The Journal of Risk and Insurance

© 2017 The Journal of Risk and Insurance. Vol. 84, No. S1, 477–493 (2017).
DOI: 10.1111/jori.12198

Modeling Multicountry Longevity Risk With Mortality Dependence: A Lévy Subordinated Hierarchical Archimedean Copulas Approach

Wenjun Zhu
Ken Seng Tan
Chou-Wen Wang

Abstract

This article proposes a new copula model known as the Lévy subordinated hierarchical Archimedean copulas (LSHAC) for multicountry mortality dependence modeling. To the best of our knowledge, this is the first article to apply the LSHAC model to mortality studies. Through an extensive empirical analysis on modeling mortality experiences of 13 countries, we demonstrate that the LSHAC model, which has the advantage of capturing the geographical structure of mortality data, yields better fit, compared to the elliptical copulas. In addition, the proposed LSHAC model generates out-of-sample forecasts with smaller standard deviations, when compared to other benchmark copula models. The LSHAC model also confirms that there is an association between geographical locations and dependence of the overall mortality improvement. These results yield new insights into future longevity risk management. Finally, the model is used to price a hypothetical survival index swap written on a weighted mortality index. The results highlight the importance of dependence modeling in managing longevity risk and reducing population basis risk.

Wenjun Zhu is an Assistant Professor at the School of Finance, Nankai University, Tianjin, China. Zhu can be contacted via e-mail: zhuwenjun@nankai.edu.cn. Ken Seng Tan is the University Research Chair Professor at the Department of Statistics and Actuarial Science, University of Waterloo 200 University Avenue West, Waterloo, N2L 3G1, ON, Canada. Tan can be contacted via e-mail: kstan@uwaterloo.ca. Chou-Wen Wang is at the Department of Finance National Sun Yat-sen University, Kaohsiung, Taiwan Fellow of Risk and Insurance Research Center, College of Commerce National Chengchi University, Taiwan. Wang can be contacted via e-mail: chouwenwang@gmail.com. Zhu acknowledges funding support from the Society of Actuaries (SoA) James C. Hickman Scholar Program and the China Scholarship Council (CSC). Tan acknowledges the research funding from the Natural Sciences and Engineering Research Council of Canada and the Society of Actuaries CAE Research Grant. Wang was also supported in part by the MOST 101-2410-H-327-029- from the Ministry of Science and Technology, R.O.C.

INTRODUCTION

It is estimated that the human life expectancy in developed countries has been increasing almost linearly over the past 150 years (Blake et al., 2013). The unanticipated increases in life expectancy create significant financial burden to both public and private pension plan sponsors. As a result, longevity risk, as attributed to the increase in life expectancy, has been recognized as one of the major risks faced by insurers, reinsurers, governments, and individuals in recent years. Hence, a traded market in longevity-linked securities and derivatives has emerged to facilitate the development of annuity markets and protect the long-term viability of global retirement income provision. For example, as the first longevity-linked derivative transaction, a *q*-forward contract between JPMorgan and the U.K. company Lucida was traded in January 2008. In addition, Swiss Re launched the Kortis longevity bond to transfer USD 50 million of longevity risk to the capital markets.

A number of two-population mortality models have been proposed (see, e.g., Li and Lee, 2005; Cairns et al., 2011; Dowd et al., 2011; Jarner and Kryger, 2011; Li and Hardy, 2011; Zhou, Li, and Tan, 2013; Zhou et al., 2014). Many innovative contracts have payoffs that are linked to broad-based population mortality indices; hence, a more sophisticated stochastic mortality model for multipopulation is critical for pricing these longevity-linked securities as well as for reducing population basis risk (Blake et al., 2013). Chen, MacMinn, and Sun (2015) introduce factor copula into multipopulation mortality modeling. They employ a two-stage procedure based on the time series analysis and a factor copula approach. Wang, Yang, and Huang (2015) model multicountry mortality using a dynamic copula framework.

We propose to use a new copula family called the Lévy subordinated hierarchical Archimedean copula (LSHAC) model for multipopulation mortality modeling. The LSHAC model has some appealing advantages. First, it overcomes some serious drawbacks of the classical Archimedean copulas (AC) and hierarchical Archimedean copulas (HAC). The AC, although it has the advantage of simplicity, suffers from a fully exchangeable structure. As the HAC model has been proposed to partially overcome the exchangeability by "nesting" two or more ACs with appropriate grouping, its generators must fulfill the *compatible conditions* to ensure that the resulting HAC has a valid multivariate distribution (Joe, 1997; Savu and Trede, 2010; McNeil, 2008). The compatible conditions, however, can be difficult (and empirically almost impossible) to verify, and hence restrict the practical application of the HACs (Savu and Trede, 2010). This difficulty is resolved by Hering et al. (2010) and Mai and Scherer (2012). With a two-layer illustration, they show that as long as the HACs are constructed from Lévy subordinators, the compatible conditions are automatically satisfied. Zhu, Wang, and Tan (2016) provide an estimation methodology for the LSHAC model in a general setting. Motivated by these findings, the goal of this article is to employ the multi-layer LSHAC to model the multipopulation mortality dependence. To the best of our knowledge, this is the first article to model the multipopulation longevity risk with the LSHAC model.

Another important advantage of the LSHAC is its ability to capture the relationship between the geographical locations and mortality dependence using its hierarchical

structure. Other existing copula models, such as elliptical copulas (including Gaussian copula and Student's t copula) and classical AC, do not have such capability. The idea that mortality dependence is associated with the distance between countries makes sense intuitively because mortality rates are related to socioeconomic, climate, and epidemical factors that are spatially correlated. In fact, it is supported by some epidemiological studies. For example, studies have shown that some geographical patterns of infectious diseases such as malaria and cardiovascular disease are at the international level (Mundial, 1993; Elliott and Best, 1998), whereas some others are at the town or county level (Pocock et al., 1980). From a multicountry longevity risk management point of view, it is beneficial, yet challenging, to capture the mortality dependence that is associated with geographical locations. The inherent hierarchical construction of the LSHAC provides a natural way of addressing this relationship. In the empirical analysis, we apply the LSHAC-based mortality model to 13 populations, comprising 12 European countries and 1 very distant country, Australia. Including Australia in our analysis allows us to assess the impact of geographical location on mortality modeling. We show that the LSHAC yields better fit than other copulas that cannot take geographical structure into consideration. In addition, the LSHAC also has more accurate and robust out-of-sample forecasting results. Finally, as an illustrative example, the 13-dimensional LSHAC model is used to price a survivor index swap contract.

The remainder of this article is organized as follows. The "Multicountry Stochastic Mortality Model" section describes in details the modeling of multicountry mortality data using the LSHAC model. The "Empirical Analysis of Multicountry Mortality Data" section applies the LSHAC model to mortality experience from 13 countries. A survival index swap contract is discussed in the "Multicountry Survivor Index Swaps" section. The "Conclusion" section concludes the article.

Multicountry Stochastic Mortality Model

The "CBD Model" subsection first reviews the two factor Cairns–Blake–Dowd (CBD; Cairns, Blake, and Dowd, 2006) model that will be used to analyze the mortality risk dynamic for each individual country. The "Multicountry Mortality Modeling With LSHAC" subsection then describes our proposed multicountry stochastic mortality model based on LSHAC.

CBD Model

A number of extrapolative stochastic mortality models have been developed for mortality modeling (Lee and Carter, 1992; Renshaw and Haberman, 2003; Cairns, Blake, and Dowd, 2006; Cairns et al., 2011). In addition to the advantages of parsimony and transparency, the CBD model is able to provide a nontrivial correlation structure; that is, changes in mortality rates at different ages are not perfectly correlated, providing greater robustness (Cairns et al., 2009). More importantly, the CBD model performs well in forecasting. Finally, the CBD model has no identification problems in estimating and satisfies the new-data-invariant property (Chan, Li, and Li, 2014). Therefore, we apply the CBD model for the marginal distributions.

We use $q^j_{x,t}$ to denote the mortality rates of the jth population at age x and time t, where $j = 1, \ldots, N$ and $t = 1, \ldots, T$. Then the CBD model can be expressed as

$$\log\left(\frac{q^j_{x,t}}{1-q^j_{x,t}}\right) = \kappa^j_{1,t} + \kappa^j_{2,t}(x - \bar{x}), \quad t = 1, \ldots, T, j = 1, \ldots, N, \tag{1}$$

where $\bar{x}$ is the mean age in the sample range. The time-varying parameters $\kappa^j_{1,t}$ and $\kappa^j_{2,t}$, which are period effects for the jth population, have economic interpretations under the CBD assumptions. In particular, $\kappa^j_{1,t}$ represents an overall mortality improvement, and $\kappa^j_{2,t}$ is the steepness of the mortality curve (in logit scale), which indicates the mortality improvement for older ages.

Following CBD, we model $\boldsymbol{\kappa}^j_t = (\kappa^j_{1,t}, \kappa^j_{2,t})'$ as a two-dimensional random walk with drift as follows:

$$\boldsymbol{\kappa}^j_t - \boldsymbol{\kappa}^j_{t-1} = \boldsymbol{\mu}^j + \boldsymbol{C}^j \boldsymbol{Z}^j_t, \quad t = 1, \ldots, T, \; j = 1, \ldots, N, \tag{2}$$

where $\boldsymbol{\mu}^j = (\mu^j_1, \mu^j_2)'$ are the drifts; C^j is a lower-triangular matrix based on Cholesky decomposition, satisfying $\boldsymbol{V}^j = \boldsymbol{C}^j \boldsymbol{C}^{j'}$, where $\boldsymbol{V}^j$ is the covariance matrix of $\boldsymbol{\kappa}^j_t - \boldsymbol{\kappa}^j_{t-1}$; and $\boldsymbol{Z}^j_t = (Z^j_{1,t}, Z^j_{2,t})'$ are independent standard normal innovations. Consequently, $\kappa^j_{1,t}$ is related to $Z^j_{1,t}$ and $\kappa^j_{2,t}$ is related to both $Z^j_{1,t}$ and $Z^j_{2,t}$.

Multicountry Mortality Modeling With LSHAC

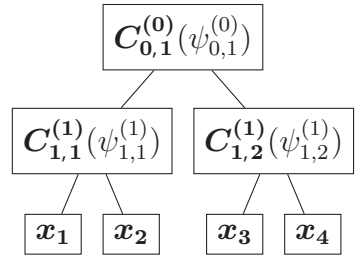
Recall that a N-dimensional AC, $C : [0,1]^N \to [0,1]$, can be defined as (Nelsen, 2006)

$$C(u_1, u_2, \ldots, u_N) = \psi\big(\psi^{-1}(u_1) + , ..., \psi^{-1}(u_N)\big), \tag{3}$$

where $\psi \in \mathcal{G} = \{\psi : [0,\infty) \to [0,1] \mid \psi_{\lim u \to \infty}(u) = 0, \psi(0) = 1, (-1)^k \frac{d^k}{du^k}\psi(u) \geq 0, k \in \mathbb{N}\}$, is called the *completely monotonic* (c.m.) generator, and ψ^{-1} is its inverse, defined as $\psi^{-1}(u) = \inf\{t : \psi(t) \leq u\}$. This copula has been widely used in risk management applications due to its parsimonious framework with a small number of parameters. One of its main drawbacks is the exchangeable structure; that is, the distribution of $(u_1, u_2, \ldots, u_d)'$ in (3) is invariant of permutation, which severely restricts the modeling capability of the AC models. In an attempt of overcoming the exchangeability, the HAC has been proposed for high-dimensional dependence modeling (Savu and Trede, 2010; McNeil, 2008). Instead of using a single generator, ψ, as in the traditional construction of AC in (3), the HAC partitions all the random variables into a hierarchical structure with various levels and subgroups.

The impact of hierarchical structure on capturing the dependence between variables is best illustrated with an example as shown in Figure 1. The hierarchical structure is used to model a four-dimensional variables $(\boldsymbol{x_1}, \boldsymbol{x_2}, \boldsymbol{x_3}, \boldsymbol{x_4})$ using a two level HAC with three generators. The bottommost level shows that the four variables are subdivided into two subgroups of $\{\boldsymbol{x_1}, \boldsymbol{x_2}\}$ and $\{\boldsymbol{x_3}, \boldsymbol{x_4}\}$. The dependence within the variables for the

FIGURE 1
A Four-Dimension HAC Example



two subgroups is induced by the inner copulas $\boldsymbol{C_{1,1}^{(1)}}$ and $\boldsymbol{C_{1,2}^{(1)}}$ with respective generators $\psi_{1,1}^{(1)}$ and $\psi_{1,2}^{(1)}$ at level 1. These two copulas, in turn, are nested together via the outer copula $\boldsymbol{C_{0,1}^{(0)}}$ with generator $\psi_{0,1}^{(0)}$ at level 0. For detailed notation system of a general L-level LSHAC, see Zhu, Wang, and Tan (2016). The resulting four-dimensional copula function is succinctly represented by

$$
\begin{aligned}
&C(\boldsymbol{x_1}, \boldsymbol{x_2}, \boldsymbol{x_3}, \boldsymbol{x_4}) \\
&\quad = C_{0,1}^{(0)}\left(C_{1,1}^{(1)}(\boldsymbol{x_1}, \boldsymbol{x_2}), C_{1,2}^{(1)}(\boldsymbol{x_3}, \boldsymbol{x_4})\right) \qquad (4) \\
&\quad = \psi_{0,1}^{(0)}\left\{\psi_{0,1}^{-1(0)}\left(\psi_{1,1}^{(1)}(\psi_{1,1}^{-1(1)}(\boldsymbol{x_1}) + \psi_{1,1}^{-1(1)}(\boldsymbol{x_2})) + \psi_{1,2}^{(1)}(\psi_{1,2}^{-1(1)}(\boldsymbol{x_3}) + \psi_{1,2}^{-1(1)})(\boldsymbol{x_4})\right)\right\}. \qquad (5)
\end{aligned}
$$

The determination of the hierarchical structure of the LSHAC is based on the hierarchical clustering analysis algorithm.[1] In particular, we follow the technique proposed in Zhu, Wang, and Tan (2016) where a new τ-Euclidean distance measure is used to determine the structure. An extensive analysis was conducted in Zhu, Wang, and Tan (2016) to show that their proposed clustering-based approach of determining an optimal hierarchical structure is reliable and robust.

In view of Equations (4), (5), and Figure 1, the HAC model constructs copula in such a way that different generators can be applied to different subgroups, and hence elements between subgroups are no longer exchangeable. As the example above indicates that hierarchial structure of a HAC can be created with mixing outer and innner copulas. For an HAC that yields a valid copula function, the copulas and their corresponding generators cannot be selected arbitrarily. They need to satisfy the so-called *compatible conditions*. In our preceding example, the compatible conditions are

$$
\psi_{0,1}^{(0)}, \psi_{1,1}^{(1)}, \psi_{1,2}^{(1)} \in \mathcal{G} \qquad \text{and} \qquad (\psi_{0,1}^{(0)-1} \circ \psi_{1,j}^{(1)})' \in \mathcal{G}, j = 1, 2,
$$

where $\circ$ represents a functional composition. In general, verifying the compatible conditions on a case-by-case basis is very tedious and mostly impossible in empirical

[1]For an introduction of hierarchical clustering analysis, refer to Ward (1963), Jain and Dubes (1988), and Zhang, Zhao, and Wang (2013).

TABLE 1
AC Generator Functions and Lévy Subordinators

AC Family	$\psi(u)$	λ_u	λ_l	Parameters
Gumbel (GM)	$\psi_{\mathrm{GM}}(u) = \exp\left(-u^{\frac{1}{\theta}}\right)$	$2-2^{\frac{1}{\theta}}$	0	$\theta \geq 1$
Clayton (CL)	$\psi_{\mathrm{CL}}(u) = (1+u)^{-\frac{1}{\theta}}$	0	$2^{-\frac{1}{\theta}}$	$\theta > 0$
Joe (JO)	$\psi_{\mathrm{JO}}(u) = 1-(1-e^{-u})^{1/\theta}$	$2-2^{\frac{1}{\theta}}$	0	$\theta \geq 1$
Subordinators	$\boldsymbol{\Psi(u)}$		Parameters	
Stable process (GM)	$\Psi_{\mathrm{GM}}(u) = u^a$		$0 < a < 1$	
Inverse Gaussian (IG)	$\Psi_{\mathrm{IG}}(u) = a\sqrt{2u+b^2} - ab$		$a > 0, b > 0$	

applications (Savu and Trede, 2010; Hering et al., 2010). It has been shown that when the generator functions are all of the type Gumbel (denoted as All-Gumbel-HAC) or Clayton (All-Clayton-HAC), the compatible conditions are satisfied (Embrechts, Lindskog, and McNeil, 2003). For this reason, to date empirical analysis of the HAC has been confined mostly to either All-Gumbel-HAC or All-Clayton-HAC. This severely restricts the application potential of HAC.

Hering et al. (2010) circumvent this hard-to-check compatible conditions by constructing the HAC models via Lévy subordinators. Mai and Scherer (2012) extend the LSHAC model to a h-extendible framework. Let $\{S_t : 0 \leq t \leq T\}$ be a Lévy subordinator, that is, a stochastically continuous nondecreasing Lévy process, which has zero start, stationary and independent increments (Proposition 3.10, Tankov, 2004). The Laplace–Stieltjes transform of S_t satisfies $E(e^{-\omega S_t}) = \exp\left(-t\Psi_S(\omega)\right), \forall \omega > 0$, where the nondecreasing function, $\Psi_S : [0,\infty) \to [0,\infty)$, is called the *Laplace exponent* of the Lévy subordinator, S_t. Table 1 lists examples of AC and Lévy subordinators, together with the coefficients of upper (λ_u) and lower (λ_l) tail dependence.

We now state the following two key propositions.

Proposition 1: *(Hering et al., 2010; Zhu, Wang, and Tan, 2016) For a L-level LSHAC, at level l, the j_l-th copula generator in position s_{l-1}, $\psi^{(l)}_{s_{l-1},j_l}$, can be expressed as:*

$$\psi^{(l)}_{s_{l-1},j_l} = \psi^{(0)}_{0,1} \bigodot_{i=1}^{l-1} \Psi^{(i+1)}_{s_i,j_{i+1}}, \tag{6}$$

where $\bigodot_{i=1}^{n} f_i := f_1 \circ \ldots \circ f_n$, $\Psi^{(i+1)}_{s_i,j_{i+1}}$ is the corresponding Laplace exponent. In addition, $\psi^{(l)}_{s_{l-1},j_l}, l = 1,\ldots,L$, satisfy compatible conditions.

Proposition 1 states that at each level of a LSHAC, generators can be constructed from composing an outer AC generator and a sequence of Laplace exponents of Lévy subordinators. This proposition asserts that as long as the copula generators are constructed from Lévy subordinators, the compatible conditions are automatically satisfied. This

dramatically expands the family of LSHAC. The following proposition assures that the most commonly studied HAC, that is, All-Gumbel-HAC, is a special case of LSHAC:

Proposition 2: *(Mai and Scherer, 2012; Zhu, Wang, and Tan, 2016) For an All-Gumbel-HAC, the l-th level copula generator $\psi^{(l)}(u)$ can be expressed as ($l \geq 1$):*

$$\psi^{(l)}(u) = \psi^{(0)} \bigodot_{k=1}^{l-1} \widetilde{\Psi}^{(k)}(u) = \exp\big(-u^{\prod_{k=0}^{l-1} \frac{1}{\theta_k}}\big).$$

From the parameterization in Table 1, $\psi^{(0)}$ represents a GM generator with $\theta = \theta_0, \theta_0 > 1$; $\widetilde{\Psi}^{(k)}$ denotes the k-th GM subordinator with $a = 1/\theta_k, \theta_k > 1$.

EMPIRICAL ANALYSIS OF MULTICOUNTRY MORTALITY DATA

In this section, we analyze the longevity risk among 13 countries based on the statistical framework introduced in preceding section. In particular, marginal and copula parameters are estimated with the inference for margin (IFM) procedure using the maximum likelihood (ML) method, which has been shown to yield asymptotically efficient estimators (Joe, 1997; Patton, 2006). The data set used in this study is described in the "Data" subsection. In the "Hierarchical Dependence Structure" subsection, the CBD model and the LSHAC model are applied to the marginal dynamics and the dependence structure of the data, respectively.

Data

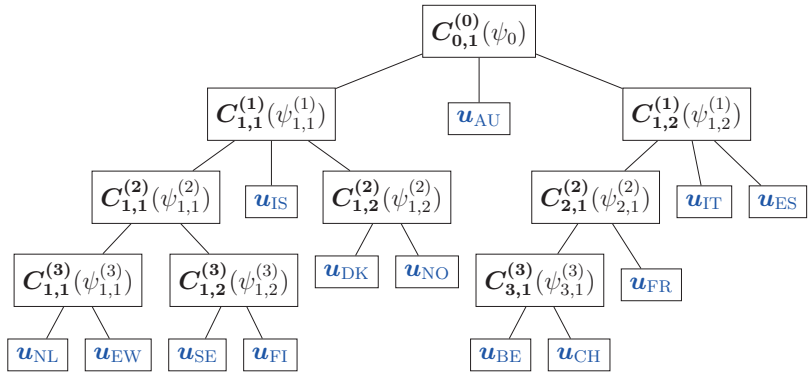
We use total population mortality rate data of 1-year cohorts for higher ages (from 65 to 90 inclusive). The data are obtained from the Human Mortality Database (HMD; see http://www.mortality.org/). The data set contains 13 countries, including Australia (AU) and 12 European countries. The earliest available Australian data starts from 1921; hence, the data set covers the period from 1921 to 2009. The 12 European countries are Belgium (BE), Netherlands (NL), England and Wales (EW), Denmark (DK), Norway (NO), Finland (FI), Sweden (SE), France (FR), Italy (IT), Spain (ES), Switzerland (CH), and Iceland (IS), which include all the countries in Europe with available data over the same period.

Hierarchical Dependence Structure

In this subsection, we first estimate the marginal dynamic of the mortality rates for each country based on the CBD model[2] and then obtain the pseudo sample by probability transform to the innovations from Equation (2), $\boldsymbol{Z}^j = (Z_1^j, Z_2^j)'$, namely, $\boldsymbol{u_k} = (u_k^1, \ldots, u_k^N)' = (\Phi(Z_k^1), \ldots, \Phi(Z_k^N))$, where $k = 1, 2$, and $\Phi(\cdot)$ represents the standard normal cumulative distribution function.

[2]The $\kappa_{1,t}$ and $\kappa_{2,t}$ estimates of the CBD model for the 13 countries are included in the Online Appendix (Zhu, Tan, and Wang, 2017).

FIGURE 2
Hierarchical Structure of 13 Countries for $\boldsymbol{u_1}$



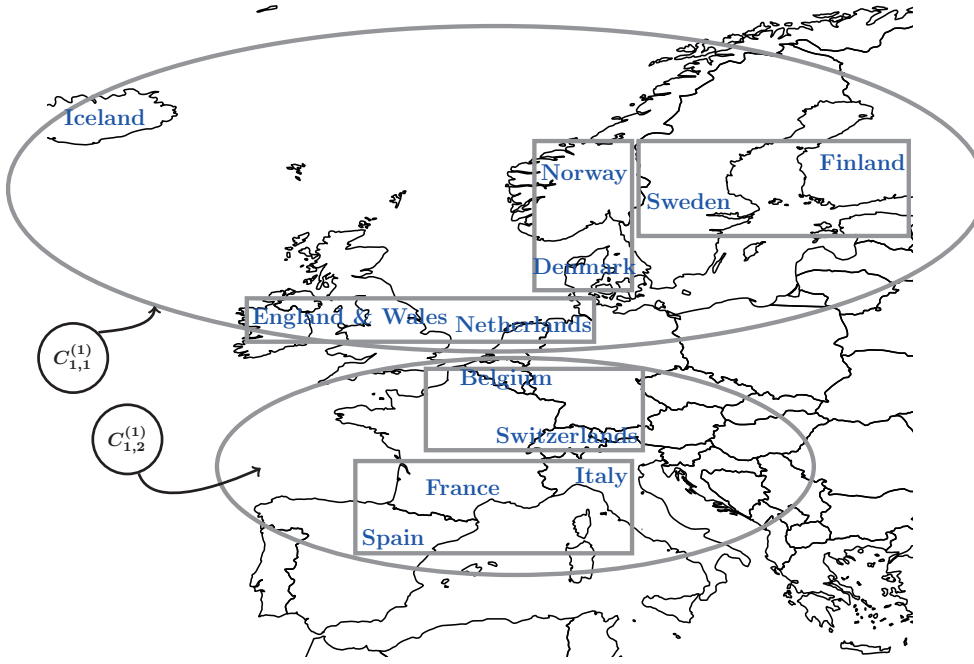
Note: Respective copula generators are in the parentheses. The 13 countries include Australia (AU), Belgium (BE), Denmark (DK), England and Wales (EW), Finland (FI), France (FR), Iceland (IS), Italy (IT), Netherlands (NL), Norway (NO), Spain (ES), Sweden (SE), and Switzerland (CH).

We then determine the underlying hierarchical structure associated with LSHAC using hierarchical clustering analysis method. Figure 2 displays the resulting hierarchical structure of $\boldsymbol{u_1}$. In this structure, the Netherlands, and England and Wales are grouped together under $\boldsymbol{C^{(3)}_{1,1}}$, whereas Sweden and Finland are grouped together under $\boldsymbol{C^{(3)}_{1,2}}$. The two groups are then linked by $\boldsymbol{C^{(2)}_{1,1}}$. In addition, Denmark and Norway are clustered together under $\boldsymbol{C^{(2)}_{1,2}}$. Joining Iceland, the three subgroups are then nested into $\boldsymbol{C^{(1)}_{1,1}}$. The other five countries are grouped into $\boldsymbol{C^{(1)}_{1,2}}$, with Belgium and Switzerland first clustered together at the bottommost level under $\boldsymbol{C^{(3)}_{3,1}}$, then joined by France and finally grouped together with Italy and Spain. At level 0, the most distant country, Australia, is isolated by itself but is nested into the outer generator $\boldsymbol{C^{(0)}_{0,1}}$, together with the other two subgroups.

Recall that $\boldsymbol{u_1}$ is transformed from $\kappa_{1,t}$, the index of overall mortality improvement. Hence, the LSHAC structure depicted in Figure 2 identifies significant geographical effect in overall mortality rates. In particular, Denmark, Finland, Norway, Sweden, and Iceland, the countries share similar history, language, social structure, and so on, belong to the first subgroup under $\boldsymbol{C^{(1)}_{1,1}}$. Meanwhile, countries in central and western Europe, that is, France, Belgium, and Switzerland, are first nested together and then joined with two southern European countries, Spain and Italy. Figure 3 visually illustrates the effect of the grouping. We also apply the same hierarchical clustering method to $\boldsymbol{u_2}$, but it was found that there is no significant geographical clustering effect.[3] The reason that the geographical effect is not as significant for the hierarchical

[3]The grouping and estimating results of $\boldsymbol{u_2}$ with the LSHAC are available in the Online Appendix (Zhu, Tan, and Wang, 2017).

FIGURE 3
Illustrative Grouping Results of 13 Countries



Note: Australia is isolated into an independent subgroup and is not displayed in this figure.

structure of $\boldsymbol{u}_2$ as for that of $\boldsymbol{u}_1$ is because $Z_{2,t}$ is orthogonal to $Z_{1,t}$ by the construction of Cholesky decomposition; therefore, it is reasonable to find that $\boldsymbol{u}_2$ does not show geographical effect. However, because $\kappa_{2,t}^j$ is a linear combination of both $Z_{1,t}$ and $Z_{2,t}$, $\kappa_{2,t}^j$ is also impacted by the geographical mortality dependence.

Statistical Validation of LSHAC

By using Bayesian information criterion (BIC), this section compares the relative efficiency of the LSHAC model proposed in the preceding subsection to other well-known copulas. The structure associates with Figure 2 is a 13-dimensional LSHAC model with nine AC generators. As there are three possible choices of AC generators (see Table 1), this implies that there are $3 \times 2^8 = 768$ candidate models. Rather than exhausting all these combinations, we first assume that the Lévy subordinators at each level of the hierarchical structure are identical so that the number of candidate models we need to consider is reduced substantially. In our example, this leads to $3 \times 2^3 = 24$ possibilities. Once we have identified the best fit LSHAC family under this assumption, we then find the best fit LSHAC by going through all the possible combinations within this family.

The estimation results for the 24 LSHAC candidate models are summarized in the lower panel of Table 2. The first four columns show the possible combinations of

TABLE 2
Estimation Results of the 13 Populations With Various Copula Models

Copula Models	No.Para	BIC
Gaussian copula	78	−165.44
Student's *t* copula	79	−206.82
AC (GM)	1	−129.49
AC (CL)	1	−84.43
AC (JO)	1	−91.87

LSHAC Model

Level 0	Level 1	Level 2	Level 3	No.Para	BIC	BIC.Imp(G)	BIC.Imp(T)	BIC.Imp(AC)
GM	GM	GM	GM	9	−198.57	+33.13	−8.25	+69.08
GM	GM	GM	IG	12	−191.05	+25.62	−15.76	+61.57
GM	GM	IG	GM	13	−197.30	+31.86	−9.51	+67.81
GM	GM	IG	IG	15	−189.16	+23.72	−17.65	+59.67
GM	IG	GM	GM	12	−207.14	+41.70	+0.32	+77.65
GM	IG	GM	IG	14	−194.63	+29.19	−12.19	+65.14
GM	IG	IG	GM	15	−189.05	+23.61	−17.77	+59.56
GM	IG	IG	IG	17	−179.32	+13.88	−27.50	+49.83
CL	GM	GM	GM	9	−195.14	+29.69	−11.68	+65.64
CL	GM	GM	IG	12	−181.08	+15.64	−25.73	+51.59
CL	GM	IG	GM	13	−192.65	+27.21	−14.17	+63.16
CL	GM	IG	IG	15	−173.89	+8.45	−32.93	+44.40
CL	IG	GM	GM	12	−187.86	+22.42	−18.96	+58.37
CL	IG	GM	IG	14	−173.33	+7.89	−33.49	+43.84
CL	IG	IG	GM	15	−162.71	−2.73	−44.11	+33.22
CL	IG	IG	IG	17	−136.95	−28.49	−69.87	+7.46
JO	GM	GM	GM	9	−190.42	+24.98	−16.40	+60.93
JO	GM	GM	IG	12	−183.57	+18.13	−23.25	+54.08
JO	GM	IG	GM	13	−189.73	+24.29	−17.09	+60.24
JO	GM	IG	IG	15	−178.45	+13.01	−28.37	+48.96
JO	IG	GM	GM	12	−207.71	+42.27	+0.89	+78.21
JO	IG	GM	IG	14	−193.62	+28.18	−12.20	+64.13
JO	IG	IG	GM	15	−190.23	+24.79	−16.59	+60.74
JO	IG	IG	IG	17	−179.72	+14.28	−27.10	+50.23

Note: The first panel lists the results for Gaussian copula, Student's *t* copula, and Archimedean copulas (AC) with GM, CL, and JO generators. The second panel shows the results for the LSHAC models. The first four columns show the AC generators/Lévy subordinator in the LSHAC structures. The fifth column shows the number of parameters in the corresponding LSHAC model. The sixth column shows the BIC values (BIC). The last three columns show the BIC improvements of the LSHAC models compared to Gaussian copula (BIC.Imp(G)), Student's *t* copula (BIC.Imp(T)), and the best AC model (BIC.Imp(AC)), with a positive sign indicating a BIC improvement.

AC generators and Lévy subordinators. The fifth column tabulates the number of parameters for the corresponding LSHAC. The sixth column gives the BIC values. To assess the relative efficiency of LSHAC, the upper panel of the same table displays

TABLE 3
Estimation Results of the Best JO-LSHAC Model

θ	$a_{1,1}^{(1)}$	$a_{1,2}^{(1)}$	$b_{1,2}^{(1)}$	$a_{1,1}^{(2)}$	$a_{1,2}^{(2)}$	$a_{2,1}^{(2)}$	$a_{1,1}^{(3)}$	$a_{1,2}^{(3)}$	$a_{3,1}^{(3)}$
1.059***	0.791***	0.790***	0.281***	0.894***	0.747***	0.792***	0.722***	0.931***	0.983***
(0.041)	(0.047)	(0.039)	(0.085)	(0.062)	(0.098)	(0.044)	(0.052)	(0.079)	(0.089)

Note: Standard errors of the estimated parameters are presented in parentheses. *** indicates that the correlations are significant at the 1 percent level.

the estimation results for Gaussian copula, Student's t copula, and the AC models with GM, CL, and JO generators. Note that the number of parameters of Gaussian and Student's t copulas increase quadratically with the dimension. In contrast, the number of parameters for the LSHAC increases only linearly. Although the ACs have the advantage of simplicity with only one parameter, they are clearly ineffective in the sense of their BIC values. The last three columns of the lower panel summarize the BIC improvements of the LSHAC models relative to Gaussian copula, Student's t copula, and the best AC model. The positive "improvement" values indicate the preference of the LSHAC model: all of the 24 LSHAC models outperform ACs, all but 2 prefer the LSHAC to Gaussian copula, and, 2 of them are better than Student's t copula. The LSHAC models present some promising competitive alternatives, providing a better trade-off between the model parsimony and good fitness.

We now focus on the LSHAC models with Joe as the outer generator (hereafter, we call JO-LSHAC for short), as this family of LSHAC yields the best estimation results. Given Joe as the outer generator, we go through all 256 (i.e., 2^8) possible combinations and then select the optimal one with the smallest BIC value.[4] Equation (7) shows the 13-dimensional copula structure, and the corresponding nine AC generators are displayed in Equations (8)–(16). The subscripts in the generator functions denote the outer generator and Lévy subordinators. For example, $\psi_{\mathrm{JO}\circ\mathrm{GM}}$ is a generator constructed by a JO outer generator and a GM Lévy subordinator. Table 3 displays the estimated parameters and standard errors of the best JO-LSHAC model. We can see that all the parameters are significant at the 1 percent level. The best JO-LSHAC model achieves a BIC value of -209.83, leading to BIC improvements of $+44.39$ and $+3.01$, compared to Gaussian copula and Student's t copula, respectively.

$$C(\boldsymbol{u_1},\ldots,\boldsymbol{u_{13}}) = C_{0,1}^{(0)}\Big(C_{1,1}^{(1)}\big(C_{1,1}^{(2)}(C_{1,1}^{(3)}(\boldsymbol{u_{NL}},\boldsymbol{u_{EW}}),C_{1,2}^{(3)}(\boldsymbol{u_{SE}},\boldsymbol{u_{FI}})),\\ C_{1,2}^{(2)}(\boldsymbol{u_{DK}},\boldsymbol{u_{NO}}),\boldsymbol{u_{IS}}\big),C_{1,2}^{(1)}\big(C_{2,1}^{(2)}(C_{3,1}^{(3)}(\boldsymbol{u_{BE}},\boldsymbol{u_{CH}}),\boldsymbol{u_{FR}}),\boldsymbol{u_{IT}},\boldsymbol{u_{ES}}\big),\boldsymbol{u_{AU}}\Big). \tag{7}$$

$$\psi_{0,1}^{(0)}(u) = \psi_{\mathrm{JO}}(u) = 1-\big(1-\exp(-u)\big)^{\frac{1}{\theta}}, \tag{8}$$

[4]Due to length restriction, we only show the estimation results for the best JO-LSHAC model. Results for the other JO-LSHAC family are included in the Online Appendix (Zhu, Tan, and Wang, 2017).

$$\psi_{1,1}^{(1)}(u) = \psi_{\mathrm{JO}\circ\mathrm{GM}}(u) = 1 - \left(1 - \exp(-u^{a_{1,1}^{(1)}})\right)^{\frac{1}{\theta}}, \tag{9}$$

$$\psi_{1,1}^{(2)}(u) = \psi_{\mathrm{JO}\circ\mathrm{GM}\circ\mathrm{GM}}(u) = 1 - \left(1 - \exp(-u^{a_{1,1}^{(1)}a_{1,1}^{(2)}})\right)^{\frac{1}{\theta}}, \tag{10}$$

$$\psi_{1,2}^{(2)}(u) = \psi_{\mathrm{JO}\circ\mathrm{GM}\circ\mathrm{GM}}(u) = 1 - \left(1 - \exp(-u^{a_{1,1}^{(1)}a_{1,2}^{(2)}})\right)^{\frac{1}{\theta}}, \tag{11}$$

$$\psi_{1,1}^{(3)}(u) = \psi_{\mathrm{JO}\circ\mathrm{GM}\circ\mathrm{GM}\circ\mathrm{GM}}(u) = 1 - \left(1 - \exp(-u^{a_{1,1}^{(1)}a_{1,1}^{(2)}a_{1,1}^{(3)}})\right)^{\frac{1}{\theta}}, \tag{12}$$

$$\psi_{1,2}^{(3)}(u) = \psi_{\mathrm{JO}\circ\mathrm{GM}\circ\mathrm{GM}\circ\mathrm{GM}}(u) = 1 - \left(1 - \exp(-u^{a_{1,1}^{(1)}a_{1,1}^{(2)}a_{1,2}^{(3)}})\right)^{\frac{1}{\theta}}, \tag{13}$$

$$\psi_{1,2}^{(1)}(u) = \psi_{\mathrm{JO}\circ\mathrm{IG}}(u) = 1 - \left(1 - \exp(-a_{1,2}^{(1)}\sqrt{2u + (b_{1,2}^{(1)})^2} + a_{1,2}^{(1)}b_{1,2}^{(1)})\right)^{\frac{1}{\theta}}, \tag{14}$$

$$\psi_{2,1}^{(2)}(u) = \psi_{\mathrm{JO}\circ\mathrm{IG}\circ\mathrm{GM}}(u) = 1 - \left(1 - \exp(-a_{1,2}^{(1)}\sqrt{2u^{a_{2,1}^{(2)}} + (b_{1,2}^{(1)})^2} + a_{1,2}^{(1)}b_{1,2}^{(1)})\right)^{\frac{1}{\theta}}, \tag{15}$$

$$\psi_{3,1}^{(3)}(u) = \psi_{\mathrm{JO}\circ\mathrm{IG}\circ\mathrm{GM}\circ\mathrm{GM}}(u) = 1 - \left(1 - \exp(-a_{1,2}^{(1)}\sqrt{2u^{a_{2,1}^{(2)}a_{3,1}^{(3)}} + (b_{1,2}^{(1)})^2} + a_{1,2}^{(1)}b_{1,2}^{(1)})\right)^{\frac{1}{\theta}}. \tag{16}$$

MULTICOUNTRY SURVIVOR INDEX SWAPS

In this section, we demonstrate the benefit of capturing the geographical structure with the LSHAC-based multicountry mortality model by applying it to pricing and hedging a survivor index swap. Such derivative has been studied extensively in Blake et al. (2013), Dawson et al. (2010), Dowd et al. (2006), and Wang, Huang, and Liu (2013). We consider a survival index for a cohort of $x_0 = 65$, which follows the framework of Wang, Yang, and Huang (2015), with the underlying survivor index based on multi-population survivor distributions in order to reduce basis risk.[5] As the survival index swap under examination is a hypothetical product with no market price observations, we first perform a backtesting, which will further verify the geographical structure from an out-of-sample forecasting perspective. We examine the best JO-LSHAC model (Table 3). For competing models, in addition to the Student's t copula, we also consider the factor copula model proposed by Chen, MacMinn, and Sun (2015).[6]

We perform the backtesting using the following three steps:

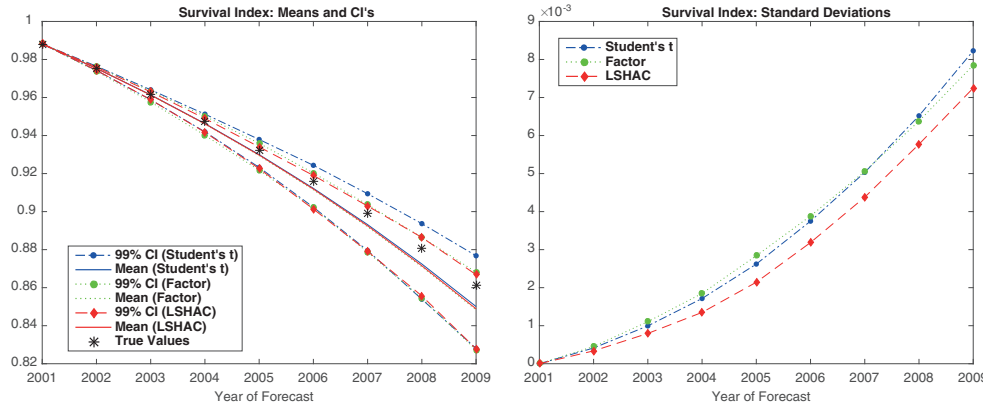
1. models are calibrated using in-sample data from 1921 to 2000;
2. models are projected to 2001–2009, with 10,000 simulation trials;
3. true values of the survival index are calculated, using out-of-sample data in 2001–2009, and compared to each copula model.

The results are displayed in Figure 4. The left plot shows the means and confidence intervals (CIs) for the projected values of survival index. The mean values estimated by

[5]Details about the survival index and the corresponding swap are included in the Online Appendix.

[6]A brief introduction of one-factor copula model and the estimation results of $\boldsymbol{u_1}$ and $\boldsymbol{u_2}$ are displayed in Section E of the Online Appendix (Zhu, Tan, and Wang, 2017).

FIGURE 4
Backtesting Results for the LSHAC Model, Factor Copula, and Student's t Copula



Note: The first plot shows the means and confidence intervals (CIs) for projected values of survival index. Stars are the true values of survival index calculated with out-of-sample data. The second plot compares the standard deviations of the projections. For both plots, curves with bullets are results for Student's t copula, curves with squares are results with the factor copula, and curves with diamonds are for the LSHAC. The complete sample period is from 1921 to 2009, with in-sample data period 1921–2000 and out-of-sample data period 2001–2009.

different models are all very close to the true values of the survival index. The LSHAC model and the factor copula model have very similar CIs with much narrower band compared to that from the Student's t copula. This indicates that although the three models are all very close in mean estimation, both the LSHAC model and the factor model have greater precision in the sense of smaller CIs. The standard deviations of the projections are compared in the right plot of Figure 4. For earlier years (2001–2007), the factor copula has slightly larger standard deviations than the Student's t copula, whereas after 2007 the opposite is true. In contrast, the LSHAC model has the lowest standard deviations of the projection over the full out-of-sample forecasting period (2001–2009).

Now, we examine the pricing and hedging implications of our proposed models by considering a survival index swap with maturity $T = 25$ years in the calendar year 2009. Table 4 compares the swap premiums under different yield rate assumptions and different levels of market price of risk $\lambda \in \{-0.1, -0.15, -0.2\}$. Table 5 reports the VaR and CTE values of the losses for different maturity times by assuming $\lambda = -0.1$. We draw the following observations based on results in Tables 4 and 5.

1. The swap premiums are clearly sensitive to the assumed market price of risk. Indeed, the more negative the market price of risk, the higher the swap premium.
2. The swap premiums are also sensitive to the assumed yield rate. The higher the interest rate, the lower the swap rate.

TABLE 4
Swap Premiums (in Basis Points) for Different Yield Rate and Market Price of Risk

	Original Yield Curve			Parallel Shift up 2%			Parallel Shift up 4%		
	LSHAC	Student's t	Factor	LSHAC	Student's t	Factor	LSHAC	Student's t	Factor
$\lambda = -0.1$	101.15	101.14	100.98	167.50	167.44	167.18	270.40	270.10	269.81
$\lambda = -0.15$	106.90	106.88	106.73	177.93	177.87	177.62	286.88	286.59	286.30
$\lambda = -0.2$	112.61	112.45	112.45	188.31	188.24	187.99	303.29	303.29	302.71

Note: The results are for time to maturity $T = 25$. Pricing results are based on the LSHAC model with the best performance in the "Hierarchical Dependence Structure" section (Table 3), factor copula, and the Student's t copula.

3. Dependence modeling is important, and different dependence models could result in diverse results. In general, the computed swap premiums are very similar among the three models, although we observe that for lower market price of risk, the pricing results of the Student's t copula are closer to the LSHAC model, whereas for higher market price of risk, the factor copula and the LSHAC model have similar pricing results. On the other hand, the estimated risk measures from the LSHAC model are smaller than the other two copula models and thus indicate that the choice of dependence model has a more pronounced impact on estimating tail risk.

CONCLUSION

In this article, we introduce the LSHAC model for modeling the mortality dependence across multiple countries to facilitate longevity-linked security pricing and longevity risk hedging. Our empirical analysis indicates that our proposed LSHAC model, which has the capability of capturing the geographical structure of the data, has better goodness of fit than the elliptical copulas. Additionally, the LSHAC model produces out-of-sample forecasts with smaller standard deviations, when compared to other benchmark copula models. In particular, our empirical results verify that

TABLE 5
Risk Measures, Including VaR and CTE, of the Losses for Different Maturity Times

	$T = 15$			$T = 20$			$T = 25$		
	LSHAC	Student's t	Factor	LSHAC	Student's t	Factor	LSHAC	Student's t	Factor
$\text{VaR}_{0.95}$	0.090	0.121	0.121	0.194	0.262	0.253	0.332	0.449	0.431
$\text{VaR}_{0.99}$	0.126	0.171	0.173	0.274	0.369	0.365	0.468	0.635	0.621
$\text{CTE}_{0.95}$	0.112	0.152	0.154	0.242	0.328	0.321	0.414	0.563	0.546
$\text{CTE}_{0.99}$	0.142	0.196	0.200	0.308	0.422	0.417	0.530	0.725	0.711

Note: The results are for market price of risk $\lambda = -0.1$. The results are based on the LSHAC model with the best performance in the "Hierarchical Dependence Structure" section (Table 3), factor copula, and Student's t copula.

geographical locations of countries are associated with the overall mortality improvement levels. The survivor index swaps pricing example demonstrates that it is critical and necessary to appropriately model the dependence structure of the mortality risk across different countries to reduce population basis risk.

As the focus of this research is to improve the dependence modeling, we apply the well-documented CBD model, which assumes normal innovations. It might be interesting in future research to investigate a CBD model with non-Gaussian residuals. Another promising future research direction is to design new mortality indexes by including geographical location as a factor. Additionally, the LSHAC model discussed in this article has constant parameters, which will lead to a static dependence structure. The benefit of the time-varying LSHAC model has not been discussed in this article but certainly deserves future research, especially when we are interested in the evolution of the dependence structure.

References

Blake, D., A. Cairns, G. Coughlan, K. Dowd, and R. MacMinn, 2013, The New Life Market, *Journal of Risk and Insurance*, 80(3): 501-558.

Cairns, A. J. G., D. Blake, and K. Dowd, 2006, A Two-Factor Model for Stochastic Mortality With Parameter Uncertainty: Theory and Calibration, *Journal of Risk and Insurance*, 73(4): 687-718.

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(1): 1-35.

Cairns, A. J. G., D. Blake, K. Dowd, G. D. Coughlan, and M. Khalaf-Allah, 2011, Bayesian Stochastic Mortality Modelling for Two Populations, *ASTIN Bulletin*, 41(1): 29-59.

Chan, W.-S., J. S.-H. Li, and J. Li, 2014, The CBD Mortality Indexes: Modeling and Applications, *North American Actuarial Journal*, 18(1): 38-58.

Chen, H., R. MacMinn, and T. Sun, 2015, Multi-Population Mortality Model: A Factor Copula Approach, *Insurance: Mathematics and Economics*, 63: 135-146.

Dawson, P., K. Dowd, A. J. G. Cairns, and D. Blake, 2010, Survivor Derivatives: A Consistent Pricing Framework, *Journal of Risk and Insurance*, 77(3): 579-596.

Dowd, K., D. Blake, A. J. Cairns, and P. Dawson, 2006, Survivor Swaps, *Journal of Risk and Insurance*, 73(1): 1-17.

Dowd, K., A. J. G. Cairns, D. Blake, G. D. Coughlan, and M. Khalaf-Allah, 2011, A Gravity Model of Mortality Rates for Two Related Populations, *North American Actuarial Journal*, 15(2): 334-356.

Elliott, P., and N. Best, 1998, *Geographic Patterns of Disease, Encyclopedia of Biostatistics* (Hoboken, NJ: Wiley).

Embrechts, P., F. Lindskog, and A. McNeil, 2003, Modelling Dependence With Copulas and Applications to Risk Management, in: S. Rachev, ed., *Handbook of Heavy Tailed Distributions in Finance* (Amsterdam: Elsevier, North-Holland), pp. 329-384.

Hering, C., M. Hofert, J.-F. Mai, and M. Scherer, 2010, Constructing Hierarchical Archimedean Copulas With Lévy Subordinators, *Journal of Multivariate Analysis*, 101(6): 1428-1433.

Jain, A. K., and R. C. Dubes, 1988, Algorithms for Clustering Data (Upper Saddle River, NJ: Prentice Hall).

Jarner, S. F., and E. M. Kryger, 2011, Modelling Adult Mortality in Small Populations: The SAINT Model, *ASTIN Bulletin*, 41(2): 377-418.

Joe, H., 1997, *Multivariate Models and Dependence Concepts* (London: Chapman & Hall/CRC).

Lee, R. D., and L. R. Carter, 1992, Modeling and Forecasting U.S. Mortality, *Journal of the American Statistical Association*, 87(419): 659-671.

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 Population: An Extension to the Classical Lee-Carter Approach, *Demography* 42(3): 575-594.

Mai, J.-F., and M. Scherer, 2012, H-Extendible Copulas, *Journal of Multivariate Analysis*, 110, 151-160.

McNeil, A. J., 2008, Sampling Nested Archimedean Copulas, *Journal of Statistical Computation and Simulation*, 78(6): 567-581.

Mundial, B., 1993, *World Development Report 1993; Investing in Health* (Oxford: Oxford University Press).

Nelsen, R. B., 2006, *An Introduction to Copulas*, 2nd edition (New York: Springer).

Patton, A. J., 2006, Modelling Asymmetric Exchange Rate Dependence, *International Economic Review*, 47(2): 527-556.

Pocock, S. J., A. G. Shaper, D. G. Cook, R. F. Packham, R. F. Lacey, P. Powell, and P. F. Russell, 1980, The British Regional Heart Study: Geographical Variations in Cardiovascular Mortality and the Role of Water Quality, *British Medical Journal*, 280(6226): 1243-1249.

Renshaw, A. E., and S. Haberman, 2003, Lee-Carter Mortality Forecasting With Age-Specific Enhancement, *Insurance: Mathematics and Economics*, 33(2): 255-272.

Savu, C., and M. Trede, 2010, Hierarchies of Archimedean Copulas, *Quantitative Finance*, 10(3): 295-304.

Tankov, P., 2004. *Financial Modelling With Jump Process* (Boca Raton, FL: CRC Press).

Wang, C.-W., H.-C. Huang, and I.-C. Liu, 2013, Mortality Modeling With Non-Gaussian Innovations and Applications to the Valuation of Longevity Swaps, *Journal of Risk and Insurance*, 80(3): 775-798.

Wang, C.-W., S. S. Yang, and H.-C. Huang, 2015, Modeling Multicountry Mortality Dependence and Its Application in Pricing Survivor Index Swaps—A Dynamic Copula Approach, *Insurance: Mathematics and Economics*, 63: 30-39.

Ward, J. H. Jr., 1963, Hierarchical Grouping to Optimize an Objective Function, *Journal of the American Statistical Association*, 58(301): 236-244.

Zhang, W., D. Zhao, and X. Wang, 2013, Agglomerative Clustering via Maximum Incremental Path Integral, *Pattern Recognition*, 46(11): 3056-3065.

Zhou, R., J. S.-H. Li, and K. S. Tan, 2013, Pricing Standardized Mortality Securitizations: A Two-Population Model With Transitory Jump Effects, *Journal of Risk and Insurance*, 80(3): 733-774.

Zhou, R., Y. Wang, K. Kaufhold, J. S.-H. Li, and K. S. Tan, 2014, Modeling Mortality of Multiple Populations With Vector Error Correction Models: Applications to Solvency II, *North American Actuarial Journal*, 18(1): 150-167.

Zhu, W., C.-W. Wang, and K. S. Tan, 2016, Structure and Estimation of Lévy Subordinated Hierarchical Archimedean Copulas (LSHAC): Theory and Empirical Tests, *Journal of Banking, & Finance*, 69: 20-36.

Zhu, W., K. S. Tan, and C.-W. Wang, 2017, Supplementary Material to Modeling Multicountry Longevity Risk With Mortality Dependence: A Levy Subordinated Hierarchical Archimedean Copulas Approach, *Journal of Risk and Insurance*, http://dx.doi.org/10.1111/jori.12198

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Appendices