estimated age and period effects for the 10 sub-populations.

Figure 3 also reveals that the historical pattern of the time series between sub-groups exhibits a narrowing of the gap over some periods as well as some divergence over other periods – features that are quite permissible within the structure of the gravity model. Nevertheless, the model does not allow for long-term trend differentials, despite this being a major risk concern for the insurance industry. However, it would require many more years of data than currently available to us to derive a statistically significant result supporting coherence or otherwise. While exploring alternative models is, of course, important, what we present is a first step which demonstrates that the current model provides a good fit to the data.

Estimates for the fitted age and period effects in Figure 3 are the underlying drivers of the next two figures.

Figure 4 shows fitted age-specific death rates for the years 1985 and 2012 across all ages and for ages 60 and 80 across all years. The upper right sub-plot can be compared to the observed age-specific death rates in Figure 2 (left-hand plot). The fitted rates smooth out the noise considerably and achieve a sharp distinction between each of the sub-groups. At the same time, this clarity is achieved without losing any of the main patterns and characteristics underpinning the observed rates, such as the change in trend around 1995. From these plots of smoothed fitted age-specific death rates (and others not included here), we can confirm the findings observed earlier from plots of observed rates:

- Falling death rates over time at all ages.
- A wide gap in age-specific death rates between the least and most affluent at younger ages, and a narrower gap at higher ages.

For comparison with alternative grouping methods, equivalent plots are provided in Figure 5, where the sub-groups are based on *income only* with *no lockdown* at age 67. The step change in the age-80 plot in the fitted rates around 1995 for Group 1 that is evident in the bottom right-hand plot in Figure 5 but not in Figure 4 is most likely to be the result of changes in social assistance and old-age pension provision in the previous year. After

https://doi.org/10.1017/asb.2019.14 Published online by Cambridge University Press

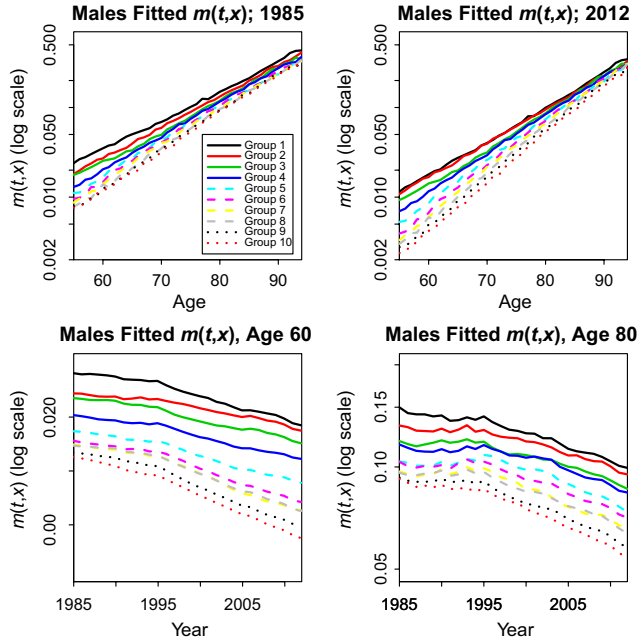


FIGURE 4: Fitted age-specific death rates, $m(i, t, x)$, for Danish males Groups 1–10 using the CBD-X model, with sub-groupings based on the affluence index (wealth+15×income) and lockdown at age 67. Top row: by age (55–94) in years 1985 and 2012. Bottom row: by year (1985–2012) for ages 60 and 80.

1995, the sub-group rankings are poor especially after retirement. Similar plots have been obtained for income only with lockdown at 67, and the affluence index (wealth + 15 × income) with no lockdown. Out of all the numerous experiments that we conducted, only affluence *with lockdown* at 67 produces a consistent ranking across all years and all ages.

Figure 6 plots the development of partial period LEs over time for ages 67 and 73 for each of Groups 1–10. The points show LEs derived from observed age-specific death rates, while the lines show LEs based on fitted age-specific death rates.

Compared with the observed LEs, the fitted LEs have more stable trends, as well as smoother gaps between the 10 groups from year to year. The model and fitting procedure achieve this without losing the essential characteristics of the observed LEs. To illustrate, the fitted LEs at both 67 and 73 (lines) track closely, but more smoothly, the level and pattern of the observed LEs (points) across all years. In particular, they pick up both the trend change in 1995 and the changing levels of inequality, such as the narrowing of the gap over time in LEs between Groups 1 and 2.

In line with earlier findings on the crude LEs, fitted LEs exhibit a wider spread at younger ages and a slight widening of spreads between sub-groups over the period 1985–2012. These plots highlight the wide gap between the rich and poor, even in a country with a strong health care and social security system.

https://doi.org/10.1017/asb.2019.14 Published online by Cambridge University Press

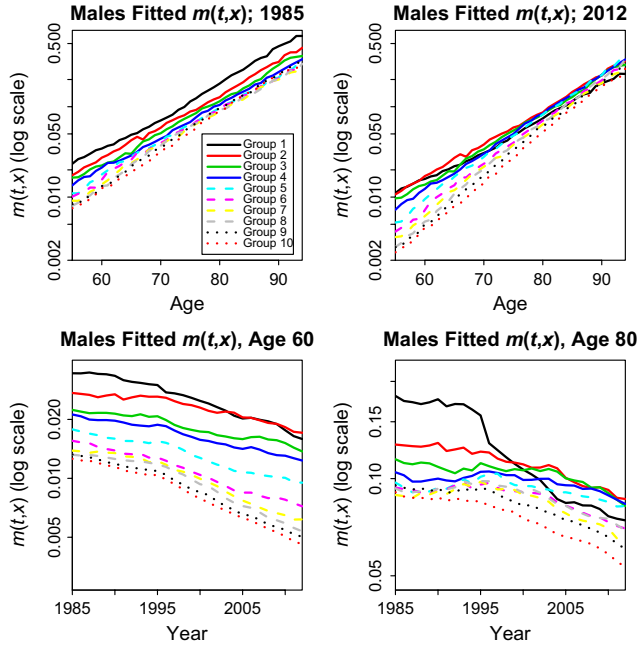


FIGURE 5: Fitted age-specific death rates, $m(i, t, x)$, for Danish males Groups 1–10 using the CBD-X model, but with sub-groupings based on income only and no lockdown at age 67. Top row: by age (55–94) in years 1985 and 2012. Bottom row: by year (1985–2012) for ages 60 and 80.

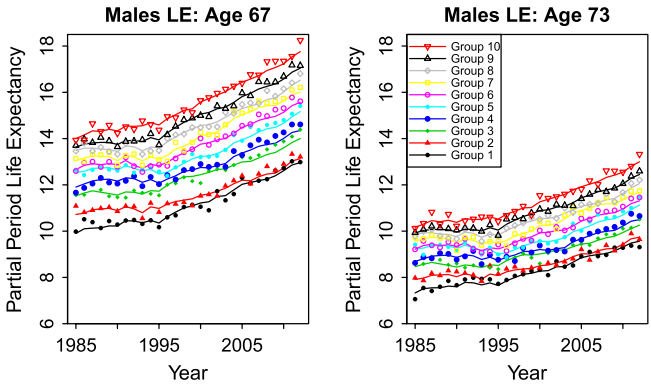


FIGURE 6: Partial period LE for Danish males in Groups 1–10 for ages 67 (left) and 73 (right). Lines: LEs based on fitted age-specific death rates. Points: LEs based on observed age-specific death rates.

4.2. Parameter uncertainty

Posterior distributions for the process parameters are presented in Figure 7. Even with 28 years of data we can see that there is significant uncertainty in the drift parameters μ_1 and μ_2: the key parameters dictating long-term improvement rates. As discussed below, uncertainties in μ_1 and μ_2 are also important

https://doi.org/10.1017/asb.2019.14 Published online by Cambridge University Press

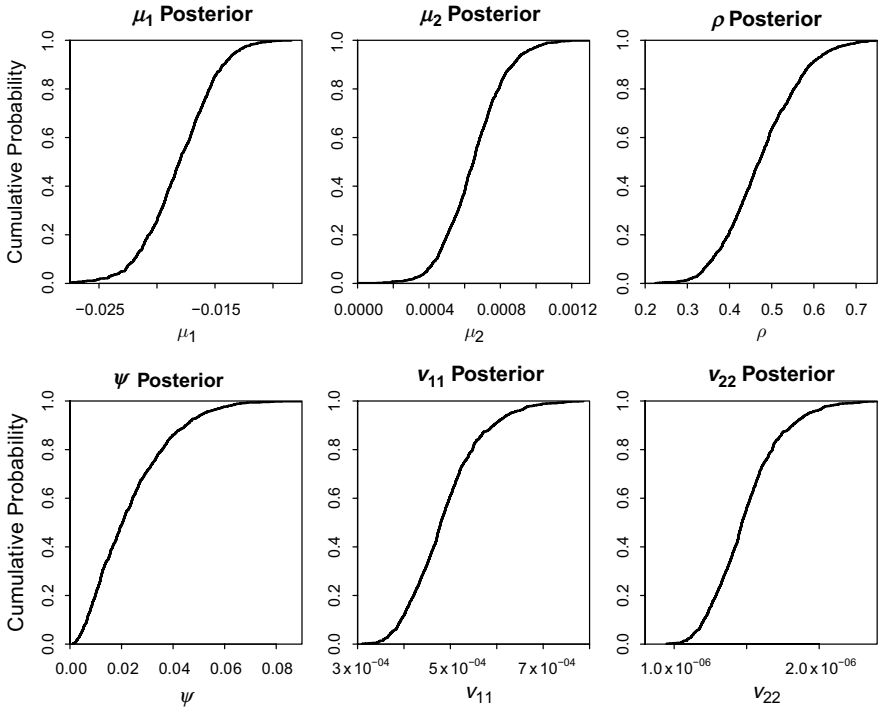


FIGURE 7: Posterior distributions for the process parameters μ_1, μ_2, ρ, ψ, v_{11}, and v_{22}.

determinants of between-group correlations because of the common dependence of future death rates on the uncertain drifts. The limited number of observations also results in relatively wide posterior distributions for ρ, v_{11}, and v_{22}. The gravity parameter, ψ, can be seen to be relatively small with, proportionately, significant variation around its median.

Although we do not plot posterior distributions for the age and historical period effects, their impact can be seen in the credibility intervals for historical death rates (i.e., up to 2012) in Figure 8. The impact of parameter uncertainty on forward correlations is discussed further in Section 5.

4.3. Model validation

The stochastic model was subjected to a number of in-sample and out-of-sample model validation tests. Standardised residuals were not found to exhibit any clustering or obvious cohort effects and had a level of variability consistent with model assumptions, indicating a good in-sample fit and providing no evidence of overfitting (which would be evident if the standardised residuals had a much lower variance than 1). Estimates of age and period effects

https://doi.org/10.1017/asb.2019.14 Published online by Cambridge University Press

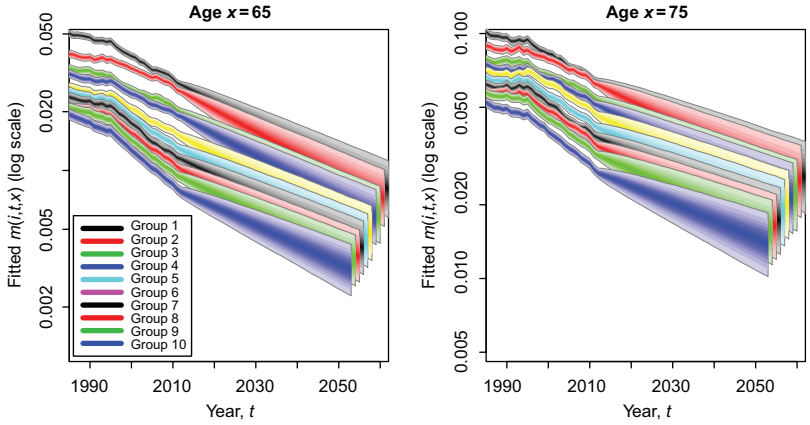


FIGURE 8: Fan charts for age-specific death rates at ages 65 and 75 for Danish males Groups 1–10. The charts show: (a) parameter uncertainty in the fitted age-specific death rates up to 2012 and (b) combined parameter uncertainty and process risk from 2013. The lower and upper edges of the fans correspond to the 5% and 95% quantiles, respectively.

were found to be robust relative to the choice of estimation period, and out-of-sample experience was found to be consistent with projections. For further details, see Appendix D.

Additionally, it is important to recognise that Danish mortality data have their own idiosyncratic features (as identified, for example, by Juel (2008) and Christensen *et al.* (2010)) and include features such as the change in trend after 1995 seen in Figure 2. These idiosyncrasies are captured well by our stochastic model across all ages, years and affluence groups.

5. PROJECTED MORTALITY

5.1. Projecting future age-specific death rates

We begin by projecting future age-specific death rates. The results in this section, unless otherwise stated, include full parameter uncertainty. This means that we draw historical values for the latent state variables (the $\beta_0^{(i)}(x)$ and the $\kappa_j^{(i)}(t)$) and the process parameters (ρ, ψ, etc.) at random from the simulated MCMC output, and then use the final year's (2012) $\kappa_j^{(i)}(t)$ plus the selected process parameters to generate each stochastic projection scenario. (For further details, see Cairns *et al.*, 2011, Appendix B).

Fan charts for the underlying historical (pre-2012) and forecast (post-2012) age-specific death rates for ages 65 and 75 are plotted in Figure 8. The fans have a reasonably constant width over the period 1985–2012, but widen after 2013, reflecting the projected uncertainty in future rates.

Forward correlations over varying time horizons are considered in Figure 9. These rise steadily the further into the future we project. Initially, the levels

https://doi.org/10.1017/asb.2019.14 Published online by Cambridge University Press

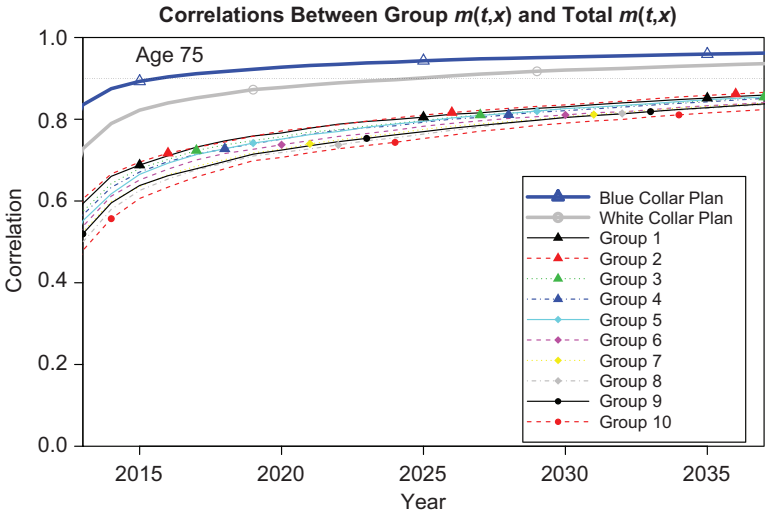


FIGURE 9: Empirical forward correlations between individual sub-groups, two pension plans, and the total population, cor $(m(i,t,x), \bar{m}(t,x))$, for age 75. Simulations incorporate full parameter uncertainty. The horizontal dotted line at a correlation of 0.9 is for reference only.

of the curves reflect the short-term contemporaneous correlations between the period effects. As the projection horizon lengthens, the shape reflects mean reversion towards the 'national' random walk (Equations (3.2) and (3.3)).

We also see that Group 10 tends to have lower forward correlations than, for example, Group 5, which, in turn, has lower forward correlations than Group 1. The reason is simply that Group 10 death rates are lower than Group 1 and so, in relative terms, contribute less than 10% of the risk to the national average. To illustrate, suppose R_1 and R_2 are identically distributed risks, and $R = pR_1 + (1-p)R_2$. If the weight, p, attached to R_1 is greater than 0.5, then R_1 will contribute more to the riskiness in R, and the correlation between R and R_1 will be higher than that between R and R_2.

Figure 9 also shows forward correlations between two stylised pension plans and the national population. The first (white-collar) pension plan is assumed to be made up of equal numbers of Groups 8–10, the high earners. The second (blue-collar) pension plan is made up of equal numbers of Groups 2–4. We exclude Group 1 from the analysis as it potentially includes more unemployed people. Both plans have much higher correlations with the national population than any of Groups 1–10 separately, reflecting the point that some of the idiosyncratic risk in each of the three contributing sub-groups has been pooled. We also see that the blue-collar plan has higher correlations than the white-collar plan, for the same reason that Group 1 has higher correlations than Group 10. It is noteworthy that the correlation term structure for the white-collar plan is similar to that for UK assured lives both in terms of level and shape (see the UK Continuous Mortality Investigation of Assured Lives vs. England & Wales males examined in Dowd *et al.* (2010), Figure 13).

https://doi.org/10.1017/asb.2019.14 Published online by Cambridge University Press

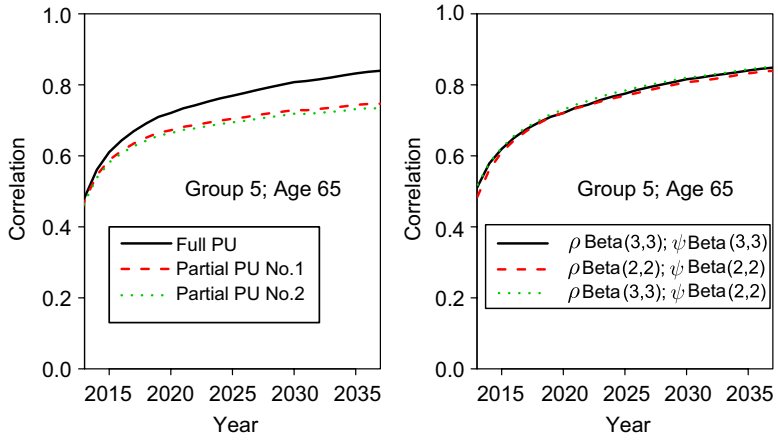


FIGURE 10: Empirical forward correlations between Group 5 and the total population, cor $(m(i, t, x), \bar{m}(t, x))$, for age 65, showing sensitivity to changes in inputs and assumptions. Left: impact of different levels of parameter uncertainty: Full parameter uncertainty (PU); Partial PU No. 1 has μ_1 and μ_2 fixed at their posterior medians; Partial PU No. 2 has μ_1, μ_2, ρ, ψ, and ν fixed at their posterior medians. Other elements of the posterior distribution remain random. Right: the three lines show sample forward correlations under three different combinations of prior distributions for ρ and ψ.

We also looked at how the ratio of death rates between different populations, $m(i, t, x)/m(j, t, x)$, evolved through time, scenario by scenario (Figure 8 does not allow this to be investigated), in order to check how strongly the principle of coherence applies in simulations. Baseline ratios between different groups and ages are defined by $\exp\left(\beta_0^{(i)}(x) - \beta_0^{(j)}(x)\right)$. The long-run simulated ratios are centred on these baseline values, typically with a standard deviation of about 0.14 (slightly wider than, but still consistent with, the observed historical ratios). Groups that were closer together historically have a greater chance of crossing over temporarily. For the same reason, small temporary crossovers are more likely at higher ages.

5.2. Sensitivity to the inclusion of parameter uncertainty

The level and shape of the forward correlation curve depend on whether we include uncertainty in the underlying process parameters (see, e.g., Cairns, 2013). We investigate this issue in Figure 10 (left-hand plot) for Group 5 by way of illustration. This plot shows forward correlations under three experiments:

- Full Parameter Uncertainty (Full PU): full allowance for uncertainty in all process parameters and latent state variables based on the posterior distribution.
- Partial PU No. 1: the drift parameters μ_1 and μ_2 are fixed at their posterior medians. All other elements of the posterior distribution remain random.
- Partial PU No. 2: the process parameters μ_1, μ_2, ρ, ψ, and ν are fixed at their posterior medians.

https://doi.org/10.1017/asb.2019.14 Published online by Cambridge University Press

From Figure 10 (left-hand plot), we see that the curves for Partial PU Nos. 1 and 2 are almost indistinguishable indicating that uncertainty in ρ, ψ, and ν has little impact on the empirical correlations. By contrast, moving from Partial PU No. 1 or Partial PU No. 2 to Full PU results in a big change. Uncertainty in μ_1 and μ_2 pushes up uncertainty in $\bar{\kappa}_1(t)$ and $\bar{\kappa}_2(t)$, with a corresponding impact on uncertainty in sub-group death rates. Since each sub-group has a common dependency on $\bar{\kappa}_1(t)$ and $\bar{\kappa}_2(t)$ (and, hence, μ_1 and μ_2), correlations rise, in line with the results in Cairns (2013).

5.3. Sensitivity to the choice of prior distribution

In Figure 10 (right-hand plot), we pick out Group 5, by way of example, and investigate the sensitivity of the correlation term structure to changes in the prior distributions for ρ and ψ. Each has either a Beta(2, 2) or Beta(3, 3) prior distribution as outlined in the legend. In each case, the three sets of priors produce similar results in each plot. Although this experiment is a limited one, it suggests that our estimates of the correlations are stable over a range of time horizons relative to the choice of prior.

5.4. Sensitivity to the gravity assumption

Appendix E outlines a modified model with different gravity effects for κ_1 and κ_2, ψ_1 and ψ_2, respectively. Although we find that estimates for ψ_1 and ψ_2 are significantly different from the single parameter ψ in the original model, we find that key outputs of interest (the distribution of future death rates and forward correlations) are largely unaltered.

6. Applications and extensions

6.1. Applications

Our analysis of mortality across a wide age range is important in the context of the intended applications. In particular, we seek to model sub-population death rates with a view to assessing how sub-groups will evolve over time relative to each other at all included ages. The resulting forward correlation term structure (as defined in Appendix A) needs to be plausible if the model is going to be useful to pension plans and insurers as a key component of their risk assessment (see, e.g., Villegas *et al.*, 2017).

A key current example is hedging an annuity provider's book of business using an index hedge, i.e., a hedging instrument that is based on the mortality experience of the national population. The difference between this and the mortality experience of the sub-population underlying the annuity provider's

https://doi.org/10.1017/asb.2019.14 Published online by Cambridge University Press

book of business introduces basis risk in the hedge and hence reduces its effectiveness. The higher the correlation between the two populations, the lower the basis risk and the greater the hedge effectiveness (see Blake *et al.*, 2018, Section 6.3). For a real-world example of an index hedge specifically designed to minimise basis risk, see the Aegon hedge designed in 2013 by Société Générale (Michaelson and Mulholland, 2015).

Another example concerns concentration and diversification of risk. Per unit of pension, a pension plan whose membership is concentrated entirely in a single affluence decile will potentially be exposed to greater levels of longevity risk than a similar-sized plan whose membership is spread more evenly across all 10 deciles. The extent of this additional risk can be measured using the approach in this paper, and assessment of strategies to mitigate this risk will require knowledge of the forward correlation term structure. Complementing this, a plausible forward correlation term structure will aid assessment of the diversification benefits from having a portfolio of lives that span a range of cohorts rather than one that is concentrated in a single cohort.

6.2. Extensions

This study has focused on older male mortality in Denmark. For females, we obtain similar results with the exception that the affluence index is less effective at producing the anticipated ranking among Groups 1–3. Group 1, in particular, seems to have rather lower mortality than Groups 2 and 3, perhaps suggesting that reported levels of income and wealth do not reflect the true level of affluence of the females in Group 1. Further work needs to be done on female mortality, although we can report that consistent rankings can be observed among Groups 4–10. In contrast, a similar study of affluence-based mortality in Norway (Kallestrup-Lamb and Rosenskjold, 2017) finds consistent rankings for females across all 10 deciles.

We have not attempted to explore the many other potential covariates and pieces of information within the SD database. Our study could be extended in a number of ways. For example, we could look at the explanatory power of including other covariates, such as marital status, occupation, and area of residence; how individuals migrate between different affluence sub-groups; and how to model mortality above age 95, in a way that exploits the convergence in death rates at high ages observed in Figure 4.

Although the model proposed here fits the data well, there are many variants and generalisations that could be considered. These include alternative one-year-ahead correlation structures; alternative gravity effects (e.g., Appendix E); and more general models for the drifts of $\bar{\kappa}_1(t)$ and $\bar{\kappa}_2(t)$ (e.g., as remarked in Appendix D.3) including allowance for trend differentials between the sub-groups.

https://doi.org/10.1017/asb.2019.14 Published online by Cambridge University Press

7. Conclusions

Understanding socio-economic differences in mortality is important both to policymakers planning state pension budgets and to private sector providers of mortality-linked products, such as pensions and annuities. We have found that socio-economic differences can be modelled more effectively through a *combination* of two elements: the choice of multi-population stochastic mortality model and the choice of predictive variable or covariate.

The multi-population stochastic mortality model we developed explains historical differences in the mortality of sub-groups of a national population and ensures that we obtain plausible projections of mortality improvements, which satisfy the following five criteria: (a) generate a projected trend that is consistent with the historical trend in mortality improvements, (b) produce levels of variability consistent with that observed in the historical data, (c) preserve sub-group rankings, (d) generate coherent projections across sub-groups, and (e) deliver plausible forward correlations between sub-groups over different ages and future horizons. We used Bayesian estimation methods to dampen the impact of sampling variation in the sub-group data – this is important for a relatively small national population such as Denmark.

We had access to a comprehensive and very high-quality data set covering the mortality experience of the Danish population. Our approach allowed us to explain mortality differences in older Danish males at a much finer level of granularity than hitherto attempted using income or education as a predictive variable – namely, at the decile level and for single ages – using just a single composite index to allocate males to a sub-group. The index, which we call the *affluence index*, is based on income and wealth data available on the SD database. The male population aged 55–94 in years 1985–2012 was subdivided into 10 sub-groups, based on the relative value of their affluence index (measured as wealth $+ 15 \times$ income). Prior to age 67, they would be allocated to a particular sub-group annually, based on the value of the index in the previous year. In their final year before reaching age 67 – the state pension age in Denmark for much of the period under investigation – they would be allocated to a sub-group and treated as remaining in that sub-group for the remainder of their lives, a procedure we call *lockdown*. Lockdown at a particular age combined with the affluence index was found to be critical for ensuring a consistent ranking of sub-group death rates across all years and all ages.

We experimented with a number of possible predictive variables: income, individual wealth, household wealth, and linear combinations of these. For most of the variants considered, results were not satisfactory across all ages and/or in all years. Typically, we would find consistent and unchanging mortality rankings across age for the higher (more wealthy) sub-groups, while lower sub-groups might be correctly ranked at some ages, but *not* at other ages. In

https://doi.org/10.1017/asb.2019.14 Published online by Cambridge University Press

most of our experiments, we found that (a) high income or high wealth is a strong predictor of low mortality, but (b) low income or low wealth is not a strong or consistent predictor of high mortality. For example, a 70 year old might have little capital, but still have a good pension that allows him to live a healthy lifestyle. Another 70 year old might have no pension, but have substantial personal savings that she can draw on to provide an adequate income in retirement – income that would not be recorded as such by SD.

Our approach fits Danish male data very well, preserving the socio-economic rankings over the historical period and without losing any of the raw data's underlying characteristics. It also generates plausible forecasts of mortality for each of the 10 sub-group and produces a plausible correlation term structure for sub-group mortality rates. In the case of Danish females, there are some crossovers in the lower socio-economic groups, but the rankings are preserved in the middle and higher groups. Our modelling framework makes no assumptions about rankings and so the model copes perfectly well if there are crossovers – although large numbers of crossovers would indicate that the predictive variable was fairly weak. To the best of our knowledge, no other investigators have been able to identify covariates that explain female mortality in the lower socio-economic groups better than our affluence index, although it remains a puzzle why females in Group 2 sometimes have worse mortality than those in Group 1. This is a clearly a subject for future research.

It is important to recognise that the focus of the paper is the affluence index and the modelling framework rather than the national peculiarities of the particular data set we have analysed. The peculiarities of the Danish data have not had any influence over the choice of affluence index or the stochastic mortality model, although it did influence the lockdown age which was set equal to the main state pension age in Denmark for most of the period. The implication of this is that with a suitable data set and with appropriate adaptation of the affluence index and lockdown conditions, the same model could be used with other countries.

Acknowledgements

The authors wish to acknowledge the numerous discussants at conference presentations of this work and especially the two referees who made many constructive and thought-provoking comments on a previous draft of the paper. MKL and CPTR gratefully acknowledge financial support from the Danish Council for Independent Research, Social Sciences, grants: 11-105548/FSE (MKL) and CREATES, Center for Research in Econometric Analysis of Time Series (DNRF78), funded by the Danish National Research Foundation. AJGC acknowledges financial support from the Actuarial Research Centre of the Institute and Faculty of Actuaries, and from Netspar under project LMVP 2012.03.

https://doi.org/10.1017/asb.2019.14 Published online by Cambridge University Press

REFERENCES

BLAKE, D., CAIRNS, A.J.G., DOWD K. and KESSLER, A. R. (2018) Still living with mortality: The longevity risk transfer market after one decade. *British Actuarial Journal*, **24**, e1, 1–80.

BÖRGER, M., FLEISCHER, D. and KUKSIN, N. (2014) Modelling the mortality trend under modern solvency regimes. *ASTIN Bulletin*, **44**, 1–38.

BOOTH, H., MAINDONALD, J. and SMITH, L. (2002) Applying Lee-Carter under conditions of variable mortality decline. *Population Studies*, **56**, 325–336.

BOUND, J., GERONIMUS, A.T., RODRIGUEZ, J.M. and WAIDMANN, T.A. (2015) Measuring recent apparent declines in longevity: The role of increasing educational attainment. *Health Affairs*, **34**, 2167–2173.

BRØNNUM-HANSEN, H. and BAADSGAARD, M. (2012) Widening social inequality in life expectancy in Denmark. A register-based study on social composition and mortality trends for the Danish population. *BMC Public Health*, **12**, 994.

CAIRNS, A.J.G. (2013) Robust hedging of longevity risk. *Journal of Risk and Insurance*, **80**, 621–648.

CAIRNS, A.J.G., BLAKE, D. and DOWD, K. (2006) A two-factor model for stochastic mortality with parameter uncertainty: Theory and calibration. *Journal of Risk and Insurance*, **73**, 687–718.

CAIRNS, A.J.G., BLAKE, D., DOWD, K., COUGHLAN, G.D., EPSTEIN, D., ONG, A. and BALEVICH, I. (2009) A quantitative comparison of stochastic mortality models using data from England & Wales and the United States. *North American Actuarial Journal*, **13**, 1–35.

CAIRNS, A.J.G., BLAKE, D., DOWD, K., COUGHLAN, G.D. and KHALAF-ALLAH, M. (2011) Bayesian stochastic mortality modelling for two populations. *ASTIN Bulletin*, **41**, 29–59.

CAIRNS, A.J.G., BLAKE, D., DOWD, K. and KESSLER, A. (2016) Phantoms never die: Living with unreliable population data. *Journal of the Royal Statistical Society, Series A*, **179**, 975–1005.

CAIRNS, A.J.G. and EL BOUKFAOUI, G. (2017) Basis risk in index based longevity hedges: A guide for longevity hedgers. *North American Actuarial Journal* (to appear).

CHEN, L., CAIRNS, A.J.G. and KLEINOW, T. (2017) Small population bias and sampling effects in stochastic mortality modelling. *European Actuarial Journal*, **7**, 193–230.

CHETTY, R., STEPNER, M., ABRAHAM, S., LIN, S., SCUDERI, B., TURNER, N., BERGERON, A. and CUTLER, D. (2016) The association between income and life expectancy in the United States, 2001-2014. *Journal of the American Medical Association*, **315**(16), 1750–1766.

CHRISTENSEN, K., DAVIDSEN, M., JUEL, K., MORTENSEN, L., RAU, R. and VAUPEL, J.W. (2010) The divergent life-expectancy trends in Denmark and Sweden – and some potential explanations. In *International Differences in Mortality at Older Ages: Dimensions and Sources* (eds. E.M. Crimmins, S.H. Preston, B. Cohen), pp. 385–407. Washington, DC: The National Academies Press.

COUGHLAN, G.D., KHALAF-ALLAH, M., YE, Y., KUMAR, S., CAIRNS, A.J.G., BLAKE, D. and DOWD, K. (2011) Longevity hedging 101: A framework for longevity basis risk analysis and hedge effectiveness. *North American Actuarial Journal*, **15**, 150–176.

CRISTIA, J.P. (2009) Rising mortality and life expectancy differentials by lifetime earnings in the United States. *Journal of Health Economics*, **28**, 984–995.

DOWD, K., CAIRNS, A.J.G., BLAKE, D., COUGHLAN, G.D. and KHALAF-ALLAH, M. (2010) Backtesting stochastic mortality models: An ex-post evaluation of multi-period-ahead density forecasts. *North American Actuarial Journal*, **14**, 281–298.

DOWD, K., CAIRNS, A.J.G., BLAKE, D., COUGHLAN, G.D., and KHALAF-ALLAH, M. (2011) A gravity model of mortality rates for two related populations. *North American Actuarial Journal*, **15**, 334–356.

GAVRILOV, L.A. AND GAVRILOVA, N.S. (1991) *The Biology of Life Span: A Quantitative Approach.* New York, NY: Harwood Academic Publisher.

GILKS, W.R., RICHARDSON, S. and SPIEGELHALTER, D.J. (1996) *Markov Chain Monte Carlo in Practice*. London: Chapman and Hall.

https://doi.org/10.1017/asb.2019.14 Published online by Cambridge University Press

HYNDMAN, R., BOOTH, H. and YASMEEN, F. (2013) Coherent mortality forecasting: The product-ratio method with functional time series models. *Demography*, **50**, 261–283.
JUEL, K. (2008) Middellevetid og dødelighed i Danmark sammenlignet med i Sverige. Hvad betyder rygning og alkohol? [Life expectancy and mortality in Denmark compared to Sweden. What is the effect of smoking and alcohol?]. *Ugeskrift Laeger*, **170**(33), 2423–2427. (In Danish).
KALLESTRUP-LAMB, M. and ROSENSKJOLD, C. (2017) Insight into the female longevity puzzle: Using register data to analyse mortality and cause of death behaviour across socio-economic groups. CREATES Research Paper 2017-08, Department of Economics and Business Economics, Aarhus University.
KITAGAWA, E.M. AND HAUSER, P.M. (1968) Education differentials in mortality by cause of death: United States, 1960. *Demography*, **5**, 318–353.
KWON, H.-S. and JONES, B. L. (2008) Applications of a multi-state risk factor/mortality model in life insurance. *Insurance: Mathematics and Economics*, **43**, 394–402.
LEE, R.D. and CARTER, L.R. (1992) Modeling and forecasting U.S. mortality. *Journal of the American Statistical Association*, **87**, 659–675.
LEE, R. and MILLER, T. (2001) Evaluating the performance of the Lee-Carter method for forecasting mortality. *Demography*, **38**, 537–549.
LI, N. and LEE, R. (2005) Coherent mortality forecasts for a group of populations: An extension of the Lee-Carter method. *Demography*, **42**, 575–594.
MACKENBACH, J.P., BOS, V., ANDERSEN, O., CARDANO, M., COSTA, G., HARDING, S., REID, A., HEMSTRÖM, Ö., VALKONEN, T. and KUNST, A.E. (2003) Widening socio-economic inequalities in mortality in six Western European countries. *International Journal of Epidemiology*, **32**, 830–837.
MICHAELSON, A. and MULHOLLAND, J. (2015) Strategy for increasing the global capacity for longevity risk transfer: Developing transactions that attract capital markets investors. *Pension and Longevity Risk Transfer for Institutional Investors*, **2015**(1), 28–37.
OFFICE FOR NATIONAL STATISTICS (2011) Trends in life expectancy by National Statistics Socio-economic Classification 1982-2006. Available at: https://www.ons.gov.uk/.
OLSHANSKY, S.J., ANTONUCCI, T., BERKMAN, L., BINSTOCK, R.H., BOERSCH-SUPAN, A., CACIOPPO, J.T., CARNES, B.A., CARSTENSEN, L.L., FRIED, L.P., GOLDMAN, D.P., JACKSON, J., KOHLI, M., ROTHER, J., ZHENG Y. and ROWE, J. (2012). Differences in life expectancy due to race and educational differences are widening, and many may not catch up. *Health Affairs*, **31**(8), 1803–1813.
PLAT, R. (2009) On stochastic mortality modelling. *Insurance: Mathematics and Economics*, **45**, 393–404.
VAN BERKUM, F., ANTONIO, K. and VELLEKOOP, M. (2016) The impact of multiple structural changes on mortality predictions. *Scandinavian Actuarial Journal*, **2016**, 581–603.
VILLEGAS, A. and HABERMAN, S. (2014) On the modeling and forecasting of socio-economic mortality differentials: An application to deprivation and mortality in England. *North American Actuarial Journal*, **18**, 168–193.
VILLEGAS, A., HABERMAN, S., KAISHEV, V. and MILLOSSOVICH, P. (2017) A comparative study of two-population models for the assessment of basis risk in longevity hedges. *ASTIN Bulletin*, **47**, 631–679.
WALDRON, H. (2013) Mortality differentials by lifetime earnings decile: Implications for evaluations of proposed social security law changes. *Social Security Bulletin*, **73**(1), 1–37.

ANDREW CAIRNS (CORRESPONDING AUTHOR)
Maxwell Institute for Mathematical Sciences,
Edinburgh EH9 3FD, UK
Department of Actuarial Mathematics and Statistics
Heriot-Watt University,
Edinburgh EH14 4AS, UK
E-mail: A.J.G.Cairns@hw.ac.uk.

https://doi.org/10.1017/asb.2019.14 Published online by Cambridge University Press

MALENE KALLESTRUP-LAMB AND CARSTEN ROSENSKJOLD
Department of Economics and Business Economics,
Center for Research in Econometric Analysis of Time Series (CREATES)
Aarhus University,
Fuglesangs Allé 4, DK-8210 Aarhus V, Denmark
E-mail: mkallestrup@econ.au.dk

DAVID BLAKE
Pensions Institute, Cass Business School,
City University of London,
London EC1Y 8TZ, UK
E-mail: D.Blake@city.ac.uk

KEVIN DOWD
Durham University Business School
Durham DH1 3LB, UK
E-mail: kevin.dowd@durham.ac.uk

APPENDIX A. FORWARD CORRELATIONS

Forward correlations can depend on a number of different random outcomes of interest. We mention two here, but others can be defined. For notational convenience assume that the last year of observation is year 0. Quantities of interest include death rates, $m(i, t, x)$, and survivor indices, $S(i, t, x) = \prod_{s=1}^{t} \exp[-m(i, s, x + s - 1)]$ (the *ex post* probability that an individual aged x at time 0 survives to time t). We define the correlation functions:

- $\rho_m(i, t, x; j, s, y) = \text{cor}(m(i, t, x), m(j, s, y))$,
- $\rho_S(i, t, x; j, s, y) = \text{cor}(S(i, t, x), S(j, s, y))$.

For both, we are likely mainly to be interested in the case $t = s$, and we focus our brief comments on this case.

A *plausible* forward correlation term structure will satisfy the following desirable criteria (building on Villegas *et al.* (2017)):

FC1: $\rho_m(i, t, x; j, s, y)$ should only be equal to 1 in the case $(i, t, x) = (j, s, y)$ (non-trivial correlations).

FC2: For a fixed age x, and time t, $\rho_m(i, t, x; i, t, y)$ should be close to 1 for y close to x, and fall away as y gets further away from x.

FC3: For $i \neq j$, $\rho_m(i, t, x; j, t, y) < \rho_m(i, t, x; i, t, y)$ (i.e., correlations between different ages in the same population are higher than equivalent correlations between different populations).

Having a reliable and plausible forward correlation term structure is important for a number of actuarial applications: for example, in assessing the diversification benefits to an insurer of holding distinctly different portfolios of lives, and in assessing the effectiveness of an index-based longevity hedge on an annuity portfolio.

A popular multi-population model that fails some of the above criteria is that of Li and Lee (2005) (see Villegas *et al.*, 2017) in which $\log m(i, t, x) = A(x)K(t) + \beta(i, x)\kappa(i, t)$, where

https://doi.org/10.1017/asb.2019.14 Published online by Cambridge University Press

$A(x)K(t)$ is an age-period component that is common to all populations, and $\beta(i, x)\kappa(i, t)$ is an age-period component that is specific to population i. For two populations i and j, for any ages x and y, where $\beta(i, x) = \beta(j, y) = 0$, we will have $\rho_m(i, t, x; j, t, y) = 1$. So the model fails criteria 1 and 2 and, almost certainly, criterion 3.

APPENDIX B. DEFINITIONS OF FINANCIAL VARIABLES USED IN THE STUDY

Gross annual income includes all taxable income, such as wage income, self-employment income, unemployment insurance benefits, social assistance (from 1994), honoraria, and all types of pension-related incomes. Note that non-taxable incomes (e.g., contributions to labour market pension schemes and private pension schemes) are not included in the gross annual income, as the taxes are deferred until the deccumulation phase.

For retired individuals, there is a break in the gross annual income variable from 1994 onwards. There are several reasons for this break. The level of old age pension benefits was increased in 1994 to compensate for the government removal of a special tax rebate previously given to retirees. Moreover, individuals living in retirement homes prior to 1994 were only given a monthly allowance, but no amount to cover rent, heating, electricity, and so on, as these expenses were paid for by the government. After 1994, all individuals, irrespective of their living situation, were given the full amount of old age pension benefits or disability pension benefits.

Household net wealth is calculated as total assets minus total liabilities at year end. Total assets include real estate, bank deposits, bonds, stocks, mortgage deeds, shares in firms, the value of cars, boats and mobile homes, as well as cash. Liabilities include credit secured by a mortgage on real estate, debt to financial institutions, credit card debt, and all other types of debt to private companies and the government (e.g., unpaid tax). Prior to 1983, all assets were assigned to the male spouse. Thereafter, 50% is allocated to each spouse, unless otherwise reported to the tax authorities. From 1987 to 1996, net wealth also included the taxable equity value of self-employed businesses (which can be negative). In 1997, a wealth tax was abolished, which resulted in the equity value of self-employed companies, the value of some mortgage deeds, shares in firms, and the value of cars, boats, and mobile homes no longer being reported in the wealth measure.

APPENDIX C. POISSON UNCERTAINTY IN EMPIRICAL LIFE EXPECTANCIES

Figure C1 (an augmented version of the right-hand plot of Figure 2) shows empirical partial period LEs with 90% confidence bands added. Confidence bands reflect uncertainty in point estimates of the LE based on the conditional Poisson model for deaths. It can be seen that the confidence intervals are consistent, relatively narrow (close to 0.25 years), and generally do not overlap. Graphically, the inclusion of confidence intervals supports the hypothesis that the groups are ranked from 1 to 10 in terms of low to high LE.

https://doi.org/10.1017/asb.2019.14 Published online by Cambridge University Press

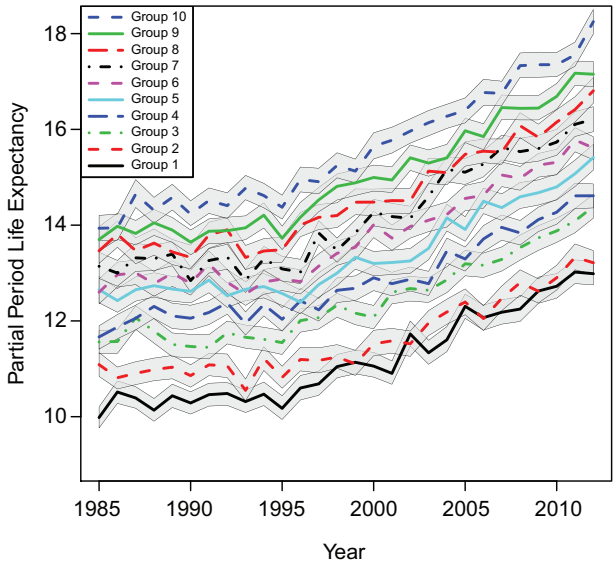


FIGURE C1: Partial period LEs for Groups 1–10 from 1985 to 2012 with 90% confidence bands (grey fans).

APPENDIX D. MODEL VALIDATION

D.1. Standardised residuals

Standardised (Pearson) residuals based on the Poisson model for deaths are defined as

$$\epsilon(i,t,x) = \left(\hat{m}(i,t,x) - m(i,t,x)\right) / \sqrt{m(i,t,x)/E(i,t,x)}.$$

If the model fits well, then the $\epsilon(i,t,x)$ should be approximately i.i.d. standard normal. Heat plots of these (see, e.g., Cairns *et al.*, 2009) for the 10 affluence groups are provided in Figure D1 using the posterior mean for the $m(i,t,x)$. As a graphical diagnostic, each of these plots looks almost completely random, with no systematic clustering, and hence consistent with the independence criterion. The $\epsilon(i,t,x)$ have mean -0.07 and variance 1.004, which is consistent with standard normality, and a QQ plot of the residuals against a standard normal is reasonably linear with only marginal evidence of a fat tail (kurtosis = 3.14). There is also no evidence of overfitting (which would be evident if the standardised residuals had a much lower variance than 1).

These results suggest that the fit is much better than can typically be achieved for some other populations or models (see, e.g., Cairns *et al.*, 2009, Section 6.1.2). This finding allows us to conclude that the model provides a satisfactory in-sample fit over the period 1985–2012.

D.2. Robustness

We tested for the robustness of parameter estimates relative to the range of years, following Cairns *et al.* (2009). Parameter estimates using data from three periods: 1985–2012;

https://doi.org/10.1017/asb.2019.14 Published online by Cambridge University Press

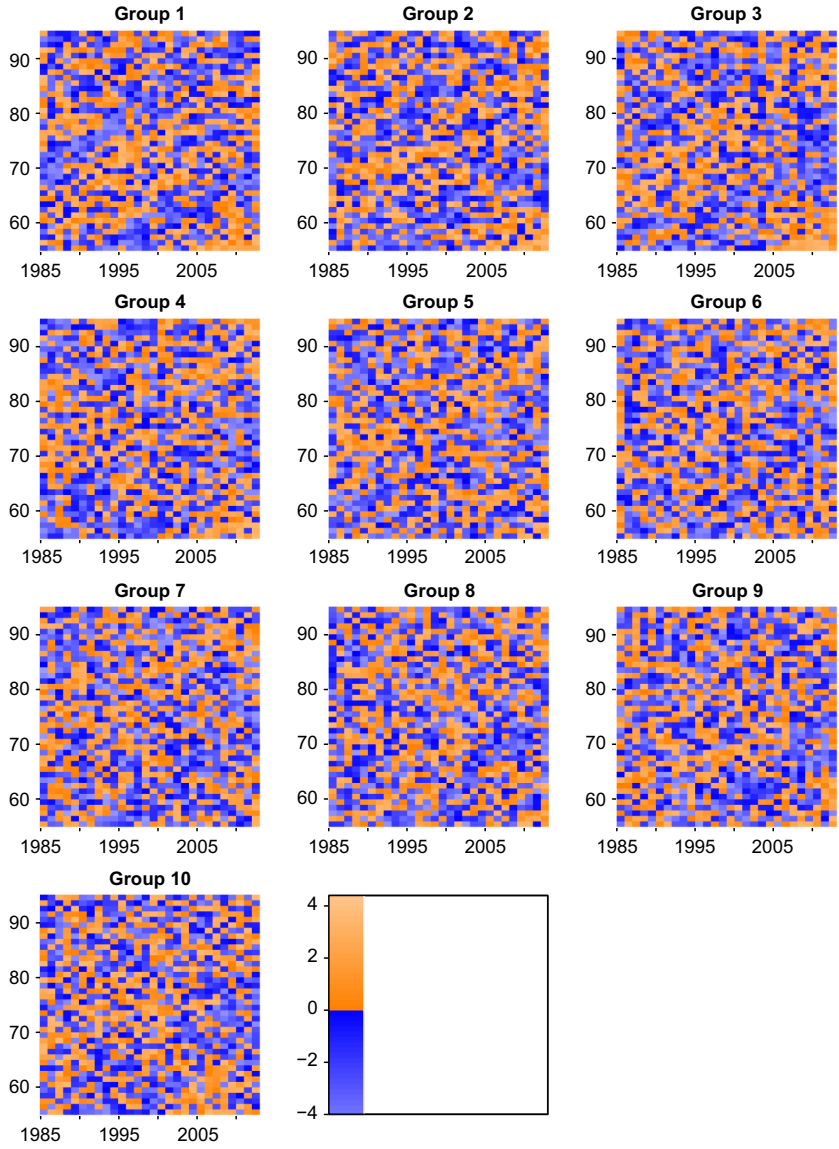


FIGURE D1: Plots of standardised (Pearson) residuals, $\epsilon(i, t, x)$, for affluence groups 1–10. Deep orange: $\epsilon > 0$, close to 0. Light orange: $\epsilon > 0$, further from 0. Deep blue: $\epsilon < 0$, close to 0. Light blue: $\epsilon < 0$, further from 0.

1995–2012; and 1995–2004. By way of example, results for Group 5 are plotted in Figure D2. We see that the broad shape of each of $\beta_0^{(5)}(x)$, $\kappa_1^{(5)}(t)$, and $\kappa_2^{(5)}(t)$ is largely unchanged in the years where estimates overlap. The approximate parallel shifts in $\kappa_1^{(5)}(t)$ and $\kappa_2^{(5)}(t)$ are the artificial result of the constraint that $\bar{\kappa}_1(t) = \bar{\kappa}_2(t) = 0$ in the first year of observation (1985 or

https://doi.org/10.1017/asb.2019.14 Published online by Cambridge University Press

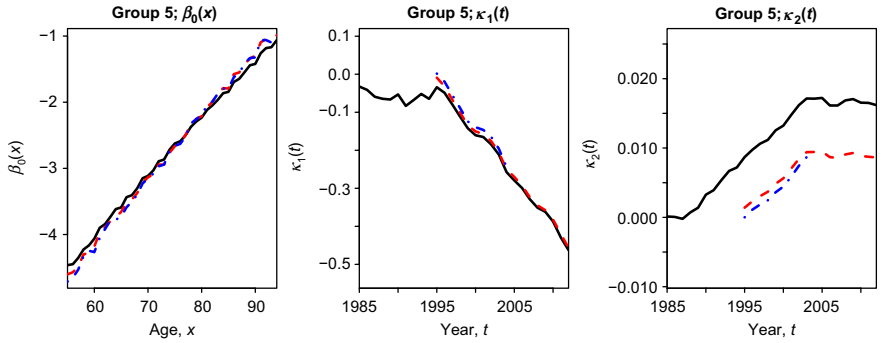


FIGURE D2: Point estimates of age and period effects for Group 5. Left: $\beta_0^{(5)}(x)$. Middle: $\kappa_1^{(5)}(t)$. Right: $\kappa_2^{(5)}(t)$. Solid lines: values based on 1985–2012 calibration. Dashed lines: 1995–2012. Dot-dashed lines: 1995–2004.

1995). The same constraints result in a small shift (from the $\bar{\kappa}_1(t)$ constraint) and a tilt (from the $\bar{\kappa}_2(t)$ constraint) in the $\beta_0^{(5)}(x)$ curves. Based on this limited test, there is no evidence to suggest that the model fit is not robust.

D.3. Optimal historical calibration period

For out of sample backtesting, we need, first, to choose an historical period (not including the most recent years) to which the model is calibrated. We then simulate from the year after the end of the calibration period up to the most recent year that has been actually observed and compare these with the forecasts. To choose the best calibration period, we follow the approach of Booth *et al.* (2002). We focused on the posterior means of the fitted $\bar{\kappa}_1(t)$, this being the main driver of uncertainty in the level of mortality in all groups. Over the full period 1985–2012, we observe a change in trend around 1995 (see Figure 3 for the individual $\kappa_1^{(i)}(t)$). Booth *et al.* (2002) remark that trend changes can and do happen (e.g., in Australia) and they propose a method for choosing the optimal start date for calibration. We found the optimal start date to be 1995.

Alternative approaches that are less sensitive to the choice of start date have been proposed that generalise the random walk to allow for structural changes or jumps in the random-walk drift: see, for example, Börger *et al.* (2014) and van Berkum *et al.* (2016).

D.4. Out-of-sample backtesting

Following Lee and Miller (2001) and Dowd (2011), we compare out-of-sample projections of partial period LE from age 67 based on a model calibration using data from 1995 to 2004. Results are presented in Figure D3. Grey fans show the out-of-sample projections of the LE with no allowance for sampling variation. Red fans show the outcome for LEs using simulations of the underlying death rates based on draws from the posterior distribution. From a graphical point of view, the actual outcomes (narrower red fans) for the 10 groups look reasonable with some groups above and some below their central projection. However, the plots exhibit relatively high LEs for Groups 9 and 10. While this is certainly compatible with the current model structure, it might indicate evidence of an alternative model structure, for example, with a different correlation pattern or with differential sub-group trends. We leave this for further work.

https://doi.org/10.1017/asb.2019.14 Published online by Cambridge University Press

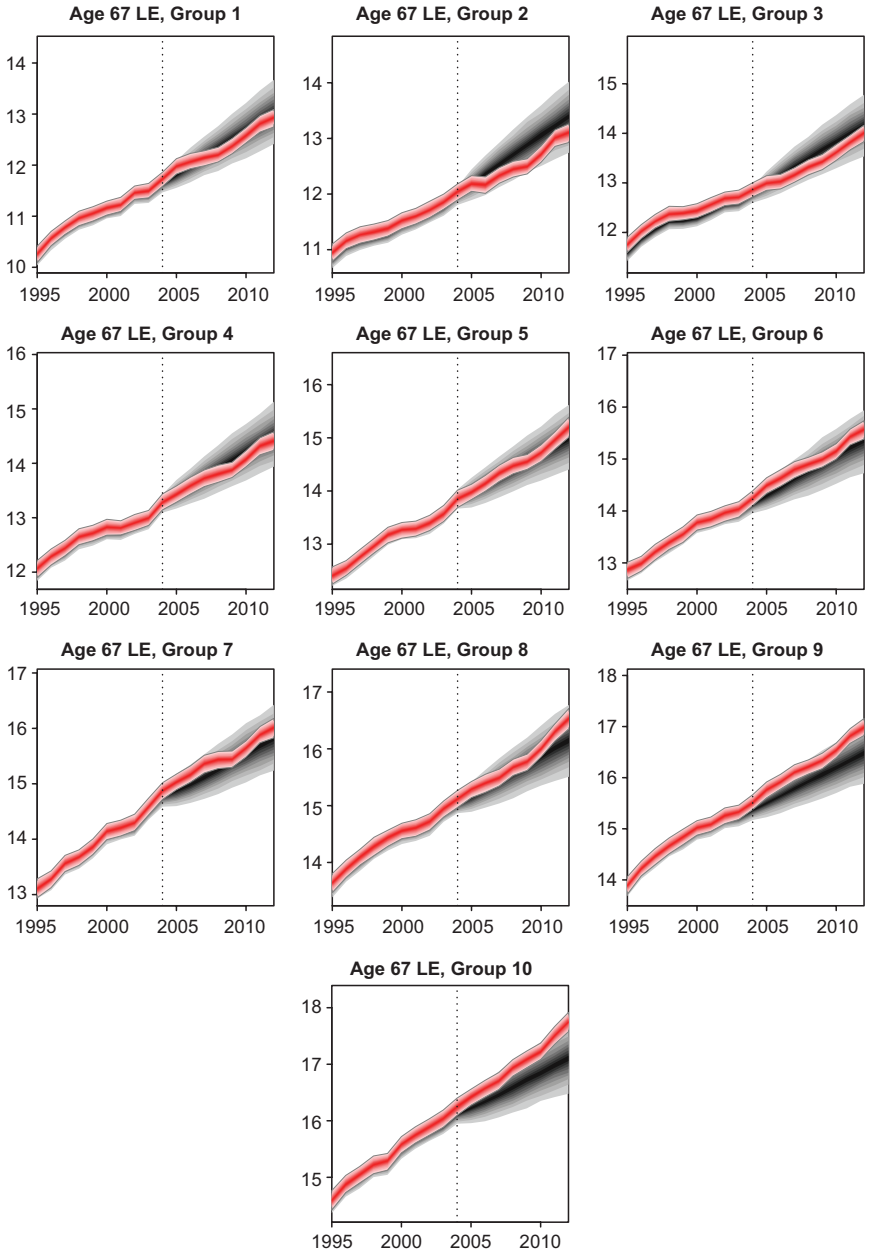


FIGURE D3: Fan charts of historical and projected partial period LE from age 67 for each of affluence Groups 1–10. Grey fan: credibility intervals (5–95% quantiles; here simulated prediction intervals with allowance for Bayesian parameter uncertainty) for underlying LE (historical and projected) using data from 1995 to 2004 and simulated mortality from 2005 to 2012. Red fan: credibility intervals using data from 1995 to 2012.

https://doi.org/10.1017/asb.2019.14 Published online by Cambridge University Press

APPENDIX E. A MORE GENERAL MODEL

A variant of the original model outlined in Equations (3.2) and (3.3) that we also considered was

$$\kappa_1^{(i)}(t) = \kappa_1^{(i)}(t-1) - \psi_1 \left(\kappa_1^{(i)}(t-1) - \bar{\kappa}_1(t-1)\right) + \mu_1 + Z_{1i}(t), \tag{E1}$$

$$\kappa_2^{(i)}(t) = \kappa_2^{(i)}(t-1) - \psi_2 \left(\kappa_2^{(i)}(t-1) - \bar{\kappa}_2(t-1)\right) + \mu_2 + Z_{2i}(t), \tag{E2}$$

where there are now two gravity parameters, ψ_1 and ψ_2, for the κ_1 and κ_2 period effects, respectively. We will refer to this new model as M2 and to the original model as M1. Although this represents a modest change from the original model, it is important to test how robust forward correlations are to variations in the model.

Outputs for M2 versus M1 are summarised briefly in Figures E1 and E2. The only significant difference between the models is that ψ_1 and ψ_2 tend to be lower than ψ in model M1 (and, to a lesser extent, ρ). In spite of this, forward correlations for M2 are very similar to M1: in part, because forward correlations are heavily dependent on uncertainty in μ_1 and μ_2 as much as ψ_1 and ψ_2.

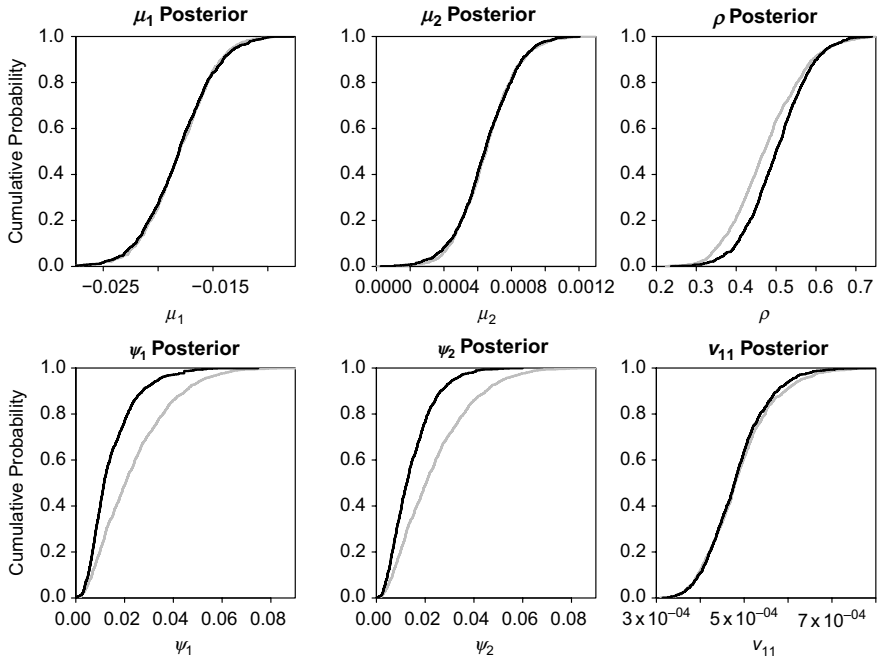


FIGURE E1: Posterior distributions for process parameters in model M2 (black) and corresponding parameters in model M1 (grey).

https://doi.org/10.1017/asb.2019.14 Published online by Cambridge University Press

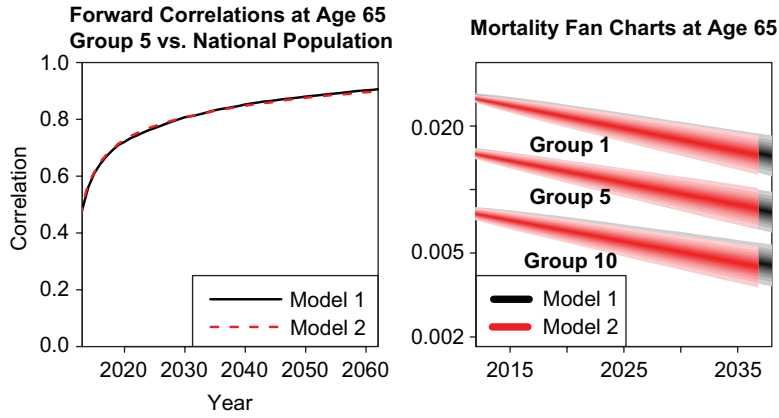


FIGURE E2: Left: forward correlations at age 65 for Group 5 versus the national population for M1 (black line) and M2 (red dashed). Right: credibility intervals for M1 (black/grey fans) and M2 (red).

https://doi.org/10.1017/asb.2019.14 Published online by Cambridge University Press