A gravity model of mortality rates for two related populations

Dowd, Kevin; Cairns, Andrew; Blake, David; Coughlan, Guy and Khalaf-Allah, Marwa

2011

Pensions Institute

DISCUSSION PAPER PI-1110

A Gravity Model of Mortality Rates for Two Related Populations

Kevin Dowd, Andrew J.G. Cairns, David Blake, Guy D. Coughlan, and Marwa Khalaf-Allah

Published in the North American Actuarial Journal, Volume 15, Number 2, 2011

ISSN 1367-580X

The Pensions Institute
Cass Business School
City University London
106 Bunhill Row
London EC1Y 8TZ
UNITED KINGDOM

http://www.pensions-institute.org/

A Gravity Model of Mortality Rates for Two Related Populations

Kevin Dowd,* Andrew J. G. Cairns†, David Blake,‡ Guy D. Coughlan§ and Marwa Khalaf-Allah‖

Abstract

The mortality rate dynamics between two related but different-sized populations are modeled consistently using a new stochastic mortality model that we call the gravity model. The larger population is modeled independently, and the smaller population is modeled in terms of spreads (or deviations) relative to the evolution of the former, but the spreads in the period and cohort effects between the larger and smaller populations depend on gravity or spread reversion parameters for the two effects. The larger the two gravity parameters, the more strongly the smaller population's mortality rates move in line with those of the larger population in the long run. This is important where it is believed that the mortality rates between related populations should not diverge over time on grounds of biological reasonableness. The model is illustrated using an extension of the Age-Period-Cohort model and mortality rate data for English and Welsh males representing a large population and the Continuous Mortality Investigation assured male lives representing a smaller related population.

1. Introduction

In most stochastic mortality models, it is implicitly assumed that the mortality rates of any one population are independent of those of other populations. Such an assumption is, however, hard to justify in the presence of common factors that would affect mortality rates across multiple populations in similar ways. Some examples might be economic progress or a cure for cancer leading to greater well-being and lower mortality across different populations, or a very bad winter or the occurrence of a pandemic leading to higher mortality. This suggests that a "biologically reasonable" mortality model should allow for the interdependence of the mortality rates of populations that are subject to common influences.

We would suggest that such a model is useful, not only because it recognizes the interdependence between related populations, but also because doing so leads to more effective risk management of mortality-related financial positions—most notably because of better estimates of basis risk and mortality correlations.

The present paper proposes such a model. The underlying assumption on which the model is based is that the state variables driving the dynamics of two related populations are attracted to each other by a "gravitational" force. In principle, we can distinguish between two different cases:

* Kevin Dowd is Visiting Professor at the Pensions Institute, Cass Business School, 106 Bunhill Row, London, EC1Y 8TZ, United Kingdom, kevin.dowd@hotmail.co.uk.

† Andrew J. G. Cairns is Professor of Financial Mathematics at the Maxwell Institute for Mathematical Sciences, and Actuarial Mathematics and Statistics, Heriot-Watt University, Edinburgh, EH14 4AS, United Kingdom, a.j.g.cairns@ma.hw.ac.uk.

‡ David Blake is Professor of Pension Economics and Director of the Pensions Institute at the Director of the Pensions Institute at the Pensions Institute, Cass Business School, 106 Bunhill Row, London, EC1Y 8TZ, United Kingdom, D.Blake@city.ac.uk.

§ Guy D. Coughlan is Managing Director of the Pension Advisory Group, J.P. Morgan, 125 London Wall, London, EC2Y 5AJ, United Kingdom, guy.coughlan@btinternet.com.

‖ Marwa Khalaf-Allah is Vice President at the Pension Advisory Group, J.P. Morgan, 125 London Wall, London, EC2Y 5AJ, United Kingdom, marwa.khalafallah@jpmorgan.com.

- *Two populations of similar size* (e.g., males and females for a given country): In this case the two populations would have (more or less) equal gravitational pull on each other. The planetary analogy here is a binary star system with two stars of more or less equal mass exerting similar gravitational pulls on each other.
- *One population much larger than the other*: In this case, the larger population exerts a pull on the smaller one, but the pull of the smaller population on the larger one is negligible. A planetary analogy here is one large object such as a star and one small object such as a planet orbiting the star. The star exerts a gravitational pull on the planet, but the planet exerts a negligible pull on the star, which for modeling purposes we can assume for convenience to be zero.[1]

This second case is arguably the most relevant to potential hedgers of mortality-dependent risks, such as pension funds or annuity providers. ''Their'' population—the one to which they are exposed—will generally be small and subject to considerable noise or other idiosyncratic factors. Typically, however, they will have the choice of hedging this exposure using an index hedge instrument predicated on some larger population, such as the national population. Thus, the large population versus small population case is key to assessing the effectiveness of (or, conversely, to quantifying the basis risk involved in) potential hedging strategies. For this reason we focus on this second case and apply it to the illustrative case of England and Wales (E&W) males versus Continuous Mortality Investigation (CMI) assured male lives estimated over a sample spanning years 1961–2005 and ages 60–84.[2]

The gravity approach suggested here has various attractions. Among these are the following:

- It gives an intuitively plausible way of thinking about the interdependence between two populations
- Given that the smaller population is much smaller than the large one, the gravity effects help to smooth out some of the imprecision in forecasts of the smaller population and
- The parameters of the model can all be estimated using maximum likelihood estimation (MLE).

The mortality model chosen is a two-population version of the Age-Period-Cohort (APC) model, alternatively known as M3B (see, e.g., Osmond 1985; Jacobsen et al. 2002; Cairns et al. 2009, Table 1). The APC model is a useful one for our purposes because it is relatively tractable and because it has a cohort effect, which evidence suggests is an important feature of the behavior exhibited in E&W mortality data and various important countries' datasets. However, the general method proposed here can be applied to most other stochastic mortality models.

The successful implementation of any two-population gravity model also requires a solution to the difficult problem of obtaining consistent estimates both of the (unobservable) state variables of the model and of the parameters governing the dynamics of those state variables. (In the case of the M3B or APC model used in this paper, the state variables are the period effects—denoted κ below—and the cohort effects—denoted γ below.) In the standard one-population context, it is relatively straightforward to obtain estimates of both the state variables and their parameters. However, in the two-population case, we also have to allow for the fact that the state variables themselves cannot be estimated without taking account of the two-population interdynamics; in other words, estimation of the state variables requires estimates of the parameters that govern them, and yet we can obtain estimates of those parameters only if we have estimates of the state variables to start with. We propose a solution to this ''chicken and egg'' problem using an iterative approach in which we start with preliminary estimates of the state variables, then estimate the parameters, then revise the estimates of the state

[1] If we believed that there was a possibility that, say, a pandemic could begin in the small population and spread to the larger population, then we would need to allow both gravity parameters to be nonzero. We might also consider a third case, namely, the pull of different country populations toward a global norm, as evidenced by the convergence of different country life expectancies over time (see, e.g., Oeppen and Vaupel 2002). We are grateful to Richard MacMinn for highlighting this case.

[2] The first population has some 186 million life-years, whereas the second has only 21 million. However, the size of the CMI ''population'' has fallen substantially over the last 15 years as demand for traditional insurance products in the United Kingdom has declined. We also emphasize that the data used were deaths and exposures data, *not* graduated q rates. The smoothing involved in the latter can lead to unreliable results, especially for the cohort effect. The E&W data are taken from the LifeMetrics website (www.lifemetrics.com) and are originally from the UK Office for National Statistics (ONS); the CMI data are available at www.actuaries.org.uk/knowledge/cmi.

variables and reestimate the parameters, and repeat this process until all parameter estimates have converged.

Although numerous studies of mortality rate modeling in single populations have been done (see, e.g., Booth and Tickle 2008; Cairns et al. 2009, 2011a; Dowd et al. 2010a, b for recent reviews of key single-population models and their properties), only a few examples are available of studies examining two-population modeling of mortality rates, for example, Li and Lee (2005), Jarner and Kryger (2009), Cairns et al. (2011b), and Li and Hardy (2011). The first study extends the Lee-Carter model by introducing the concept of a global improvement process together with mean-reverting idiosyncratic variations for each country. In the long run, the global improvement process dominates, resulting in consistent long-term projections in different countries. The second study concentrates on modeling a small population's mortality (Denmark) alongside a much larger supranational (Europe-wide) population. The large population is assumed to follow a deterministic trend over time. The relationship between the large and small population log mortality rates is modeled using a multifactor, mean-reverting spread. The mean reversion level is set at zero, reflecting the assumed similar trends in the two populations' long-run trajectories. The third study also employs a mean-reverting spread that permits different short-run trends in mortality rates, but parallel long-run improvements in line with the principles of biological reasonableness.[3] The model in that study is fitted using a Bayesian framework that allows the estimation of the unobservable state variables and the parameters of the stochastic processes driving them to be combined in a single procedure. This procedure employs Markov chain Monte Carlo techniques that both enable the uncertainty in the estimates of the historical age, period, and cohort effects to be quantified and help to smooth out the noise in parameter estimates attributable to small populations. The final study considers four types of interdependence: (1) both populations are jointly driven by the same single time-varying index, (2) the two populations are cointegrated, (3) the populations depend on a common age factor, and (4) an augmented common factor model in which a population-specific time-varying index is added to the common factor model with the property that it will tend toward a certain constant level over time.

The organization of this paper is as follows. Section 2 begins by outlining the two-population model used. Section 3 carries out preliminary empirical analysis on our E&W and CMI datasets. Section 4 examines the estimates of the latent state variables under the gravity model and compares them to the estimates obtained when both populations' mortality rates are modeled independently. Section 5 deals with the estimation of the parameters of the gravity model. Section 6 examines the gravity model's projections of the state variables and mortality (q) rates. The remaining sections consider two important applications of the gravity model: Section 7 investigates how it might be extended to examine the prices of financial instruments, and Section 8 investigates how it might be used to project forward mortality-improvement correlations, which are critical to the design of hedging strategies. Section 9 concludes.

The paper is followed by Appendices A–D, dealing with the estimation of the state variables and the parameters of the state variable processes.

2. The Gravity Two-Population Model

In this section we develop the gravity two-population version of the APC model. Before doing so, however, we analyze the single-population version of this model.

2.1. A Single-Population Model

For the single-population case, the APC model postulates that the true underlying death rate $m_{t,x}$ satisfies

$$\log m_{t,x} = \beta_x + n_a^{-1}\kappa_t + n_a^{-1}\gamma_c, \tag{1}$$

[3] A method of reasoning used to establish a causal association (or relationship) between two factors that is consistent with existing medical knowledge.

where β_x are age-dependent parameters, κ_t is a time-dependent state variable that represents the period effect, γ_c is a year-of-birth-dependent state variable that represents the cohort effect, n_a is the number of ages in the sample data used to estimate the parameters, and the variables x, t, and $c = t - x$ represent the age, current year, and year of birth, respectively. We assume κ_t follows a one-dimensional random walk with drift

$$\kappa_t = \kappa_{t-1} + \mu + CZ_t \tag{2}$$

in which μ is a constant drift term, C a constant volatility, and Z_t a one-dimensional iid $N(0,1)$ error. We follow Cairns et al. (2009) and model γ_c as an ARIMA(1,1,0) process:

$$\Delta\gamma_c = (1 - \alpha^{(\gamma)})\mu^{(\gamma)} + \alpha^{(\gamma)}\Delta\gamma_{c-1} + C^{(\gamma)}Z_c^{(\gamma)}, \tag{3}$$

where $\mu^{(\gamma)}$, $\alpha^{(\gamma)}$, and $C^{(\gamma)}$ are given parameters and $Z_c^{(\gamma)}$ is iid $N(0,1)$. The parameters of this model can be estimated using MLE.

This model requires the imposition of identifiability constraints. The first two are

$$\sum_t \kappa_t = 0 \text{ and } \sum_c \gamma_c = 0. \tag{4}$$

We also need a third identifiability constraint based on a tilting parameter δ.[4] We propose that the tilting parameter δ be chosen within an iterative scheme to minimize

$$S(\delta) = \sum_x (\beta_x + \delta(x - \bar{x}) - \bar{\beta}_x)^2,$$

where $\bar{\beta}_x = n_y^{-1} \Sigma_t \log m(t, x)$. This implies that

$$\delta = -\frac{\sum_x (x - \bar{x})(\beta_x - \bar{\beta}_x)}{\sum_x (x - \bar{x})^2}. \tag{5}$$

Given that the κ_t and γ_c already satisfy the first two constraints, we now revise our parameter estimates according to the following formulas:

$$\begin{aligned} \tilde{\kappa}_t &= \kappa_t - n_a\delta(t - \bar{t}), \\ \tilde{\gamma}_c &= \gamma_c + n_a\delta((t - \bar{t}) - (x - \bar{x})), \\ \tilde{\beta}_x &= \beta_x + \delta(x - \bar{x}). \end{aligned} \tag{6}$$

2.2 A Gravity Two-Population Model

We now wish to model death rates for two related populations, focusing on the case where one population is much larger than the other. Let us denote the large population as population 1 (''pop-1'') and the small population as population 2 (''pop-2'').

Using obvious notation, consider the pop-1 and pop-2 equivalents of equations (1)–(2):

$$\log m_{t,x}^{(1)} = \beta_x^{(1)} + n_a^{-1}\kappa_t^{(1)} + n_a^{-1}\gamma_c^{(1)}, \tag{7}$$

$$\kappa_t^{(1)} = \kappa_{t-1}^{(1)} + \mu^{(1)} + C^{(1)}Z_t^{(1)}, \tag{8}$$

and

$$\log m_{t,x}^{(2)} = \beta_x^{(2)} + n_a^{-1}\kappa_t^{(2)} + n_a^{-1}\gamma_c^{(2)}, \tag{9}$$

$$\kappa_t^{(2)} = \kappa_{t-1}^{(2)} + \mu^{(2)} + C^{(2)}Z_t^{(2)}. \tag{10}$$

---

[4] We need this constraint because we can otherwise add $n_a\delta((t - \bar{t}) - (x - \bar{x}))$ to γ_c, subtract $n_a\delta(t - \bar{t})$ from κ_t, and add $\delta(x - \bar{x})$ to β_x with no impact on fitted death rates.

The problem with applying these two models separately is that it makes no allowance for the relationship between the mortality rates of the two populations; in particular, there is nothing to stop the pop-1 and pop-2 κ_t terms drifting away from each other. Such a drift would violate the notion of biological reasonableness by which we would expect the two populations' mortality rates to be affected by common factors not explicitly modeled here (see Li and Lee 2005; Cairns et al. 2011b; Li and Hardy 2011).

We can avoid this problem by first replacing equations (8) and (10) with the following:

$$\kappa_t^{(1)} = \kappa_{t-1}^{(1)} + \mu^{(1)} + C^{(11)}Z_t^{(1)}, \tag{11}$$

$$\kappa_t^{(2)} = \kappa_{t-1}^{(2)} + \phi^{(\kappa)}(\kappa_{t-1}^{(1)} - \kappa_{t-1}^{(2)}) + \mu^{(2)} + C^{(21)}Z_t^{(1)} + C^{(22)}Z_t^{(22)}. \tag{12}$$

Note the two new terms in equation (12). The first introduces a dependency on the spread, or difference, between the kappas for the two populations, and the second introduces a stochastic factor common to $\kappa_t^{(1)}$. The first of these new terms, $\phi^{(\kappa)} \geq 0$, is a "gravity" or mean reversion parameter: If $\phi^{(\kappa)} > 0$, this parameter works to reduce any emerging spread between the 2-pop κ_t state variables back toward 0 by pulling the pop-2 κ_t state variables back toward the pop-1 κ_t state variables, which are themselves unaffected by the pop-2 κ_t state variables.

Some noteworthy features can be pointed out about this 2-pop κ_t specification:

1. The greater the gap between the κ_t state variables, the greater the "gravitational pull" on the pop-2 κ_t to force them toward the pop-1 κ_t.
2. The special case $\phi^{(\kappa)} = 0$ results in a simple bivariate random walk model with drift and correlated innovations.
3. In the alternative special case $\phi^{(\kappa)} = 1$, the deterministic part of the pop-2 κ_t process (12) reduces to $\kappa_t^{(2)} = \kappa_{t-1}^{(1)} + \mu^{(2)}$, implying that the evolution of $\kappa_t^{(2)}$ is driven off $\kappa_{t-1}^{(1)}$ and is *entirely unaffected* by $\kappa_{t-1}^{(2)}$. This would seem to be unreasonably extreme. We should therefore regard $\phi^{(\kappa)} = 1$ as an extreme upper bound. The noise processes in equations (11)–(12) have been generalized to allow for correlation between the two populations.

The parameters of the κ_t process are estimated by MLE, and details of the algorithms are provided in Appendix B.

In addition to the period effects, we also have the cohort effects, and, were these allowed to drift apart, we would again have a model that violated biological reasonableness. We therefore assume that the two populations' gamma state variables are dependent as well.

Let $\mu^{(\gamma1)}$ be the mean reversion level for $\Delta\gamma_c^{(1)}$, $\mu^{(\gamma2)}$ the mean reversion level for $\Delta\gamma_c^{(2)}$, $\alpha^{(\gamma1)}$ the AR(1) parameter for $\Delta\gamma_c^{(1)}$, and $\alpha^{(\gamma2)}$ the AR(1) parameter for $\Delta\gamma_c^{(2)}$. Now let $\phi^{(\gamma)} \geq 0$ be the gravity parameter for the 2-pop γ_c process. The 2-pop AR(1) γ_c processes can then be expressed as

$$\gamma_c^{(1)} = (1 + \alpha^{(\gamma1)})\gamma_{c-1}^{(1)} - \alpha^{(\gamma1)}\gamma_{c-2}^{(1)} + \mu^{(\gamma1)}(1 - \alpha^{(\gamma1)}) + C^{(\gamma11)}Z_c^{(\gamma1)}, \tag{13a}$$

$$\begin{aligned}\gamma_c^{(2)} = {} & (1 + \alpha^{(\gamma2)} - \phi^{(\gamma)})\gamma_{c-1}^{(2)} - \alpha^{(\gamma2)}\gamma_{c-2}^{(2)} + \phi^{(\gamma)}\gamma_{c-1}^{(1)} + \mu^{(\gamma2)}(1 - \alpha^{(\gamma2)}) \\ & + C^{(\gamma21)}Z_c^{(\gamma1)} + C^{(\gamma22)}Z_c^{(\gamma2)}.\end{aligned} \tag{13b}$$

The first new term in equation (13b) introduces a dependency on the spread between the gammas for the two populations, and the second introduces a stochastic factor common to $\gamma_c^{(1)}$. The first of these new terms $\phi^{(\gamma)} \geq 0$ is a "gravity" or mean reversion parameter: If $\phi^{(\gamma)} > 0$, this parameter works to reduce any emerging spread between the 2-pop γ_c state variables back toward 0 by pulling the pop-2 γ_c state variables back toward the pop-1 γ_c state variables, which are themselves unaffected by the pop-2 γ_c state variables.

Again, the parameters (in this case, the γ_c parameters) of pop-1 are independent of, but exert a gravitational pull on, the pop-2 γ_c parameters.

A couple of other features should also be noted about this 2-pop γ_c process:

1. For exactly the same reasons as with the κ_t process, we would regard $\phi^{(\gamma)} = 1$ as an extreme upper bound and would usually expect $\phi^{(\gamma)}$ to be toward the lower end of the range [0,1).
2. As with the 2-pop κ_t, we have generalized the noise process to allow for noise correlation between the two populations.

The parameters of the γ_c process are estimated by MLE, and details of the algorithms are provided in Appendix C.

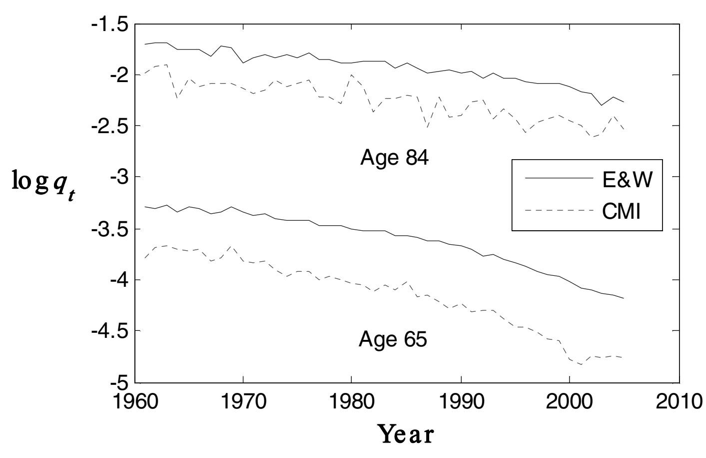
3. Preliminary Empirical Analysis

This section provides a preliminary empirical analysis of the E&W and CMI q rates and state variables. Figure 1 shows plots of the historical E&W and CMI male q rates over the period 1961–2005.[5] (The q rates are given by $q = 1 - \exp(-m)$, where m is the death rate.) We see that the CMI rates are always below the E&W rates, although the difference seems to be declining over time. We also see that age-84 rates are notably higher and more volatile than age-65 rates,[6] and the CMI rates are more volatile than the E&W rates for both ages.

Figure 2 shows the historical evolution of κ_t for single-population APC models of each population for the period 1961–2005, each estimated independently. We can see that the historical E&W and CMI follow each other fairly closely, but the CMI κ_t tends to fall a little more rapidly.

Figure 3 shows corresponding historical single-population γ_c for E&W and CMI populations against birth years. CMI historical γ_c are noticeably more volatile than E&W historical γ_c, but (except for the very early birth years and the very latest birth year) they follow each other fairly closely.

Figure 1
Historical Mortality (q_t) Rates: E&W versus CMI



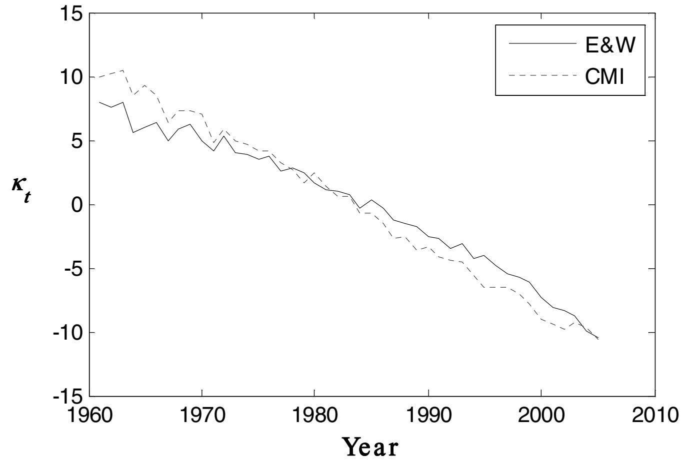
Note: Based on LifeMetrics E&W males data and CMI assured male lives data for 1961–2005.

[5] It is well known that exposures for the E&W 1886 birth cohort are unreliable for reasons that are not understood (see, for example, Cairns et al. 2009). For this analysis, we wished to include this cohort, but in a way that did not produce an outlier in the fitted cohort effect. Therefore we chose to adjust the E&W exposures, $E(t, x)$, for the 1886 cohort effect in the calendar years 1961–1970 (after 1970 the unadjusted exposures seem to be correct). This was achieved by assuming that the logarithms of the crude death rates $m(t, x)$ for the 1886 cohort were equal to the average of the log death rates for the 1885 and 1887 cohorts in the same calendar year. See Coughlan et al. (2011) for a detailed model-independent empirical analysis of the relationship between these two populations.

[6] Note, however, that the exposure sizes for the CMI at higher ages are very low and have been declining significantly.

Figure 2

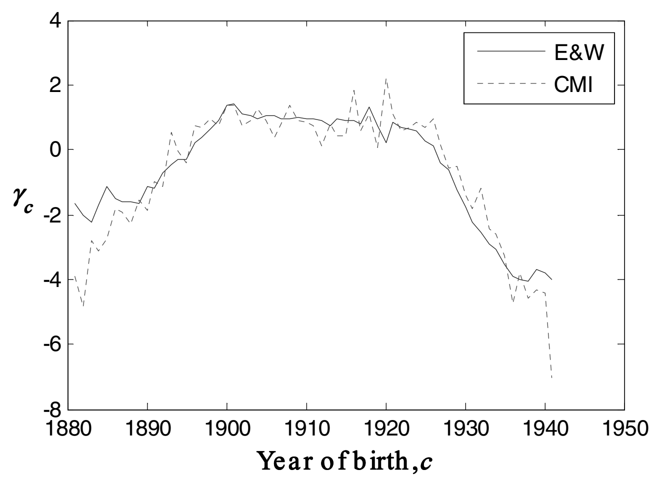
Historical κ_t State Variables Estimated Independently: E&W versus CMI



Note: Based on LifeMetrics E&W males data and CMI assured male lives data for 1961–2005 and estimated over ages 60–84.

Figure 3

Historical γ_c State Variables Estimated Independently: E&W versus CMI



Note: Based on LifeMetrics E&W males data and CMI assured male lives data for 1961–2005 and estimated over ages 60–84.

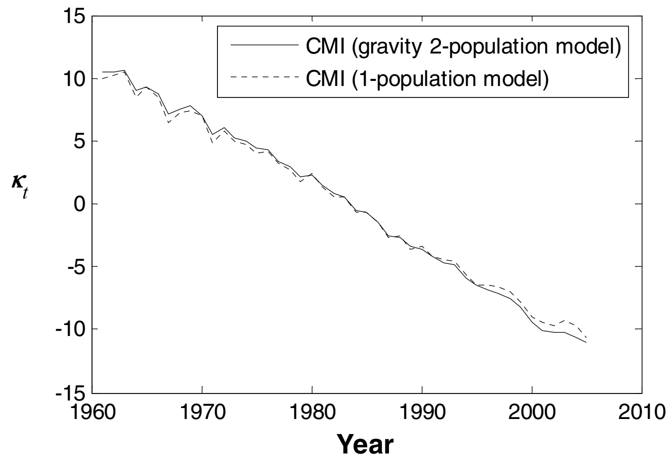
4. ESTIMATES OF THE STATE VARIABLES UNDER THE GRAVITY TWO-POPULATION MODEL

This section discusses the estimates of the state variables under the gravity two-population model, obtained using the iterative algorithm explained in Appendices A and D.

Figure 4 compares the single-population and two-population gravity model estimates of the kappa state variables. We see that there is a small difference between estimates of these state variables.

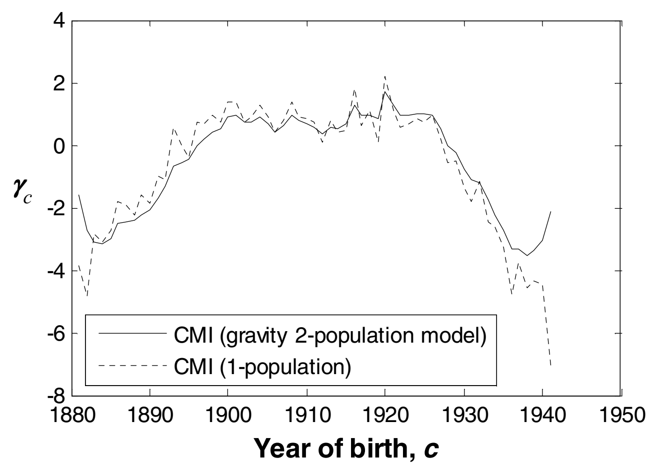
Figure 5 compares the corresponding estimates of the gamma state variables. In this case the differences are more obvious: The gravity estimates of the gammas are smoother than the independently estimated gammas, and the big drop at the end of the series (referring to year of birth 1941) is replaced with a large increase.

Figure 4
One-Population and Gravity Two-Population Estimates of the CMI κ_t State Variables



Note: Based on LifeMetrics E&W males data and CMI assured male lives data for 1961–2005 and estimated over ages 60–84.

Figure 5
One-Population and Gravity Two-Population Estimates of the CMI γ_c State Variables



Note: Based on LifeMetrics E&W males data and CMI assured male lives data for 1961–2005 and estimated over ages 60–84.

5. Estimates of the Parameters of the Gravity Two-Population Model

This section deals with the estimation of the parameters of the gravity two-population model. These parameters are estimated using a Bayesian MLE approach that is explained further in Appendices A, B, and C.

Table 1 gives estimates of the parameters of the bivariate κ_t process fitted to the E&W and CMI populations. The table shows results for both the gravity model and the single-population model, in which the mortality rates of the two populations are modeled independently of each other, and our focus of interest is on how the $\kappa_t^{(2)}$ parameters might differ between the two cases. Note that $\hat{V}^{(11)}$, $\hat{V}^{(12)}$, $\hat{V}^{(21)}$, and $\hat{V}^{(22)}$ are the elements of the estimated variance-covariance matrix for the two-population κ_t process. In the two-population case, we get $\hat{\mu}^{(2)} = -0.4900$, $\hat{V}^{(12)} = 0.2559$,

Table 1
Estimated κ_t Parameters

Parameter	Parameter Estimate
Gravity Model (E&W and CMI)	
$[\hat{\mu}^{(1)}, \hat{\mu}^{(2)}]$	$[-0.4205, -0.4900]$
$\begin{bmatrix} \hat{V}^{(11)}, \hat{V}^{(12)} \\ \hat{V}^{(21)}, \hat{V}^{(22)} \end{bmatrix}$	$\begin{bmatrix} 0.4745, 0.2559 \\ 0.2559, 0.2999 \end{bmatrix}$
$\hat{\phi}^{\kappa}$	0.0928
E&W Modeled Independently	
$\hat{\mu}^{(1)}$	−0.4205
$\hat{V}^{(11)}$	0.4745
CMI Modeled Independently	
$\hat{\mu}^{(2)}$	−0.4670
$\hat{V}^{(22)}$	0.5298

Note: Based on LifeMetrics E&W males data and CMI assured male lives data for 1961–2005 and estimated over ages 60–84.

$\hat{V}^{(22)} = 0.2999$, and $\hat{\phi}^{(\kappa)} = 0.0928$, whereas, in the one-population case, we get $\hat{\mu}^{(2)} = -0.4670$ and $\hat{V}^{(22)} = 0.5298$ (and, implicitly, $\hat{V}^{(12)} = 0$ and $\hat{\phi}^{(\kappa)} = 0$).

Table 2 gives estimates of the parameters of the γ_c process. As in Table 1, this table shows results for both the two-population and one-population models, where $\hat{V}^{(\gamma 11)}$, $\hat{V}^{(\gamma 12)}$, $\hat{V}^{(\gamma 21)}$, and $\hat{V}^{(\gamma 22)}$ are the elements of the estimated variance-covariance matrix for the two-population γ_c process. In the two-population case, we get $\hat{\alpha}^{(\gamma 2)} = 0.0626$, $\hat{\mu}^{(\gamma 2)} = 0.0066$, $\hat{V}^{(\gamma 12)} = 0.0081$, $\hat{V}^{(\gamma 22)} = 0.0858$, and $\hat{\phi}^{(\gamma)} = 0.3133$, whereas, in the single-population case, we get $\hat{\alpha}^{(\gamma 2)} = -0.3440$, $\hat{\mu}^{(\gamma 2)} = -0.0168$,

Table 2
Estimated γ_c Parameters

Parameter	Parameter Estimate
Two-Population Gravity Model (E&W and CMI)	
$[\hat{\alpha}^{(\gamma 1)}, \hat{\alpha}^{(\gamma 2)}]$	$[0.3024, 0.0626]$
$[\hat{\mu}^{(\gamma 1)}, \hat{\mu}^{(\gamma 2)}]$	$[-0.0318, 0.0066]$
$\begin{bmatrix} \hat{V}^{(\gamma 11)}, \hat{V}^{(\gamma 12)} \\ \hat{V}^{(\gamma 21)}, \hat{V}^{(\gamma 22)} \end{bmatrix}$	$\begin{bmatrix} 0.0872, 0.0081 \\ 0.0081, 0.0858 \end{bmatrix}$
$\hat{\phi}^{(\gamma)}$	0.3133
E&W Modeled Independently	
$\hat{\alpha}^{(\gamma 1)}$	0.2743
$\hat{\mu}^{(\gamma 1)}$	−0.0378
$V^{(\gamma 11)}$	0.0901
CMI Modeled Independently	
$\hat{\alpha}^{(\gamma 2)}$	−0.3440
$\hat{\mu}^{(\gamma 2)}$	−0.0168
$\hat{V}^{(\gamma 22)}$	0.5174

Note: Based on LifeMetrics E&W males data and CMI assured male lives data for 1961–2005 and estimated over ages 60–84.

$\hat{V}^{(\gamma 22)} = 0.5174$, and (again implicitly) $\hat{V}^{(\gamma 12)} = 0$ and $\hat{\phi}^{(\gamma)} = 0$. The negative value of $\hat{\alpha}^{(\gamma 2)}$ reflects the greater noise in the estimates of the CMI cohort effects (see Figure 3).[7]

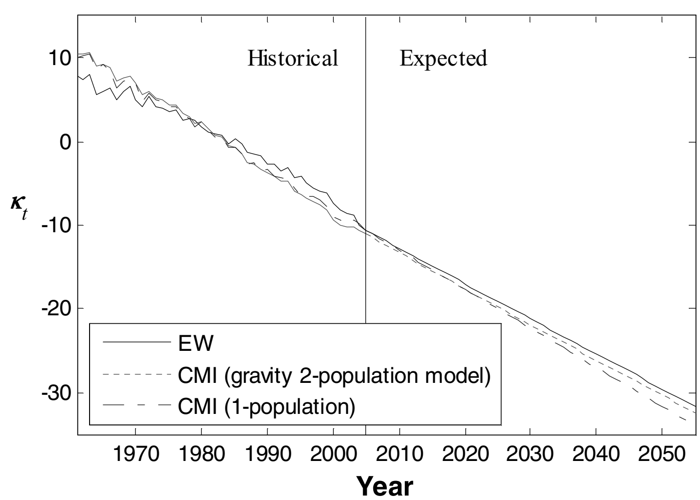
6. PROJECTIONS OF THE GRAVITY TWO-POPULATION MODEL

Based on the estimates for the κ_t parameters in Table 1, Figure 6 shows the projections of expected κ_t over horizons of up to 50 years ahead for the E&W population (the continuous line) modeled as an independent single population and for the CMI population using both the two-population gravity model (dotted) and an independent one-population model (dash-dotted) cases, respectively. All three curves move down over time, but for the single-population case, the expected CMI κ_t curve moves farther away from the E&W curve over time, whereas, for the gravity case, the expected CMI κ_t curve moves close to it. The kappa gravity effect is thus operating along the lines anticipated.

To provide further insight, Figure 7 shows the fan chart projections for the same two cases: We see that the gravity two-population projections are a little higher than the single- or one-population projections, but also a little less dispersed.

Figure 8 shows the expected γ_c for the E&W population (the continuous line) and for the CMI population under both the gravity two-population model (dotted) and the one-population model (dash-dot). We see how, under the latter, the CMI expected γ_c starts off lower than the E&W γ_c but declines at a slower rate (and would, in fact, eventually cross the E&W curve), whereas under the gravity model, the CMI curve moves closer toward the E&W curve and then shadows it in parallel. Thus, once again, the gravity effect is operating as anticipated.

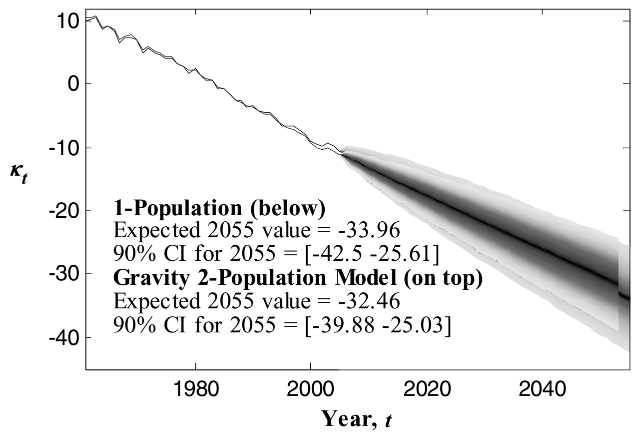
Figure 6
Historical and Expected Future κ_t



Note: Based on LifeMetrics E&W males data and CMI assured male lives data for 1961–2005 and estimated over ages 60–84 and 10,000 simulation trials.

[7] The alert reader will note three anomalies in Table 2: $\hat{\alpha}^{(\gamma 1)}$ is 0.3024 under the gravity model but 0.2743 when modeled independently; $\hat{\mu}^{(\gamma 1)}$ is -0.0318 under the gravity model but -0.0378 when modeled independently, and $V^{(\gamma 11)}$ is 0.0872 under the gravity model but 0.0901 when modeled independently. These are anomalies because these particular parameters should in theory be the same under the gravity model or when modeled independently. However, they differ here for two quite subtle reasons, both of which can lead to unimportant differences between the two model estimates of these parameters. First, when modeled independently, these parameters are estimated using a univariate AR(1) process, but when estimated under the gravity model, they are estimated using the (somewhat different) algorithm described in Appendix C. Second, this algorithm takes as its inputs estimates of $\alpha^{(\gamma 1)}$, $\alpha^{(\gamma 2)}$, and $\phi^{(\gamma)}$, and these will typically differ between the single-population and gravity model cases.

Figure 7
Fan Chart Projections of One-Population and Gravity Two-Population Model CMI κ_t



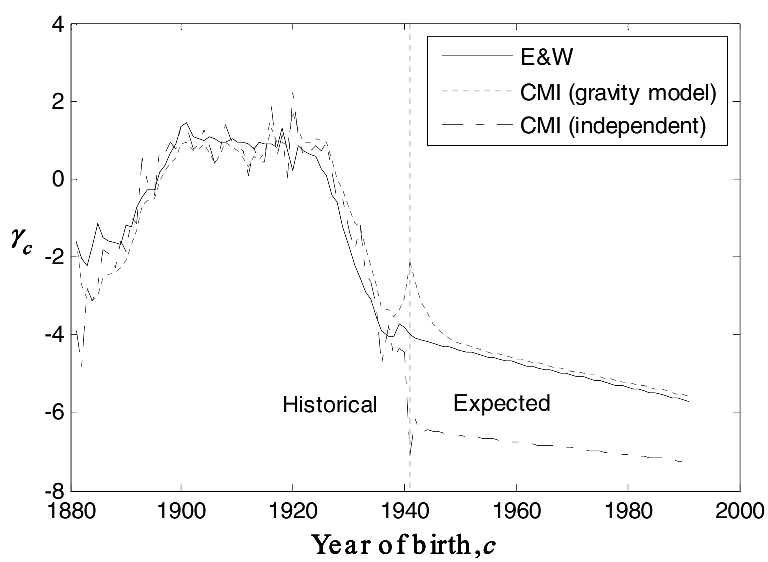
Note: Based on LifeMetrics E&W males data and CMI assured male lives data for 1961–2005 and estimated over ages 60–84 and 10,000 simulation trials.

Figure 9 shows the corresponding γ_c fan chart projections. The gravity projections are somewhat higher than the projections under the one-population case (as we would expect from the above) but are also noticeably narrower.

Having established the impact of the gravity effects on the expected values of the state variables, Figure 10 shows the fan charts for the two populations considered independently, Figure 11 shows the fan charts for the two populations under the gravity model, and Figure 12 shows the CMI fan charts for both models, all for age 65.

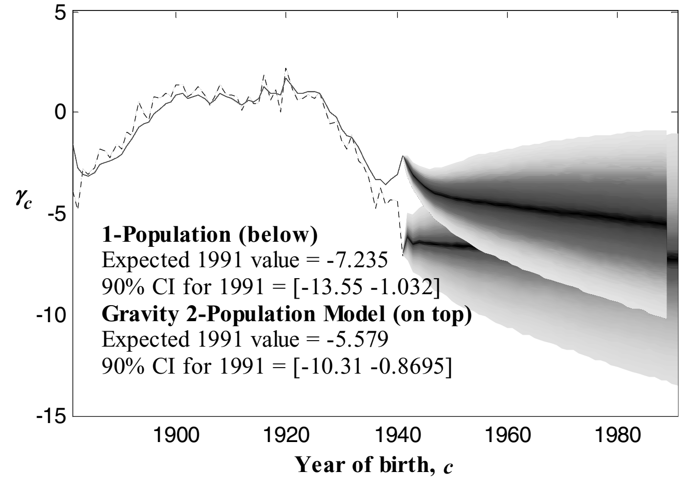
We can see that, relative to considering the two populations independently, the gravity model narrows the CMI projections a little (e.g., the 90% prediction bounds for $T = 50$ change from [0.0021 0.0050]

Figure 8
Historical and Expected Future γ_c



Note: Based on LifeMetrics E&W males data and CMI assured male lives data for 1961–2005 and estimated over ages 60–84 and 10,000 simulation trials.

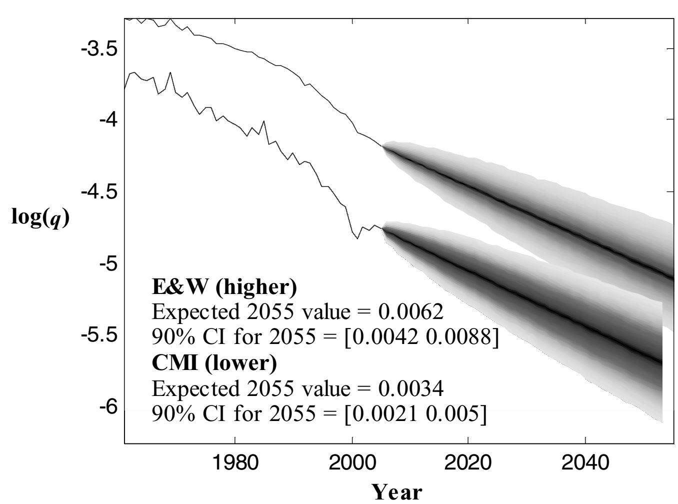
Figure 9
Fan Chart Projections of One-Population and Gravity Two-Population Model CMI γ_c



Note: Based on LifeMetrics E&W males data and CMI assured male lives data for 1961–2005 and estimated over ages 60–84 and 10,000 simulation trials.

to [0.0023 0.0047]).[8] To put this another way, if the gravity model is correct and if one ignores this when modeling CMI male q, then one will obtain fan charts that underestimate somewhat the uncertainty in the projections.

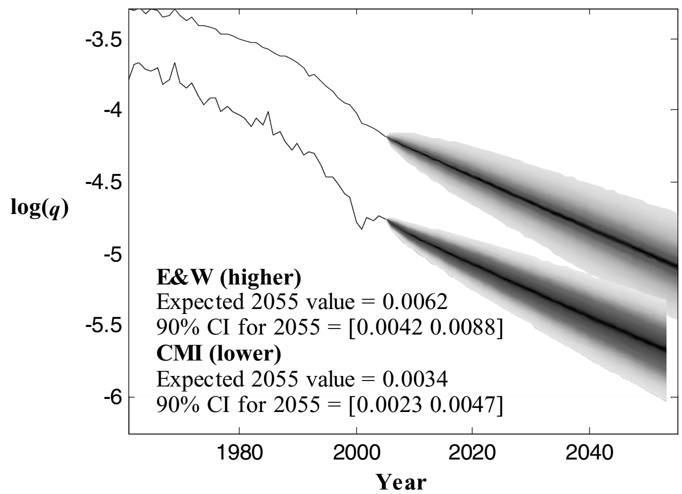
Figure 10
Fan Chart Projections of E&W and CMI q Modeled Independently: Males Age 65



Note: Based on LifeMetrics E&W males data and CMI assured male lives data for 1961–2005 and estimated over ages 60–84 and 10,000 simulation trials.

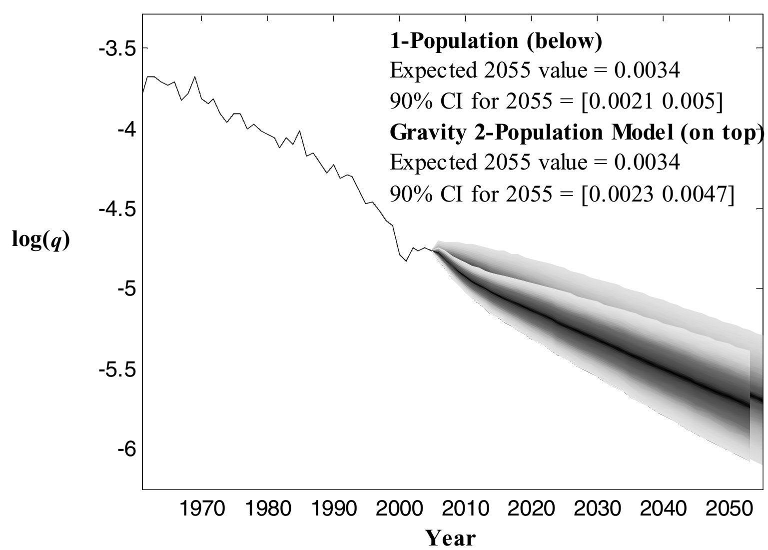
[8] The explanation for the gravity effect narrowing the CMI projections is that the one-population CMI gammas are rather more volatile than the E&W gammas, and so the impact of the gravity model is to smooth the CMI gammas, and it this smoothing that leads to narrower CMI projections. These results are also broadly consistent with those of Cairns et al. (2011b).

Figure 11
Fan Chart Projections of E&W and CMI *q* under the Gravity Model: Males Age 65



Note: Based on LifeMetrics E&W males data and CMI assured male lives data for 1961–2005 and estimated over ages 60–84 and 10,000 simulation trials.

Figure 12
Fan Chart Projections of One-Population and Gravity Two-Population CMI *q*: Males Age 65



Note: Based on LifeMetrics E&W males data and CMI assured male lives data for 1961–2005 and estimated over ages 60–84 and 10,000 simulation trials.

7. VALUING ANNUITIES

The next two sections discuss applications of the model. One application is to the estimation of mortality-related variables (such as expected future lifetimes [EFLs] or annuity prices). Unfortunately, the APC model does not permit us to directly simulate q probabilities for very high ages (in the case of our sample, beyond age 84) because the model requires age (beta) parameters for each age in our

Table 3
Illustrative Term Annuity Prices

Population	Age = 65
E&W	10.8996
CMI (gravity model)	11.7627
CMI (independent)	11.8373

Note: Estimated over ages 60–85 and 1961–2005. Refers to the price of a term annuity that pays 1 unit for each year survived to age 85. Based on an interest rate of 4% and 10,000 simulation trials.

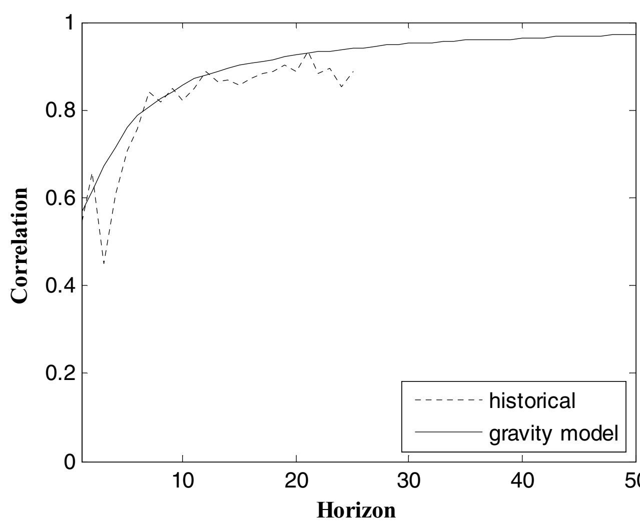
sample range.[9] This means that we cannot directly estimate EFLs or the prices of lifetime annuities, but we can still estimate the prices of term annuities whose payments end by the 85th birthday.

Table 3 shows some illustrative prices for age 65 for term annuities that end at age 85. The cases considered are E&W and CMI both with and without gravity effects. We assume a constant interest rate of 4%. We find that the estimated price for a CMI term annuity is 11.8373 in the case where CMI mortality is modeled independently with the one-population model and 11.7627 where the two-population gravity model is used. Thus, ignoring the gravity effect in this case leads to an overestimation of the price of the annuity of 0.63%.[10]

8. CORRELATIONS BETWEEN *q* MORTALITY IMPROVEMENTS

A second application of the model is the analysis of mortality improvement correlations. Figure 13 shows plots of correlations between mortality or q improvements, for age 65 for horizons running from 1 to 25 years ahead, from starting year 2005. For any given population and age x, the q reduction

Figure 13
Correlations between E&W and CMI Improvements: Males Age 65



Note: Based on LifeMetrics E&W males data and CMI assured male lives data for 1961–2005 and estimated over ages 60–84 and 10,000 simulation trials.

[9] This said, we can still simulate q probabilities for higher ages if we have a method to extrapolate estimates of higher age death rates. An example is provided by Coughlan et al. (2011).

[10] As a comparator, this difference is smaller than that between the annuity prices offered by different annuity providers (see, e.g., Gunawardena et al. 2008).

factor for a horizon H years ahead is defined to be $q(t + H, x)/q(t, x)$ for year $t = 2005$. We see that the correlations between the mortality improvements for the E&W and CMI populations under the gravity model start from a little under 0.6 for $H = 1$ and rise sharply as H increases and eventually approach 1 in the limit.[11] By contrast, the corresponding correlations in the zero-gravity case are, in fact, all equal to 0, reflecting the fact that, in this case, the populations are independent. Thus, taking account of the gravity effect is essential if we are to produce sensible correlations in mortality improvements between related populations.

In addition, as a reality check, the figure also includes a plot of the corresponding historical reduction-factor correlations over our sample period.[12] We can see that the model-based correlations are quite close to the historical correlations present in the data.

9. CONCLUSIONS

This paper has introduced a new type of model, which we have called a gravity model, to handle the interdependence between the mortality rates of two related populations. This model has a number of attractions:

- The gravity concept is intuitively plausible: In the large-population versus small-population context, gravity effects serve to bring the state variables of the secondary population toward those of the dominant primary population in a manner consistent with biological reasonableness.
- Given that the secondary population (in our illustration, the CMI population) is much smaller, an additional benefit of the gravity model is to smooth out some of imprecision in the forecasts of the secondary population.
- The approach is fast and easy to implement using MATLAB: A single fan chart with 10,000 simulation trials takes about 15 seconds to produce.
- The approach can be used to price mortality-related financial contracts (e.g., annuity prices) or model other mortality-related variables (e.g., mortality-improvement correlations, which are useful for hedging purposes).

This study applied the model to the case of E&W males versus CMI assured male lives mortality rates. Applied to this dataset, our results suggest that, relative to the case where CMI mortality is modeled independently, the gravity model tends to slightly raise the expected mortality projection (and, relatedly, slightly increase annuity prices), smooth somewhat the dispersion (or volatility) of projections of CMI q rates, and, perhaps most importantly, can produce forward mortality-improvement correlations that closely match those found in the historical data. These findings would suggest that the use of a gravity model has the potential, at least, to produce considerable improvements in the effectiveness of the risk management of mortality-linked exposures.

APPENDIX A

ESTIMATING THE STATE VARIABLES AND THEIR PARAMETERS

This appendix explains the estimation of the latent state variables and their parameters. The estimation process involves the following stages:

Stage 1: Preliminary estimation of the state variables ignoring any interdependence. We estimate equations (8) and (10) separately to obtain estimates of the beta, kappa, and gamma parameters, ignoring any interdependence between the two populations' state variables.

---

[11] The correlations in Figure 13 are also very close to those reported in Figure 13 in Cairns et al. (2011b) and Coughlan et al. (2011).

[12] For our sample of size 45 years and horizon H, these are the correlations between the reduction factors $q(t + H, 65)/q(t, 65)$ for years $t = 1961 - 2005 : H$, thus giving us the correlation between two samples of 44 observations for $H = 1$, the correlation between two samples of 43 observations for $H = 2$, and so on. We do not calculate the historical correlations beyond $H = 25$ because of the small samples involved.

Stage 2: Estimation of the parameters of the state-variable processes. We then estimate the parameters of the two-population κ_t processes, taking account of possible interdependence, but taking the estimates of the state variables as given. The algorithms used to do this are set out in Appendices B and C. Stages 1 to 2 constitute the first complete estimation cycle.

Stage 3: Reestimation of the state variables taking account of possible interdependence. We then take the parameters of the state variables as given and reestimate the state variables using the algorithms set out in Appendix D.

Stage 4: Repeat stages 2 and 3 until parameter estimates converge. The second and subsequent estimation cycles repeat stages 2 and 3. We input our state variable estimates from stage 3 and repeat stage 2 to obtain a new set of parameter estimates, and we use these to repeat stage 3 to obtain a new set of state variable estimates. We then repeat the estimation cycle as many times as required until parameter estimates converge.

This gives us estimates of both latent state variables and the parameters of the processes governing the state variables that are consistent with each other and take appropriate account of the interdependence between the two populations.

APPENDIX B

ESTIMATING THE PARAMETERS OF THE κ_t PROCESS

This appendix explains the algorithms used to estimate the parameters of the two populations' κ_t processes to take account of their interdependence, while taking the estimates of the state variables as given.

We start with equations (11)–(12) reproduced below:

$$\kappa_t^{(1)} = \kappa_{t-1}^{(1)} + \mu^{(1)} + C^{(12)} Z_t^{(2)}, \tag{11}$$

$$\kappa_t^{(2)} = \kappa_{t-1}^{(2)} + \phi^{(\kappa)}(\kappa_{t-1}^{(1)} - \kappa_{t-1}^{(2)}) + \mu^{(2)} + C^{(21)} Z_t^{(1)} + C^{(22)} Z_t^{(22)}. \tag{12}$$

Imagine to begin that we know the value of $\phi^{(\kappa)}$. Note that (with the exception of $\mu^{(1)}$ and $C^{(1)}$) the μ and C parameters are dependent on $\phi^{(\kappa)}$, so we make this dependence explicit and then, in turn, replace these parameters with their estimates. This gives us

$$\kappa_t^{(1)} = \kappa_{t-1}^{(1)} + \hat{\mu}^{(1)}|\phi^{(\kappa)} + (\hat{C}^{(12)}|\phi^{(\kappa)}) Z_t^{(2)}, \tag{B1}$$

$$\kappa_t^{(2)} = \kappa_{t-1}^{(2)} + \phi^{(\kappa)}(\kappa_{t-1}^{(1)} - \kappa_{t-1}^{(2)}) + \hat{\mu}^{(2)}|\phi^{(\kappa)} + (\hat{C}^{(21)}|\phi^{(\kappa)}) Z_t^{(1)} + (\hat{C}^{(22)}|\phi^{(\kappa)}) Z_t^{(22)}. \tag{B2}$$

Our estimates of the μ parameters conditional on $\phi^{(\kappa)}$ are now

$$\hat{\mu}^{(1)}|\phi^{(\kappa)} = \frac{1}{n_y - 1} \sum_{t=2}^{n_y} (\kappa_t^{(1)} - \kappa_{t-1}^{(1)}), \tag{B3}$$

$$\hat{\mu}^{(2)}|\phi^{(\kappa)} = \frac{1}{n_y - 1} \sum_{t=2}^{n_y} (\kappa_t^{(2)} - (1 - \phi^{(\kappa)})\kappa_{t-1}^{(2)} - \phi^{(\kappa)}\kappa_{t-1}^{(1)}). \tag{B4}$$

Now let

$$\hat{D}_t^{(1)}|\phi^{(\kappa)} = \kappa_t^{(1)} - \kappa_{t-1}^{(1)} - \hat{\mu}^{(1)}|\phi^{(\kappa)}, \tag{B5}$$

$$\hat{D}_t^{(2)}|\phi^{(\kappa)} = \kappa_t^{(2)} - (1 - \phi^{(\kappa)})\kappa_{t-1}^{(2)} - \phi^{(\kappa)}\kappa_{t-1}^{(1)} - \hat{\mu}^{(2)}|\phi^{(\kappa)}, \tag{B6}$$

for $t = 2, \ldots, n_y$. For $i, j = 1, 2$:

$$\hat{V}^{(ij)}|\phi^{(\kappa)} = \frac{1}{n_y - 1 + \xi} \left\{ \sum_{t=2}^{n_y} (\hat{D}_t^{(i)}|\phi^{(\kappa)})(\hat{D}_t^{(j)}|\phi^{(\kappa)}) + \xi\omega_{ij} \right\}, \tag{B7}$$

where ξ is a weight (set to 1) attached to a Bayesian *prior* estimate for $\hat{V}$ given by $\omega = \begin{pmatrix} \omega_{11} & \omega_{12} \\ \omega_{21} & \omega_{22} \end{pmatrix}$, and we take this $\hat{V}$ estimate to be $\omega = \begin{pmatrix} \hat{V}^{(11)} & 0 \\ 0 & \hat{V}^{(22)} \end{pmatrix}$ using the original one-population estimate of $\hat{V}$ in Table 1.[13] This gives us our estimate of V conditional on $\phi^{(\kappa)}$:

$$\hat{V}|\phi^{(\kappa)} = \begin{bmatrix} \hat{V}^{(11)}|\phi^{(\kappa)}, \hat{V}^{(12)}|\phi^{(\kappa)} \\ \hat{V}^{(21)}|\phi^{(\kappa)}, \hat{V}^{(22)}|\phi^{(\kappa)} \end{bmatrix}. \tag{B8}$$

$\hat{C}$ is then the Choleski decomposition of $\hat{V}$.

Finally, we need an MLE method to estimate $\phi^{(\kappa)}$. If we let $\hat{W}|\phi^{(\kappa)} = \begin{bmatrix} \hat{W}^{(11)}|\phi^{(\kappa)}, \hat{W}^{(12)}|\phi^{(\kappa)} \\ \hat{W}^{(21)}|\phi^{(\kappa)}, \hat{W}^{(22)}|\phi^{(\kappa)} \end{bmatrix} = (\hat{V}|\phi^{(\kappa)})^{-1}$, the log-likelihood function is now

$$\begin{aligned} l_k(\phi^{(\kappa)}) = &-\frac{n_y - 1}{2} \ln[\det(\hat{V}|\phi^{(\kappa)})] - \frac{1}{2}\sum_{t=2}^{n_y} (\hat{D}_t^{(1)}|\phi^{(\kappa)})(\hat{D}_t^{(1)}|\phi^{(\kappa)})\hat{W}^{(11)}|\phi^{(\kappa)}) \\ &- \frac{1}{2}\sum_{t=2}^{n_y} (\hat{D}_t^{(1)}|\phi^{(\kappa)})(\hat{D}_t^{(2)}|\phi^{(\kappa)})\hat{W}^{(12)}|\phi^{(\kappa)} - \frac{1}{2}\sum_{t=2}^{n_y} (\hat{D}_t^{(2)}|\phi^{(\kappa)})(\hat{D}_t^{(1)}|\phi^{(\kappa)})\hat{W}^{(21)}|\phi^{(\kappa)} \\ &- \frac{1}{2}\sum_{t=2}^{n_y} (\hat{D}_t^{(2)}|\phi^{(\kappa)})(\hat{D}_t^{(2)}|\phi^{(\kappa)})\hat{W}^{(22)}|\phi^{(\kappa)}. \end{aligned} \tag{B9}$$

We then add to this the log of the prior distribution, which consists of the sum of: the log of the prior density for $\hat{V}(\theta)$ alluded to in the discussion below equation (B7), taken to be the Inverse-Wishart distribution with inverse scaling matrix $\psi = (\psi_{ij}) = \xi\omega = \xi(\omega_{ij})$ and $m = \xi - 3$ degrees of freedom, and the log of the prior density for $\phi^{(\kappa)}$, taken to be a beta $(\xi(v - 1) + 1, \xi(w - 1) + 1)$ distribution with $v = 2$, $w = 2$ evaluated at $\hat{\phi}^{(\kappa)}$. The log of the prior distribution is therefore

$$\begin{aligned} l_{\kappa}^{(prior)}(\phi^{(\kappa)}) = &- \frac{\xi}{2}\ln[\det(\hat{V}|\phi^{(\kappa)})] - \frac{1}{2}\operatorname{trace}(\xi\omega(\hat{V}|\phi^{(\kappa)})^{-1}) \\ &+ \ln[\beta(\hat{\phi}^{(\kappa)}, \xi(v - 1) + 1, \xi(w - 1) + 1)]. \end{aligned} \tag{B10}$$

We can then obtain an MLE estimate of $\phi^{(\kappa)}$ by searching across the permissible $\phi^{(\kappa)}$ range [0,1] and selecting the $\phi^{(\kappa)}$ value that maximizes the sum of equations (B9) and (B10). Once we have this value, we use equations (B1), (B2), and (B6) to estimate μ and V.

The results reported in this paper were based on an illustrative value of $\xi = 5$.

APPENDIX C

ESTIMATING THE PARAMETERS OF THE γ_c PROCESS

This appendix explains the algorithms used to estimate the parameters of the two-population γ_c process, taking estimates of the gamma state variables as given.

Recall the two-population γ_c process (13):

$$\gamma_c = \Phi_1\gamma_{c-1} + \Phi_2\gamma_{c-2} + \begin{pmatrix} \mu^{(\gamma 1)}(1 - \alpha^{(\gamma 1)}) \\ \mu^{(\gamma 2)}(1 - \alpha^{(\gamma 2)}) \end{pmatrix} + C^{(\gamma)}Z_c^{(\gamma)}, \tag{13}$$

[13] The inclusion of the ξ and ω_{ij} terms in equation (B7) serves to penalize the oversmoothing of the secondary population κ_t process that might otherwise occur over the course of repeated estimation cycles.

where $\Phi_1 = \begin{bmatrix} 1 + \alpha^{(\gamma 1)}, 0 \\ \phi^{(\gamma)}, 1 + \alpha^{(\gamma 2)} - \phi^{(\gamma)} \end{bmatrix}$ and $\Phi_2 = \begin{bmatrix} -\alpha^{(\gamma 1)}, 0 \\ 0, -\alpha^{(\gamma 2)} \end{bmatrix}$.

We now expand equation (13) to obtain

$$\gamma_c^{(1)} = (1 + \alpha^{(\gamma 1)})\gamma_{c-1}^{(1)} - \alpha^{(\gamma 1)}\gamma_{c-2}^{(1)} + \mu^{(\gamma 1)}(1 - \alpha^{(\gamma 1)}) + C^{(\gamma 11)}Z_c^{(\gamma 1)} + C^{(\gamma 12)}Z_c^{(\gamma 2)}, \quad \text{(C1)}$$

$$\gamma_c^{(2)} = \phi^{(\gamma)}\gamma_{c-1}^{(1)} + (1 + \alpha^{(\gamma 1)} - \phi^{(\gamma)})\gamma_{c-1}^{(2)} - \alpha^{(\gamma 2)}\gamma_{c-2}^{(2)} + \mu^{(\gamma 2)}(1 - \alpha^{(\gamma 2)}) + C^{(\gamma 21)}Z_c^{(\gamma 1)} + C^{(\gamma 22)}Z_c^{(\gamma 2)}. \quad \text{(C2)}$$

Given $(\alpha^{(\gamma 1)}, \alpha^{(\gamma 2)}, \phi^{(\gamma)}) = \theta$, we can then obtain estimates of the $\mu^{(\gamma)}$ parameters conditional on θ:

$$\hat{\mu}^{(\gamma 1)}|\theta = \frac{1}{(n_c - 2)(1 - \alpha^{(\gamma 1)})} \sum_{c=3}^{n_c} (\gamma_c^{(1)} - (1 + \alpha^{(\gamma 1)})\gamma_{c-1}^{(1)} + \alpha^{(\gamma 1)}\gamma_{c-2}^{(1)}), \quad \text{(C3)}$$

$$\hat{\mu}^{(\gamma 2)}|\theta = \frac{1}{(n_c - 2)(1 - \alpha^{(\gamma 2)})} \sum_{c=3}^{n_c} (\gamma_c^{(2)} - \phi^{(\gamma)}\gamma_{c-1}^{(1)} - (1 + \alpha^{(\gamma 2)} - \phi^{(\gamma)})\gamma_{c-1}^{(2)} + \alpha^{(\gamma 2)}\gamma_{c-2}^{(2)}), \quad \text{(C4)}$$

for $c = 3, \ldots, n_c$. Now let

$$\hat{D}_c^{(\gamma 1)}|\theta = \gamma_c^{(1)} - (1 + \alpha^{(\gamma 1)})\gamma_{c-1}^{(1)} + \alpha^{(\gamma 1)}\gamma_{c-2}^{(1)} - (\hat{\mu}^{(\gamma 1)}|\theta)(1 - \alpha^{(\gamma 1)}), \quad \text{(C5)}$$

$$\hat{D}_c^{(\gamma 2)}|\theta = \gamma_c^{(2)} - \phi^{(\gamma)}\gamma_{c-1}^{(1)} - (1 + \alpha^{(\gamma 2)} - \phi^{(\gamma)})\gamma_{c-1}^{(2)} + \alpha^{(\gamma 2)}\gamma_{c-2}^{(2)} - (\hat{\mu}^{(\gamma 2)}|\theta)(1 - \alpha^{(\gamma 2)}). \quad \text{(C6)}$$

For $i, j = 1, 2$:

$$\hat{V}^{(\gamma ij)}|\theta = \frac{1}{n_c - 2 + \xi^{(\gamma)}} \left\{ \sum_{c=3}^{n_c} (\hat{D}_c^{(\gamma i)}|\theta)(\hat{D}_c^{(\gamma j)}|\theta) + \xi^{(\gamma)}\omega_{ij}^{(\gamma)} \right\}, \quad \text{(C7)}$$

where $\xi^{(\gamma)}$ is a weight (set to 1) attached to a Bayesian *prior* estimate for $\hat{V}^{(\gamma)}$ given by $\omega^{(\gamma)} = \begin{pmatrix} \omega_{11}^{(\gamma)} & \omega_{12}^{(\gamma)} \\ \omega_{21}^{(\gamma)} & \omega_{22}^{(\gamma)} \end{pmatrix}$, and we take this $\hat{V}^{(\gamma)}$ estimate to be $\omega^{(\gamma)} = \begin{pmatrix} \hat{V}^{(\gamma 11)} & 0 \\ 0 & \hat{V}^{(\gamma 11)} \end{pmatrix}$ using the original one-population estimate of $\hat{V}^{(\gamma 11)}$ in Table 2.[14]

Equation (C7) gives us estimates of the $V^{(\gamma)}$ parameters conditional on θ:

$$\hat{V}^{(\gamma)}|\theta = \begin{bmatrix} \hat{V}^{(\gamma 11)}|\theta, \hat{V}^{(\gamma 12)}|\theta \\ \hat{V}^{(\gamma 21)}|\theta, \hat{V}^{(\gamma 22)}|\theta \end{bmatrix}. \quad \text{(C8)}$$

$\hat{C}^{(\gamma)}$ is then the Choleski decomposition of $\hat{V}^{(\gamma)}|\theta$.

Now let

$$\hat{W}^{(\gamma)}|\theta = \begin{bmatrix} \hat{W}^{(\gamma 11)}|\theta, \hat{W}^{(\gamma 12)}|\theta \\ \hat{W}^{(\gamma 21)}|\theta, \hat{W}^{(\gamma 22)}|\theta \end{bmatrix} = (\hat{V}^{(\gamma)}|\theta)^{-1}.$$

The log-likelihood function is then

$$\begin{aligned} l_\gamma(\theta) = &-\frac{n_c - 2}{2} \ln[\det(\hat{V}^{(\gamma)}|\theta)] - \frac{1}{2}\sum_{c=3}^{n_c} (\hat{D}_c^{(\gamma 1)}|\theta)(\hat{D}_c^{(\gamma 1)}|\theta)\hat{W}^{(\gamma 11)}|\theta \\ &- \frac{1}{2}\sum_{c=3}^{n_c} (\hat{D}_c^{(\gamma 1)}|\theta)(\hat{D}_c^{(\gamma 2)}|\theta)\hat{W}^{(\gamma 12)}|\theta - \frac{1}{2}\sum_{c=3}^{n_c} (\hat{D}_c^{(\gamma 2)}|\theta)(\hat{D}_c^{(\gamma 1)}|\theta)\hat{W}^{(\gamma 21)}|\theta \\ &- \frac{1}{2}\sum_{c=3}^{n_c} (\hat{D}_c^{(\gamma 2)}|\theta)(\hat{D}_c^{(\gamma 2)}|\theta)\hat{W}^{(\gamma 22)}|\theta. \end{aligned} \quad \text{(C9)}$$

[14] The inclusion of the $\xi^{(\gamma)}$ and $\omega_{ij}^{(\gamma)}$ terms in equation (C7) serves to penalize the oversmoothing of the secondary population cohort process that might otherwise occur over the course of repeated estimation cycles.

We then add to this the log of the prior distribution, which consists of the sum of: the log of the prior density for $\hat{V}^{(\gamma)}(\theta)$, taken to be the Inverse-Wishart distribution with inverse scaling matrix $\psi^{(\gamma)} = (\psi_{ij}^{(\gamma)}) = \xi^{(\gamma)}\omega^{(\gamma)} = \xi(\omega_{ij}^{(\gamma)})$ and $m = \xi - 3$ degrees of freedom, and the log of the prior density for $\phi^{(\gamma)}$, taken to be a beta $(\xi(v - 1) + 1, \xi(w - 1) + 1)$ distribution with $v = 2$, $w = 2$ evaluated at $\hat{\phi}^{(\gamma)}$. The log of the prior distribution is therefore

$$l_{\gamma}^{(prior)}(\phi^{(\gamma)}) = -\frac{\xi^{(\gamma)}}{2}\ln[\det(\hat{V}^{(\gamma)}|\phi^{(\gamma)})] - \frac{1}{2}\operatorname{trace}(\xi^{(\gamma)}\omega^{(\gamma)}(\hat{V}^{(\gamma)}|\phi^{(\gamma)})^{-1}) + \ln[\beta(\hat{\phi}^{(\gamma)}, \xi(v - 1) + 1, \xi(w - 1) + 1)]. \quad \text{(C10)}$$

We can then obtain MLE estimates of $\alpha^{(\gamma 1)}$, $\alpha^{(\gamma 2)}$, and $\phi^{(\gamma)}$ by searching across the permissible ranges of these parameters, and selecting those that maximize the sum of equations (C9) and (C10).

The results reported in this paper were based on an illustrative value of $\xi^{(\gamma)} = 5$.

APPENDIX D

REESTIMATING THE SMALLER POPULATION STATE VARIABLES TAKING ACCOUNT OF THEIR INTERDEPENDENCE WITH THE LARGER POPULATION

This appendix explains the procedure used to reestimate the CMI state variables, taking account of their interdependence with the E&W state variables, but taking the parameters governing the state variables as given. Note, of course, that there is no need to reestimate the E&W state variables because they are independent of the CMI state variables.

Before doing this, we define the standard deviations and correlations of the noise processes governing each state variable using obvious notation: $\sigma^{(1)} = C^{(11)}$, $\sigma^{(2)} = \sqrt{C^{(21)2} + C^{(22)2}}$, $\rho = C^{(21)}/\sqrt{C^{(21)2} + C^{(22)2}})$, $\sigma^{(\gamma 1)} = C^{(\gamma 11)}$, $\sigma^{(\gamma 2)} = \sqrt{C^{(\gamma 21)2} + C^{(\gamma 22)2}}$, and $\rho^{(\gamma)} = C^{(\gamma 21)}/\sqrt{C^{(\gamma 21)2} + C^{(\gamma 22)2}}$. At this point we also drop the ''hat'' notation to avoid cluttering up the equations.

We now go through the following four stages:

Stage 1: Updating the CMI kappas. We begin by backing out historical standardized errors for population-1 using

$$Z_{t+1} = (\kappa_{t+1}^{(1)} - \kappa_t^{(1)} - \mu^{(1)})/\sigma^{(1)} \quad \text{(D1)}$$

for $t = 1, \ldots, n_y - 1$ and setting $Z_1 = 0$.

Let D, E, and m, respectively, be the deaths, exposures, and latest available APC death-rate matrices for CMI, each of which has dimension $n_y \times n_a$. Now let l_P be the log-likelihood of the CMI Poisson deaths process given by

$$l_P = \sum_{t,x} (D_{tx} \log m_{tx} - E_{tx}m_{tx}) + \text{const.} \quad \text{(D2)}$$

The first and second derivatives of l_P with respect to $\kappa_t^{(2)}$ are

$$\frac{\partial l_P}{\partial \kappa_t^{(2)}} = \sum_x (D_{tx} - E_{tx}m_{tx})/n_a, \quad \text{(D3)}$$

$$\frac{\partial^2 l_P}{\partial \kappa_t^{(2)2}} = \sum_x - E_{tx}m_{tx}/n_a^2. \quad \text{(D4)}$$

Now let l_κ be the log-likelihood of the CMI kappa process, which is given as

$$l_\kappa = \sum_{t=2}^{n_y} \left\{ -\frac{1}{2}\log(\sigma^{(2)2}(1 - \rho_\kappa^2)) - \frac{(\kappa_t^{(2)} - (1 - \phi^{(\kappa)})\kappa_{t-1}^{(2)} + \phi^{(\kappa)}\kappa_{t-1}^{(1)} - \mu^{(\kappa 2)} - \sigma^{(2)}\rho_\kappa Z_t^{(\kappa 1)})^2}{2\sigma^{(2)2}(1 - \rho_\kappa^2)} \right\}. \quad \text{(D5)}$$

Its first derivatives with respect to $\kappa_t^{(2)}$ are

$$\frac{\partial l_\kappa}{\partial \kappa_1^{(2)}} = \frac{(1 - \phi^{(\kappa)})}{\sigma^{(2)2}(1 - \rho^{(\kappa)2})} \{\kappa_2^{(2)} - (1 - \phi^{(\kappa)})\kappa_1^{(2)} - \phi^{(\kappa)}\kappa_1 - \mu^{(2)} - \sigma^{(2)}\rho^{(\kappa)}Z_2^{(1)}\}, \tag{D6a}$$

$$\frac{\partial l_\kappa}{\partial \kappa_t^{(2)}} = \frac{1}{\sigma^{(2)2}(1 - \rho^{(\kappa)2})} \{(1 - \phi^{(\kappa)})\kappa_{t+1}^{(2)} - (1 + (1 - \phi^{(\kappa)})^2)\kappa_t^{(2)} + (1 - \phi^{(\kappa)})\kappa_{t-1}^{(2)} \tag{D6b}$$

$$+ \phi^{(\kappa)}\kappa_{t-1}^{(1)} - \phi^{(\kappa)}(1 - \phi^{(\kappa)})\kappa_t^{(1)} + \phi^{(\kappa)}\mu^{(2)} + \sigma^{(2)}\rho^{(\kappa)}(Z_t^{(1)} - (1 - \phi^{(\kappa)})Z_{t+1}^{(1)})\},$$

$$\frac{\partial l_\kappa}{\partial \kappa_{n_y}^{(2)}} = \frac{-1}{\sigma^{(2)2}(1 - \rho^{(\kappa)2})} \{\kappa_{n_y}^{(2)} - (1 - \phi^{(\kappa)})\kappa_{n_y-1}^{(2)} - \phi^{(\kappa)}\kappa_{n_y-1}^{(1)} - \mu^{(2)} - \sigma^{(2)}\rho^{(\kappa)}Z_{n_y}^{(1)}\}, \tag{D6c}$$

for the cases $t = 1$, $1 < t < n_y$, and $t = n_y$ respectively.

Its corresponding second derivatives are

$$\frac{\partial^2 l_\kappa}{\partial \kappa_1^{(2)2}} = \frac{-(1 - \phi^{(\kappa)})^2}{\sigma^{(2)2}(1 - \rho^{(\kappa)2})}, \tag{D7a}$$

$$\frac{\partial^2 l_\kappa}{\partial \kappa_t^{(2)2}} = \frac{-(1 + (1 - \phi^{(\kappa)})^2)}{\sigma^{(2)2}(1 - \rho^{(\kappa)2})}, \tag{D7b}$$

$$\frac{\partial^2 l_\kappa}{\partial \kappa_{n_y}^{(2)2}} = \frac{-1}{\sigma^{(2)2}(1 - \rho^{(\kappa)2})}. \tag{D7c}$$

The updating routine for κ_t is then

$$\kappa_t^{(2)new} = \kappa_t^{(2)old} - \frac{\partial l_P/\partial \kappa_t^{(2)} + \partial l_\kappa/\partial \kappa_t^{(2)}}{\partial^2 l_P/\partial \kappa_t^{(2)2} + \partial^2 l_\kappa/\partial \kappa_t^{(2)2}}, \tag{D8}$$

where the partial derivatives are evaluated at the old value of κ_t.

Stage 2: Updating the CMI gammas. The corresponding gamma process is a little more involved. We begin by backing out historical gamma standardized errors for population-1 using a rearranged version of equation (13a):

$$Z_1^{(\gamma 1)} = 0, \tag{D9a}$$

$$Z_2^{(\gamma 1)} = (\gamma_2^{(1)} - (1 + \alpha^{(\gamma 1)})\gamma_1^{(1)} - \mu^{(\gamma 1)}(1 - \alpha^{(\gamma 1)}))/\sigma^{(\gamma 1)}, \tag{D9b}$$

$$Z_c^{(\gamma 1)} = (\gamma_c^{(1)} - (1 + \alpha^{(\gamma 1)})\gamma_{c-1}^{(1)} + \alpha^{(\gamma 1)}\gamma_{c-2}^{(1)} - \mu^{(\gamma 1)}(1 - \alpha^{(\gamma 1)}))/\sigma^{(\gamma 1)}, \tag{D9c}$$

for $c = 3, \ldots, n_c$.

The first and second derivatives of l_P with respect to $\gamma_c^{(2)}$ are

$$\frac{\partial l_P}{\partial \gamma_c^{(2)}} = \sum_{t,x} (D_{tx} - E_{tx}m_{tx})I_{tx}(c)/n_\alpha, \tag{D10}$$

$$\frac{\partial^2 l_P}{\partial \gamma_c^{(2)2}} = \sum_{t,x} - E_{tx}m_{tx}I_{tx}(c)/n_\alpha^2, \tag{D11}$$

where the identity function $I_{tj}(c)$ takes the value 1 if $c = t - x$ and otherwise takes the value 0.

Now let l_γ be the log-likelihood of the CMI gamma process, given by

$$l_\gamma = -\frac{n_c - 2}{2} \ln(\sigma^{(\gamma 2)2}(1 - \rho^{(\gamma)2})$$

$$- \frac{1}{\sigma^{(\gamma 2)2}(1 - \rho^{(\gamma)2})} \sum_{c=3}^{n_c} (\gamma_c^{(2)} - (1 + \alpha^{(\gamma 2)} - \phi^{(\gamma)})\gamma_{c-1}^{(2)} + \alpha^{(\gamma 2)}\gamma_{c-2}^{(2)} - \varepsilon_c - \psi), \tag{D12}$$

where $\varepsilon_1 = 0$, $\varepsilon_c = \phi^{(\gamma)}\gamma^{(1)}_{c-1} + \sigma^{(\gamma2)}\rho^{(\gamma)}Z^{(\gamma1)}_c$ for $c = 2, \ldots, n_c$, $\psi = \mu^{(\gamma2)}(1 - \alpha^{(\gamma2)})$, and $SR2 = \sigma^{(\gamma2)2}(1 - \rho^{(\gamma)2})$.

Its first derivatives are

$$\frac{\partial l_\gamma}{\partial \gamma^{(2)}_1} = \frac{-1}{SR2}\{\alpha^{(\gamma2)}[\gamma^{(2)}_3 - (1 + \alpha^{(\gamma2)} - \phi^{(\gamma)})\gamma^{(2)}_2 + \alpha^{(\gamma2)}\gamma^{(2)}_1 - \varepsilon_3 - \psi]\}, \tag{D13a}$$

$$\frac{\partial l_\gamma}{\partial \gamma^{(2)}_2} = \frac{-1}{SR2}\{\alpha^{(\gamma2)}[\gamma^{(2)}_4 - (1 + \alpha^{(\gamma2)} - \phi^{(\gamma)})\gamma^{(2)}_3 + \alpha^{(\gamma2)}\gamma^{(2)}_2 - \varepsilon_4 - \psi] - (1 + \alpha^{(\gamma2)} - \phi^{(\gamma)})[\gamma^{(2)}_3 - (1 + \alpha^{(\gamma2)} - \theta^{(\gamma)})\gamma^{(2)}_2 + \alpha^{(\gamma2)}\gamma^{(2)}_1 - \varepsilon_3 - \psi]\}, \tag{D13b}$$

$$\frac{\partial l_\gamma}{\partial \gamma^{(2)}_{n_c-1}} = \frac{-1}{SR2}\{\alpha^{(\gamma2)}[\gamma^{(2)}_{c+2} - (1 + \alpha^{(\gamma2)} - \phi^{(\gamma)})\gamma^{(2)}_{c+1} + \alpha^{(\gamma2)}\gamma^{(2)}_c - \varepsilon_{c+2} - \psi] - (1 + \alpha^{(\gamma2)} - \phi^{(\gamma)})[\gamma^{(2)}_{c+1} - (1 + \alpha^{(\gamma2)} - \phi^{(\gamma)})\gamma^{(2)}_c + \alpha^{(\gamma2)}\gamma^{(2)}_{c-1} - \varepsilon_{c+1} - \psi] + [\gamma^{(2)}_c - (1 + \alpha^{(\gamma2)} - \phi^{(\gamma)})\gamma^{(2)}_{c-1} + \alpha^{(\gamma2)}\gamma^{(2)}_{c-2} - \varepsilon_c - \psi]\}, \tag{D13c}$$

$$\frac{\partial l_\gamma}{\partial \gamma^{(2)}_{n_c-1}} = \frac{-1}{SR2}\{-(1 + \alpha^{(\gamma2)} - \phi^{(\gamma)})[\gamma^{(2)}_{n_c} - (1 + \alpha^{(\gamma2)} - \phi^{(\gamma)})\gamma^{(2)}_{n_c-1} + \alpha^{(\gamma2)}\gamma^{(2)}_{n_c-2} - \varepsilon_{n_c} - \psi] + [\gamma^{(2)}_{n_c-1} - (1 + \alpha^{(\gamma2)} - \phi^{(\gamma)})\gamma^{(2)}_{n_c-2} + \alpha^{(\gamma2)}\gamma^{(2)}_{n_c-3} - \varepsilon_{n_c-1} - \psi]\}, \tag{D13d}$$

$$\frac{\partial l_\gamma}{\partial \gamma^{(2)}_{n_c}} = \frac{-1}{SR2}\{[\gamma^{(2)}_{n_c} - (1 + \alpha^{(\gamma2)} - \phi^{(\gamma)})\gamma^{(2)}_{n_c-1} + \alpha^{(\gamma2)}\gamma^{(2)}_{n_c-2} - \varepsilon_{n_c} - \psi]\}, \tag{D13e}$$

for the cases $c = 1$, $c = 2$, $2 < c < n_c - 2$, $c = n_c - 1$, and $c = n_c$ respectively.

Its corresponding second derivatives are

$$\frac{\partial^2 l_\gamma}{\partial \gamma^{(2)2}_1} = -\frac{\alpha^{(\gamma2)2}}{SR2}, \tag{D14a}$$

$$\frac{\partial^2 l_\gamma}{\partial \gamma^{(2)2}_2} = -\frac{\alpha^{(\gamma2)2} + (1 + \alpha^{(\gamma2)} - \phi^{(\gamma)})^2}{SR2}, \tag{D14b}$$

$$\frac{\partial^2 l_\gamma}{\partial \gamma^{(2)2}_c} = -\frac{\alpha^{(\gamma2)2} + (1 + \alpha^{(\gamma2)} - \phi^{(\gamma)})^2 + 1}{SR2}, \tag{D14c}$$

$$\frac{\partial^2 l_\gamma}{\partial \gamma^{(2)2}_{n_c-1}} = -\frac{(1 + \alpha^{(\gamma2)} - \phi^{(\gamma)})^2 + 1}{SR2}, \tag{D14d}$$

$$\frac{\partial^2 l_\gamma}{\partial \gamma^{(2)2}_{n_c}} = -\frac{1}{SR2}. \tag{D14e}$$

The updating routine for $\gamma^{(2)}_c$ is then

$$\gamma^{(2)new}_c = \gamma^{(2)old}_c - \frac{\partial l_P/\partial \gamma^{(2)}_c + \partial l_\gamma/\partial \gamma^{(2)}_c}{\partial^2 l_P/\partial \gamma^{(2)2}_c + \partial^2 l_\gamma/\gamma^{(2)2}_c}, \tag{D15}$$

where the partial derivatives are evaluated at the old value of γ_t.

Stage 3: Updating the CMI betas. We now update the CMI betas. The first derivative of the Poisson log-likelihood is

$$\frac{\partial l_P}{\partial \beta^{(2)}_x} = \sum_{t=1}^{n_y}\left(D_{tx} - E_{tx}\exp\left[\frac{\kappa^{(2)}_t}{n_a} + \frac{\gamma^{(2)}_{t-x}}{n_a}\right]\exp[\beta^{(2)}_x]\right). \tag{D16}$$

Setting the left-hand side of (D16) to zero then yields the following solution for the revised beta parameter:

$$\beta_x^{(2)new} = \ln\left\{\frac{\sum_{t=1}^{n_y} D_{tx}}{\sum_{t=1}^{n_y} E_{tx}\exp\left[\frac{\kappa_t^{(2)}}{n_a}+\frac{\gamma_{t-x}^{(2)}}{n_a}\right]}\right\}. \qquad \text{(D17)}$$

Stage 4: Applying identifiability constraints. We now need to impose the identifiability constraints, and in the case of the $\kappa_t^{(2)}$, this requires that we take our $\kappa_t^{(2)new}$ series from equation (D7), relabel it as $\kappa_t^{(2)old}$, and then obtain a new series using

$$\kappa_t^{(2)new} = \kappa_t^{(2)old} - \left(\frac{1}{n_y}\sum_{s=1}^{n_y}\kappa_s^{(2)old}\right) \qquad \text{(D18)}$$

with corresponding adjustment

$$\beta_x^{(2)new} = \beta_x^{(2)old} + \frac{1}{n_a}\left(\frac{1}{n_y}\sum_{s=1}^{n_y}\kappa_s^{(2)old}\right). \qquad \text{(D19)}$$

In the case of the $\gamma_c^{(2)}$, applying the identifiability constraint requires that we take our $\gamma_c^{(2)new}$ series from (D15), relabel it as $\gamma_c^{(2)old}$, and then obtain a new series using

$$\gamma_c^{(2)new} = \gamma_c^{(2)old} - \left(\frac{1}{n_a n_y}\sum_{t,x}\gamma_{t-x}^{(2)old}\right) \qquad \text{(D20)}$$

with corresponding adjustment

$$\beta_x^{(2)new} = \beta_x^{(2)old} + \frac{1}{n_a}\left(\frac{1}{n_a n_y}\sum_{t,x}\gamma_{t-x}^{(2)old}\right). \qquad \text{(D21)}$$

Stage 5: Tilting. Finally, we now estimate δ using equation (5) and tilt our parameter estimates using equation (6).

ACKNOWLEDGMENTS AND DISCLAIMER

We would like to thank Liang Zhao for invaluable research assistance. We would also like to thank Richard MacMinn and participants of workshops at Longevity 5: Fifth International Longevity Risk and Capital Market Solutions Conference and Humboldt University for very useful comments on earlier drafts of this paper. This report has been partially prepared by the Pension Advisory Group, and not by any research department, of JPMorgan Chase & Co. and its subsidiaries (‘‘JPMorgan’’). Information herein is obtained from sources believed to be reliable, but JPMorgan does not warrant its completeness or accuracy. Opinions and estimates constitute JPMorgan’s judgment and are subject to change without notice. Past performance is not indicative of future results. This material is provided for informational purposes only and is not intended as a recommendation or an offer or solicitation for the purchase or sale of any security or financial instrument.

REFERENCES

BOOTH, H., AND L. TICKLE. 2008. Mortality Modelling and Forecasting: A Review of Methods. *Annals of Actuarial Science* 3: 3–43.

CAIRNS, A. J. G., D. BLAKE, K. DOWD, G. D. COUGHLAN, D. EPSTEIN, A. ONG, AND I. BALEVICH. 2009. A Quantitative Comparison of Stochastic Mortality Models Using Data from England and Wales and the United States. *North American Actuarial Journal* 13(4): 1–35.

CAIRNS, A. J. G., D. BLAKE, K. DOWD, G. D. COUGHLAN, AND M. KHALAF-ALLAH. 2011a. Mortality Density Forecasts: An Analysis of Six Stochastic Mortality Models. *Insurance: Mathematics & Economics* 48: 355–367.

CAIRNS, A. J. G., D. BLAKE, K. DOWD, G. D. COUGHLAN, AND M. KHALAF-ALLAH. 2011b. Bayesian Stochastic Mortality Modelling for Two Populations. *ASTIN Bulletin* 41(1): 29–59.

COUGHLAN, G. D., M. KHALAF-ALLAH, Y. YE, S. KUMAR, A. J. G. CAIRNS, D. BLAKE, AND K. DOWD. 2011. Longevity Hedging 101: A Framework for Longevity Basis Risk Analysis and Hedge Effectiveness. *North American Actuarial Journal* 15(2): 150–176.

DOWD, K., A. J. G. CAIRNS, D. COUGHLAN, D. EPSTEIN, AND M. KHALAF-ALLAH. 2010a. Evaluating the Goodness of Fit of Stochastic Mortality Models. *Insurance: Mathematics & Economics* 33: 29–47.

DOWD, K., A. J. G. CAIRNS, D. BLAKE, G. D. COUGHLAN, D. EPSTEIN, AND M. KHALAF-ALLAH. 2010b. Backtesting Stochastic Mortality Models: An Ex-Post Evaluation of Multiperiod-Ahead Density Forecasts. *North American Actuarial Journal* 14(3): 281–298.

GUNAWARDENA, D., C. HICKS, AND D. O'NEILL. 2008. Pension Annuities: Pension Annuities and the Open Market Solution. Association of British Insurers Research Paper No. 8.

JACOBSEN, R., N. KEIDING, AND E. LYNGE. 2002. Long-Term Mortality Trends behind Low Life Expectancy of Danish Women. *Journal of Epidemiology and Community Health* 56: 205–208.

JARNER, S. F., AND E. M. KRYGER. 2009. Modelling Adult Mortality in Small Populations: The SAINT Model. Pensions Institute Discussion Paper PI-0902.

LI, J. S. H., AND M. R. HARDY. 2011. Measuring Basis Risk in Longevity Hedges. *North American Actuarial Journal* 15(2): 177–200.

LI, N., AND R. LEE. 2005. Coherent Mortality Forecasts for a Group of Populations: An Extension of the Lee-Carter Method. *Demography* 42(3): 575–594.

OEPPEN, J., AND J. W. VAUPEL. 2002. Broken Limits to Life Expectancy. *Science* 296(5570): 1029–1031.

OSMOND, C. 1985. Using Age, Period and Cohort Models to Estimate Future Mortality Rates. *International Journal of Epidemiology* 14: 124–129.

Discussions on this paper can be submitted until October 1, 2011. The authors reserve the right to reply to any discussion. Please see the Submission Guidelines for Authors on the inside back cover for instructions on the submission of discussions.