Economic Pricing of Mortality-Linked Securities: A Tâtonnement Approach

Rui Zhou, Johnny Siu-Hang Li and Ken Seng Tan

Abstract

In previous research on pricing mortality-linked securities, the no-arbitrage approach is often used. However, this approach, which takes market prices as given, is difficult to implement in today’s embryonic market where there are few traded securities. In this article, we tackle the pricing problem from a different angle by considering methods that are more related to fundamental economic concepts. Specifically, we treat the pricing work as a Walrasian tâtonnement process, in which prices are determined through a gradual calibration of supply and demand. We illustrate the proposed pricing framework with a hypothetical mortality-linked security and mortality data from the US population.

Keywords: Longevity risk; Mortality bonds; Securitization; Stochastic mortality
1 Introduction

Many financial institutions are exposed to the risk of unexpected changes in human mortality. The risk is two-sided. On the one hand, life insurers paying death benefits will suffer an economic loss if actual rates of mortality are in excess of those expected, due to catastrophic events such as a severe outbreak of influenza and a major terrorist attack. This side of the risk is referred specifically to as mortality risk. On the other hand, pension plan sponsors, as well as insurance companies providing retirement annuities, are subject to longevity risk, that is, the risk that people outlive their expected lifetimes. For these institutions, the longer people live, the greater the period of time over which retirement income must be paid and, hence, the larger the financial liability.

An unexpected change in mortality rates will affect all policies in force. Therefore, as opposed to the random variations between lifetimes of individuals, it cannot be diversified away by increasing the size of the portfolio. Reinsurance is one possible solution to the problem, but its capacity is usually limited. Alternatively, the risk may be naturally hedged or reduced through balancing products. For example, an insurance company may sell life insurance to the same lives who are buying life annuities. The resulting combination would then reduce the company’s exposure to future changes in mortality, consequently permitting a reduction of capital reserves held in respect of mortality or longevity risk. However, this strategy, as Cox and Lin (2007) pointed out, may be cost prohibitive and may not be practical in some circumstances.

In recent years, financial markets have also produced solutions by providing a variety of securities with payoffs tied to certain mortality or longevity indexes. In December 2003 we saw the first mortality bond, issued by Swiss Re. The principal of the bond would have been reduced if there had been a catastrophic mortality event during the life of the bond, therefore allowing Swiss Re to reduce some of its exposure to extreme mortality risk. Then in November 2004, BNP Paribas and the European Investment Bank (EIB) announced a 25-year longevity bond, which was intended for UK pension funds with exposures to longevity risk. This bond took the form of an annuity bond with annual coupon payments tied to the realized survival rates for some English and Welsh males. Other financial institutions such as Goldman Sachs and JP Morgan have also issued mortality-linked securities. We refer readers to Blake et al. (2008), Coughlan (2009) and Zhou and Li (2013) for further details.

A fundamental question in the study of mortality-linked securities is how to place
a value on them. In previous research on pricing mortality-linked securities, the no-arbitrage approach is often used. Generally speaking, to implement the no arbitrage approach, the first step is to estimate the distribution of future mortality rates in the real-world probability measure. Then the real-world distribution is transformed to its risk-neutral counterpart, on the basis of the actual prices of mortality-linked securities we observe in the market. Finally, the price of a mortality-linked security can be calculated by discounting, at the risk-free interest rate, its expected payoff under the identified risk-neutral probability measure. Note that this approach takes actual prices as given. As we see from the following discussion, the need of market prices makes the approach difficult to implement in today’s embryonic market.

One way to implement the no arbitrage approach is to use a stochastic mortality model, which is, at the very beginning, defined in the real-world measure and fitted to past data. The model is then calibrated to market prices, yielding a risk-neutral mortality process from which security prices are calculated. For instance, Cairns et al. (2006) calibrate a two-factor mortality model to the price of the BNP/EIB longevity bond, which is, as of this writing, the only long-term longevity security with pricing information available in the public domain. The resulting risk-neutral mortality process contains two market prices of risk, $\lambda_1$ and $\lambda_2$, one for each stochastic factor. With only one longevity bond price, they cannot be uniquely identified. As a result, an arbitrary assumption, for example, $\lambda_1 = \lambda_2$, must be made before any pricing work can be performed.

We may also make use of a distortion operator such as the Wang transform (Wang, 1996, 2000, 2002) to create a risk-neutral measure, under which mortality-linked securities can be priced. The Wang transform was first applied to mortality-linked securities by Lin and Cox (2005), and subsequently by other researchers including Denuit et al. (2007) and Dowd et al. (2006). Unless a very simple mortality model is assumed, parameters in the distortion operator are not unique if we are not given sufficient market price data. For example, when Chen and Cox (2009) used their extended Lee-Carter model with transitory jump effects to price a mortality bond, they were required to estimate three parameters in the Wang transform. To solve for the three parameters, Chen and Cox assumed that they were equal, but such an assumption is not easy to justify.

Recently, some researchers, for example, Li (2010) and Li and Ng (2011), have implemented the no arbitrage approach by a method called canonical valuation. This
method identifies a risk-neutral measure by minimizing the Kullback-Leibler information criterion, subject to market price constraints. It can be applied without making the arbitrary decisions needed in the aforementioned methods, even if we are given only the market price of the BNP/EIB bond. However, a few problems still remain. In particular, using a product that is very much bond-like is prone to distortions in the identification of pure longevity risk premia. One may doubt if the resulting risk-neutral measure is appropriate for pricing products with different liquidity profiles. For similar reasons, one may also question if the identified risk-neutral measure is applicable to securities that are linked to other reference populations.

In this paper, we approach the pricing problem from a different angle by considering a tâtonnement approach, an approach that is based on the most fundamental economic concept: demand and supply. The tâtonnement approach is highly transparent, since by working on the demand and supply from different economic agents, we know where the price of a mortality-linked security comes from. It also spares us from an arbitrary choice of a risk-neutral probability measure and other problems associated with the no-arbitrage approach when there is a lack of market price data. The tradeoff is that we need to impose more structure than in the no-arbitrage approach. For example, we have to specify a utility function for each party involved in the trade of a mortality-linked security. Another limitation is that the tâtonnement approach requires the assumption of a competitive market. This criterion is not met by the current mortality/longevity market, which is still in its infancy. Nevertheless, prices calculated with the tâtonnement approach can at least be used as a benchmark, particularly in situations when standard no-arbitrage methods are difficult to implement.

In more detail, the method we propose models the trade between two economic agents, one of which suffers mortality or longevity risk and issues a mortality-linked security to offset the risk, and the other of which invests in the mortality-linked security, possibly for earning a risk premium. It is assumed that, given a price, both agents maximize their expected terminal utilities by altering their demand or supply of the security. The estimated price of the security is the price at which the demand and supply are equal, that is, the market clears. On top of the estimated price, our pricing framework provides us with a pair of demand and supply curves. These curves can tell us the optimal quantity of a mortality-linked security to be traded. They also indicate how the supply and demand of the security will evolve with respect to a change in price. This piece of information is particularly useful when we analyze
a new security that has never been traded.

The remainder of this paper is organized as follows. In Section 2, we introduce the idea of a tâtonnement process. Given this idea, we set up our pricing framework in a single-period setting, and then extend it to a multi-period one. In Section 3, we use our proposed framework to price a hypothetical mortality-linked security. The assumed mortality model, the resulting prices, and a comparison with the results from the work of Chen and Cox (2009) are presented. In Section 4 we perform sensitivity tests on different assumptions we have made in the pricing process. In Section 5, we further generalize our pricing framework by allowing trades between the counterparties before the mortality-linked security matures. We detail in this section the required sequential decision process and an algorithm for implementing the process. Finally, in Section 6, we conclude the paper with some suggestions for further research.

2 A Tâtonnement Approach

2.1 The Idea

The idea of tâtonnement in an exchange economy was first proposed by Walras (1874). A Walrasian tâtonnement process assumes that there exists a Walrasian auctioneer who matches supply and demand in a market with perfect information and no transaction costs. There exists a fictitious auctioneer who matches supply and demand from different economic agents. The agents’ behaviors emerge from utility maximization, subject to budget constraints. The auctioneer cries a price, and the agents act to the price by determining how much they would like to offer (supply) or purchase (demand). Transactions only take place at equilibrium price, which equilibrates supply and demand. Otherwise, the price is lowered if there is an excess supply, or raised if there is an excess demand, until an equilibrium price is reached.

The theory of tâtonnement has different interpretations. Some economists view it as a dynamic theory of the equilibrating behavior of real competitive markets, while some treat it as a mathematical solution to the equations of general equilibrium. We refer interested readers to Goodwin (1951), Jaffé (1981) and Walker (1987) for

\footnote{This seminal work of Walras (1874) was translated to English by William Jaffe in 1954. The translated version is cited as Walras (1954).}
extensive discussions on how a tâtonnement may be interpreted. Note that in a tâtonnement process, the equilibrium price might not exist, and if it exists, it might not be unique. The existence and uniqueness of a tâtonnement equilibrium price have been studied by researchers including Arrow and Hurwicz (1958, 1960) and Arrow et al. (1959). Recently, the idea of tâtonnement has received much attention in the areas of operations research and computer science. For example, Cole and Fleischer (2008) analyzed fast-converging tâtonnement algorithms for one-time and ongoing market problems.

In what follows, we will formulate the pricing of a mortality-linked security as a tâtonnement process. We will first present the formulation in a single-period set-up, in which we assume that there is only one payout from the mortality-linked security in question. The single-period set-up is quite restrictive, but it allows the readers to capture the basic ideas behind our pricing framework. We will then extend it to a multi-period set-up, which is applicable to a wide variety of mortality-linked securities.

2.2 A Single-Period Set-up

We use a Walrasian tâtonnement process to model the trade of a mortality-linked security between two economic agents, Agents A and B. Suppose that Agent A has a life contingent liability that is due at time 1. We denote this amount by \( f(q_1) \), which is a deterministic function of \( q_1 \), the mortality index for a certain reference population at time 1. Agent A can be a life insurer which has some death benefits due at time 1, or a pension plan provider which has some living benefits due at time 1. At time 0, \( q_1 \) is not known and is governed by an underlying stochastic process.

To mitigate its exposure to mortality or longevity risk, Agent A sells a mortality-linked security maturing in one year. The payout from one unit of this security at time 1 is \( g(q_1) \), which is also a deterministic function of \( q_1 \). The other economic agent, Agent B, is an investor who might invest in the mortality-linked security, possibly for earning a risk premium.

Following the specification of a tâtonnement process, we suppose that there exists an imaginary auctioneer who cries an arbitrary price, say \( P \), at the beginning. Given this price, Agents A and B then decide their supply \( \theta^A \) and demand \( \theta^B \) of the mortality-linked security, respectively. Their decisions are based on a certain crite-
rion. In this paper, we assume that the agents will choose a supply or demand of the security that will maximize their expected terminal utilities.

Let $\omega^A$ and $\omega^B$ be the initial wealths of Agents A and B, respectively. It is assumed that the wealth of each agent can only be invested in either the mortality-linked security or a bank account which yields a continuously compounded risk-free interest rate of $r$ per annum. We allow a negative wealth, which means that the agent borrows money from a bank account and pays an interest rate of $r$ to the bank. Other than the bank account, the mortality-linked security and the life contingent liability, there is no other sources of income or payout.

We denote the utility functions for Agents A and B by $U^A$ and $U^B$, respectively. At time 0, Agent A sells $\theta^A$ units of the mortality-linked security and deposit both the proceeds and its initial wealth into its bank account. On the other hand, Agent B uses part of its initial wealth to purchase $\theta^B$ units of the security and deposits the rest of its wealth into the bank account. At time 1, the terminal wealth for Agent A would be the amount in its bank account less the payout arising from the mortality-linked security and its life contingent liability, while that for Agent B would be the amount in its bank account plus the payout from the mortality-linked security purchased from the other agent.

As mentioned earlier, at time 0, each agent chooses a supply or a demand of the security that maximizes its expected terminal utility. Using the notation defined above, the objectives of the two agents can be formulated as follows:

Agent A: \[ \sup_{\theta^A} E[U^A((\omega^A + \theta^A P)e^r - \theta^A g(q_1) - f(q_1))] \]  \hspace{1cm} (2.1)

Agent B: \[ \sup_{\theta^B} E[U^B((\omega^B - \theta^B P)e^r + \theta^B g(q_1))] \]  \hspace{1cm} (2.2)

We emphasize that the supply, $\theta^A$, and demand, $\theta^B$, are functions of the price $P$. We suppress the argument of these functions for simplicity.

Usually, the first guess of the price does not clear the market. If the market does not clear, the auctioneer has to adjust the price. It is obvious that the price needs to be raised if demand exceeds supply (i.e., $\theta^B - \theta^A > 0$), and vice versa. Mathematically, the $(k+1)$th update of the price can be expressed as

\[ P_{k+1} = P_k + h_k, \quad k = 1, 2, \ldots, \]  \hspace{1cm} (2.3)

where $P_0$ is the initial guess of the price, and $h_k$ is a function that always has the same sign as the excess demand, $\theta^B - \theta^A$. 

7
In our numerical illustrations, we assume that the auctioneer adjusts the price in a way linear to the excess demand. More specifically, we assume that the price is adjusted according to equation (2.3) with \( h_k = \gamma |P_k|(\theta^B - \theta^A) \), where \( \gamma \) is a positive real constant. Such a linear function is intuitive and is also considered by, for example, Katzner (1999), Kitti (2010) and Uzawa (1960). The constant \( \gamma \) has to be chosen carefully. If \( \gamma \) is too large, the changes in \( P \), \( \theta^A \), and \( \theta^B \) in each iteration tend to be large. This may lead us to missing the equilibrium. In contrast, if \( \gamma \) is too small, the adjustment process tends to be slow. Hence, there is a tradeoff between speed and accuracy. Some experiments have been done to find an appropriate value of \( \gamma \). We will revisit this problem in Section 4.

It is interesting to note that we are essentially solving the equation \( Z(P) = \theta^B - \theta^A = 0 \) for \( P \). If Newton’s method is used to solve the equation, then the update of \( P \) will be have the same form as equation (2.3), with

\[
 h_k = \left( \frac{\partial \theta^A}{\partial P} \bigg|_{P=P_k} - \frac{\partial \theta^B}{\partial P} \bigg|_{P=P_k} \right)^{-1} (\theta^B - \theta^A). 
\]

Since we have \( \frac{\partial \theta^A}{\partial P} \geq 0 \) and \( \frac{\partial \theta^B}{\partial P} \leq 0 \), \( h_k \) always has the same sign as \( \theta^B - \theta^A \). We have experimented this alternative method and found that it works for the pricing process we present in this section. Nevertheless, it is no longer applicable to the generalized pricing process which we will present in Section 5, as that pricing process involves some discretization.

Summing up, the tâtonnement process for pricing a mortality-linked security can be carried out by the algorithm below:

**Algorithm 1**

1. Guess a price \( P_0 \).
2. Determine the demand, \( \theta^B \), and supply, \( \theta^A \), on the basis of the current estimate of the price and the optimizing criteria specified by equations (2.1) and (2.2).
3. If \( \theta^A = \theta^B \), stop. Otherwise, adjust the price using equation (2.3).
4. Repeat Steps 2 and 3.

\(^2\)The law of demand and supply implies that \( \frac{\partial \theta^A}{\partial P} \geq 0 \) and \( \frac{\partial \theta^B}{\partial P} \leq 0 \). To have Newton’s method work, we require that the partial derivatives are not both zero.
Figure 1: Possible situations when Algorithm 1 terminates.

Let \( P^* \) be the price at which the algorithm terminates. There are two possible situations. We may obtain \( P^* > 0 \) and \( \theta^A = \theta^B \), which means that the market attains equilibrium and trade happens between the economic agents (see the left panel of Figure 1). However, it is also possible that trade will not occur. Examples of such a situation are illustrated in the middle and right panels of Figure 1.

2.3 A Multi-period Set-up

We now extend the tâtonnement process for pricing a mortality-linked security to a multi-period set-up. In this set-up, we allow payments to be made before the mortality-linked security matures at time \( T \). Denote \( Q_t = (q_1, \ldots, q_t) \), where \( q_t \) is the mortality index for a certain reference population over the period of \( t-1 \) to \( t \). At time 0, the values of \( q_t \) for \( t > 0 \) are not known and are governed by an underlying stochastic process.

Again we model the trade of a mortality-linked security between two economic agents, Agents A and B. Agent A has life contingent liabilities that are due at \( t = 1, 2, \ldots, T \). The amount due at time \( t \) is \( f_t(Q_t) \), where \( f_t \) is a deterministic function of \( Q_t \). To hedge the liabilities, Agent A issues a mortality-linked security, which pays an amount of \( g_t(Q_t) \) at time \( t \), where \( g_t \) is a deterministic function of \( Q_t \).\(^3\) Agent B is

\(^3\)We assume that \( f_t \) and \( g_t \) are functions of \( Q_t \) rather than \( q_t \) because the payouts can be path dependent.
an investor who may invest in the mortality-linked security. It receives at time $t$ an amount of $g_t(Q_t)$ per unit of the mortality-linked security invested.

The supply from Agent A is $\theta^A$ and the demand from Agent B is $\theta^B$. At this stage, we do not allow the agents to trade the mortality-linked security during the term of the security. This assumption will be relaxed in the generalization in Section 5.

We keep all other assumptions in the single-period set-up. Let $W^A_t$ and $W^B_t$ be the time-$t$ wealths for Agents A and B, respectively. Given the assumptions we made, the wealth process for each agent can be represented as follows:

Agent A

\[
\begin{align*}
W^A_0 &= \omega^A \\
W^A_1 &= (W^A_0 + \theta^A P)e^r - \theta^A g_1(Q_1) - f_1(Q_1) \\
W^A_2 &= W^A_1 e^r - \theta^A g_2(Q_2) - f_2(Q_2) \\
& \vdots \\
W^A_T &= W^A_{T-1} e^r - \theta^A g_T(Q_T) - f_T(Q_T)
\end{align*}
\]  

Agent B

\[
\begin{align*}
W^B_0 &= \omega^B \\
W^B_1 &= (W^B_0 - \theta^B P)e^r + \theta^B g_1(Q_1) \\
W^B_2 &= W^B_1 e^r + \theta^B g_2(Q_2) \\
& \vdots \\
W^B_T &= W^B_{T-1} e^r + \theta^B g_T(Q_T)
\end{align*}
\]  

Each agent chooses a demand or supply that maximizes its expected terminal utility. As such, given a price $P$, the supply from Agent A is $\theta^A = \arg\sup_{\theta^A} \mathbb{E}[U^A(W^A_T)]$, while the demand from Agent B is $\theta^B = \arg\sup_{\theta^B} \mathbb{E}[U^B(W^B_T)]$. Of course, in general, the initial guess of the price will not lead to $\theta^A = \theta^B$. We may adjust the price by using Algorithm 1, with equations (2.1) and (2.2) replaced by $\sup_{\theta^A} \mathbb{E}[U^A(W^A_T)]$ and $\sup_{\theta^B} \mathbb{E}[U^B(W^B_T)]$, respectively.

Instead of maximizing each agent’s expected terminal utility, one may be interested in maximizing each agent’s expected lifetime utility based on a discount factor.
Interestingly, under the assumption of exponential utility functions, both optimization criteria would lead to the same equilibrium price and quantity traded. A proof of this result is presented in the Appendix.

3 An Illustration

3.1 The Mortality Model

Before we implement the tâtonnement process, we need a stochastic process to model the randomness of the mortality index. We consider the Lee-Carter family, which is quite well-known in the insurance industry. The Lee-Carter model in its original form (Lee and Carter, 1992) can be expressed mathematically as

\[
\ln(m_{x,t}) = \beta_{x}^{(0)} + \beta_{x}^{(1)} \kappa_t + \epsilon_{x,t}, \tag{3.1}
\]

where \(m_{x,t}\) denotes the central death rate at age \(x\) and in year \(t\), \(\beta_{x}^{(0)}\) is the average level of mortality (in log scale) over time, \(\beta_{x}^{(1)}\) is the age-specific sensitivity to the time-varying factor, \(\kappa_t\), which governs the dynamics of central death rates at all ages. The error term \(\epsilon_{x,t}\), which captures all remaining variations, is assumed to have no trend over both age and time dimensions. To forecast future mortality, we need to further model \(\kappa_t\) by a time-series process. Usually, a random walk with drift, that is,

\[
\kappa_{t+1} = \kappa_t + \mu + \sigma Z_{t+1},
\]

where \(\mu\) and \(\sigma\) are constants, and \(\{Z_t\}\) is a sequence of iid standard normal random variables, is used.

As shown in the work of Li and Chan (2005, 2007), the series of \(\kappa_t\) may be contaminated with outliers, which correspond to events such as wars and pandemics. The outliers should not be neglected in pricing mortality-linked securities, especially those for hedging extreme mortality risk. Ignoring these outliers may lead to underestimating the probability of having a catastrophic event, which may bring us a large pricing error. Therefore, rather than the original Lee-Carter model, we use an extension proposed by Chen and Cox (2009). The extension has the same structure as the specification in equation (3.1), but permits transitory jumps in the evolution of \(\kappa_t\). In particular, it models the time-varying factor \(\kappa_t\) as the sum of two components. The first component, denoted by \(\tilde{\kappa}_t\), is a time-varying factor that is free of any jump
<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu$</td>
<td>$-0.2173$</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>$0.3733$</td>
</tr>
<tr>
<td>$m$</td>
<td>$0.8393$</td>
</tr>
<tr>
<td>$s$</td>
<td>$1.4316$</td>
</tr>
<tr>
<td>$p$</td>
<td>$0.0436$</td>
</tr>
</tbody>
</table>

Table 1: Estimates of parameters in the Lee-Carter model with transitory jump effects. (Source: Chen and Cox (2009).)

It is assumed that, in each year, there is at most one jump event with probability $p$; that is,

$$N_t = \begin{cases} 
1, & \text{with probability } p, \\
0, & \text{with probability } 1 - p,
\end{cases}$$

where $N_t$ denotes the number of jumps occurring in year $t$. It is also assumed that the jump severity variable, $Y_t$, at time $t$ is a normal random variable with mean $m$ and standard deviation $s$, and that $Y_t$ is independent of the jump frequency variable $N_t$. In effect, the entire stochastic process for $\kappa_t$ can be expressed as follows:

$$\begin{align*}
\tilde{\kappa}_{t+1} &= \tilde{\kappa}_t + \mu + \sigma Z_{t+1}, \\
\kappa_{t+1} &= \tilde{\kappa}_{t+1} + N_{t+1} Y_{t+1},
\end{align*}$$

where $\mu$ and $\sigma$ are constants, and $Z_t$ is a standard normal random variable that is independent of both $Y_t$ and $N_t$.

Chen and Cox (2009) fitted the extended Lee-Carter model to US mortality data from 1900 to 2003, which were provided by the National Center for Health Statistics (NCHS). The data contain age-specific death rates for age 0, age group 1-4, 10-year age groups from 5-14 to 75-84, and age group 85 and over. The resulting estimates of the parameters in equation (3.3) are displayed in Table 1. The fitted values of $\beta_x^{(0)}$ and $\beta_x^{(1)}$ can be found on p.734 of Chen and Cox (2009).

In what follows, we will price an illustrative mortality-linked security. We assume that this security is linked to the population from which the NCHS data were ob-
tained. The extended Lee-Carter model, with parameters shown in Table 1, will be used in the tâtonnement pricing process.

3.2 Pricing a Mortality-Linked Security

We assume that the illustrative mortality-linked security is linked to a mortality index $q_t$, which is defined by the average of the mortality rates for different age groups. The weights are based on the 2000 US standard population: 0.013818 for age under 1 year, 0.055317 for ages 1-4, 0.145565 for ages 5-14, 0.138646 for ages 15-24, 0.135573 for ages 25-34, 0.162613 for ages 35-44, 0.134834 for ages 45-54, 0.087247 for ages 55-64, 0.066037 for ages 65-74, 0.044842 for ages 75-84, and 0.015508 for age 85 and over.\footnote{Source: NCHS report GMWK293R (http://www.cdc.gov/nchs/nvss/mortality/gmwk293r.htm).}

The mortality index we use is exactly the same as that in Chen and Cox (2009). Assume further that Agent A has sold life insurance policies which pay a total benefit $f_t(Q_t) = 1000q_t$ at time $t$. To hedge its exposure to mortality risk, Agent A issues a mortality-linked security with a face value of $\$1$. The security is fairly similar to the mortality bond issued by Swiss Re in December 2003. In particular, the security pays a coupon at the end of each year at a rate of $r + 1.5\%$, where $r$ is the risk-free interest rate, which is assumed to be 3\% in our baseline calculations. The principal repayment at maturity depends on the values of $q_t$ over the term of the security. Specifically, the principal repayment is specified as follows:

$$\text{Principal Repayment} = \max\left(1 - \sum_{t=1}^{3} \text{loss}_t, 0\right),$$

where $\text{loss}_t$ is defined by

$$\text{loss}_t = \max(q_t - 1.1q_0, 0) - \max(q_t - 1.2q_0, 0) \over 0.1q_0.$$

In using Algorithm 1 to obtain the price of the mortality-linked security, there is a need to evaluate the expected terminal utility for each agent. The expectation can be calculated by Monte Carlo simulations as follows:

1. simulate 50 000 paths for $N_t$, $Z_t$, and $Y_t$;
2. calculate $\kappa_t$ and $q_t$ using the simulated paths;
3. calculate the terminal utility for each path;

4. take arithmetic average of all the simulated terminal utilities as the expected terminal utility.

We assume an exponential utility function, \( U(x) = 1 - e^{-kx} \), for each agent. In the utility function, parameter \( k \) is the absolute risk aversion for all wealth levels. A larger \( k \) means that the agent is more conservative and risk averse. In a study of an insurer’s optimal premium strategy, Emms and Haberman (2009) assume \( k = 1.0 \) for an insurer. We also assume in our baseline calculations that Agent A, an insurer, has an absolute risk aversion of \( k_A = 1.0 \). It is reasonable to assume that Agent A is more conservative than Agent B, because Agent A wants to hedge its mortality risk exposure while Agent B is willing to take mortality risk in return of a risk premium. Moreover, there is a good chance that Agent B is a hedge fund, which should have a low absolute risk aversion (see, e.g., Zhu (2009)). In our baseline calculations, the assumed value of \( k \) for Agent B is \( k_B = 0.5 \).

Using Algorithm 1, the estimated price of the mortality-linked security is $1.0319 and the optimal quantity of the security traded is 1.86 units. For each price level, we can calculate the demand \( \theta^B \) and supply \( \theta^A \). This allows us to plot a curve of \( \theta^B \) against \( P \) (the demand curve) and a curve of \( \theta^A \) against \( P \) (the supply curve). The resulting demand and supply curves are shown in Figure 2. We observe that the curves intersect at one single point, giving a unique price of the security.

We also observe that the supply is 0 when price is less than $1.0131, indicating that Agent A is not willing to sell any mortality-linked security for less than $1.0131. When the price exceeds $1.0131, the supply increases strictly with price. The demand has an opposite trend. It decreases with the price of the security until $1.0361, after which it remains at 0.

The use of exponential utility functions has pros and cons. When exponential utility functions are used, the agents’ initial wealths do not affect the equilibrium (see Section 4.1). This property facilitates optimization. Further, in our set-up, an agent’s wealth might become negative in some future time. A negative wealth is permitted when exponential utility functions are assumed.

In principle, we may implement the pricing framework with other utility functions, for example, power utility functions. However, to work with power utility functions, we need an assumed value of the initial wealth of each agent. Also, power utility
functions do not permit a negative wealth.

3.3 Comparing with an Alternative Method

What price would an existing pricing method give to our illustrative mortality-linked security? To answer this question, we reprice the security using the pricing method in Chen and Cox (2009), which is based on a mortality model that is completely identical to what we assume in our tâtonnement pricing process. However, rather than an economic approach, Chen and Cox use the Wang transform to identify a risk-neutral probability measure, from which the price of the mortality-linked security can be calculated.

More specifically, Chen and Cox apply the Wang transform to random variables $Z_t$, $Y_t$, and $N_t$ in equation (3.3) individually to obtain the following jump mortality process in a risk-neutral probability measure:

$$
\begin{align*}
\tilde{\kappa}_{t+1}^* &= \tilde{\kappa}_t^* + \mu + \sigma Z_{t+1}^*, \\
\tilde{\kappa}_{t+1}^* &= \tilde{\kappa}_{t+1}^* + N_{t+1}^* Y_{t+1}^*,
\end{align*}
$$
where $Z_t^* \sim N(\lambda_1, 1), Y_t^* \sim N(m + \lambda_2 s, s^2)$,

$$\tilde{N}_t = \begin{cases} 
1, & \text{with probability } p^*,
0, & \text{with probability } 1 - p^*,
\end{cases}$$

and $p^* = 1 - \Phi(\Phi^{-1}(1 - p) - \lambda_3)$. Here, $\Phi$ is the cumulative distribution function (cdf) for a standard normal random variable and $\Phi^{-1}$ is the inverse of the cdf for a standard normal random variable.

In the above, the unknown constants $\lambda_1$, $\lambda_2$, and $\lambda_3$ may be viewed as the market prices of risk associated with $Z_t$, $Y_t$, and $N_t$, respectively. Chen and Cox solve the unknowns by equating the actual price of the Swiss Re mortality bond and the price of the bond implied by the risk-neutral jump mortality process. As there are three unknowns but only one equation, there exists infinitely many possible combinations of $\lambda_1$, $\lambda_2$, and $\lambda_3$. In their calculations, Chen and Cox assume that the market prices of risk are equal. They also show in their paper the value of $\lambda_1$ when $\lambda_2 = \lambda_3 = 0$ and the value of $\lambda_2$ when $\lambda_1 = \lambda_3 = 0$, but these two sets of values are not used in their pricing work. The first three rows of Table 2 display the market prices of risk calculated in the work of Chen and Cox.

We calculate, with Monte Carlo simulations, the expected payoff from the illustrative mortality-linked security on the basis of the jump mortality process in the identified risk-neutral measure. By discounting the expected payoff at the assumed risk-free interest rate, we obtain the estimated price of the security. The prices under different assumed market prices of risk are shown in the last row of Table 2. Assuming $\lambda_1 = \lambda_2 = \lambda_3$, the method of Wang transform would give a price of $0.4976$.

The simulated results depicted in Table 2 clearly suggest the estimated price of the security is highly sensitive to the assumed values of $\lambda_1$, $\lambda_2$, and $\lambda_3$. For instance, if we assume $\lambda_2 = \lambda_3 = 0$, the estimated price would be as small as $0.1318$. The

<table>
<thead>
<tr>
<th>$\lambda_1$</th>
<th>5.1449</th>
<th>0</th>
<th>1.5000</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\lambda_2$</td>
<td>0</td>
<td>3.4808</td>
<td>1.5000</td>
</tr>
<tr>
<td>$\lambda_3$</td>
<td>0</td>
<td>0</td>
<td>1.5000</td>
</tr>
<tr>
<td>Price</td>
<td>$0.1318$</td>
<td>$0.9384$</td>
<td>$0.4976$</td>
</tr>
</tbody>
</table>

Table 2: Prices of the illustrative mortality-linked security implied by different market prices of risk. (The values of $\lambda_1$, $\lambda_2$ and $\lambda_3$ are obtained from Chen and Cox (2009).)
range of arbitrage-free prices is huge, and the method of Wang transform leaves us no clue to choose a price from this range.

Among the three prices in Table 2, the price based on the assumption \( \lambda_1 = \lambda_3 = 0 \) is the closest to the price estimated by our proposed method ($1.0319). This is rather intuitive, as \( \lambda_2 \) is associated with jumps in mortality, which is exactly the risk that the security intends to hedge. It would also be interesting to see how the estimated price would change if \( \lambda_1 = \lambda_2 = 0 \) is assumed, as \( \lambda_3 \) is also associated with jumps. However, this set of market prices of risk is not provided by Chen and Cox (2009).

4 Sensitivity Tests

A few assumptions have been made in our tâtonnement pricing process. In this section, we examine how changes to these assumptions may affect the estimated price of the security in question.

4.1 Initial Wealths

Recall that we allow both agents to borrow money from the bank. This means that the initial wealth does not limit the quantity of the mortality-linked security that Agent B can purchase at time 0. Moreover, if we assume exponential utility functions for both agents, the initial wealth of each agent has no effect on the estimated price of the security.

**Proposition 1.** If exponential utility functions are assumed, the equilibrium price \( P \) is independent of the initial wealths \( \omega_A \) and \( \omega_B \).

**Proof.** We aim to prove that \( \omega_A \) and \( \omega_B \) do not affect the demand and supply curves for the security. It is easy to prove by induction that

\[
W_T^A = (W_0^A + \theta^A P)e^{rT} - \sum_{i=0}^{T-1} e^{(T-i-1)r} [\theta^A g_{i+1}(Q_{i+1}) + f_{i+1}(Q_{i+1})],
\]

and that

\[
W_T^B = (W_0^B - \theta^B P)e^{rT} + \sum_{i=0}^{T-1} e^{(T-i-1)r} \theta^B g_{i+1}(Q_{i+1}).
\]
For convenience, we let

$$F(\theta^A, Q_1, Q_2, \ldots, Q_T) = \theta^A Pe^{rT} - \sum_{i=0}^{T-1} e^{(T-i-1)r}[\theta^A g_{i+1}(Q_{i+1}) + f_{i+1}(Q_{i+1})]$$

and

$$G(\theta^B, Q_1, Q_2, \ldots, Q_T) = -\theta^B Pe^{rT} + \sum_{i=0}^{T-1} e^{(T-i-1)r}\theta^B g_{i+1}(Q_{i+1}).$$

Given a price $P$, the supply $\theta^A$ is given by

$$\text{argsup}_{\theta^A} \mathbb{E}[U^A(W^*_A)] = \text{argsup}_{\theta^A} \mathbb{E}[-e^{-k^AW^*_A}]$$

$$= \text{argsup}_{\theta^A} \mathbb{E}[-e^{-k^AW^*_A}e^{rT}e^{-k^AF(\theta^A, Q_1, Q_2, \ldots, Q_T)}]$$

$$= \text{argsup}_{\theta^A} \mathbb{E}[-e^{-k^AF(\theta^A, Q_1, Q_2, \ldots, Q_T)}],$$

which does not depend on $W^*_0$ and hence $\omega^A$. Similarly, given $P$, the demand $\theta^B$ is

$$\text{argsup}_{\theta^B} \mathbb{E}[U^B(W^*_B)] = \text{argsup}_{\theta^B} \mathbb{E}[-e^{-k^BG(\theta^A, Q_1, Q_2, \ldots, Q_T)}],$$

which again does not depend on $W^*_0$ and hence $\omega^B$. Since both initial wealths have no effect on the demand and supply curves, they do not affect the equilibrium price of the security.

### 4.2 Risk Aversion Parameters

In the baseline calculations, we assume that the risk aversion parameters for Agents A and B are $k^A = 1.0$ and $k^B = 0.5$, respectively. We now reprice the illustrative mortality-linked security using different combinations of $k^A$ and $k^B$. The estimated prices and the corresponding numbers of unit traded are displayed in Table 3.

Recall that a larger risk aversion parameter means that the agent is more risk adverse. From Table 3 we observe that at a higher $k^A$, more units would be sold at time 0 and the price at which the security would be sold is lower. This is because if Agent A is more risk adverse, it would have a greater intention to reduce its mortality risk exposure, leading to a higher supply of the mortality-linked security and hence a lower price, other things equal. On the other hand, at a higher $k^B$, less units would be sold at time 0 and the price at which the security would be sold is lower. This is because if Agent B is more risk adverse, it would have a smaller intention to invest...
Table 3: Prices and numbers of unit traded when different risk aversion parameters are assumed.

<table>
<thead>
<tr>
<th>( k^A )</th>
<th>( k^B )</th>
<th>Price</th>
<th>Units traded</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.25</td>
<td>$1.0337</td>
<td>2.31</td>
</tr>
<tr>
<td>1</td>
<td>0.5</td>
<td>$1.0319</td>
<td>1.86</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>$1.0289</td>
<td>1.35</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>$1.0250</td>
<td>0.87</td>
</tr>
<tr>
<td>0.5</td>
<td>0.5</td>
<td>$1.0331</td>
<td>1.41</td>
</tr>
<tr>
<td>2</td>
<td>0.5</td>
<td>$1.0306</td>
<td>2.23</td>
</tr>
<tr>
<td>4</td>
<td>0.5</td>
<td>$1.0299</td>
<td>2.44</td>
</tr>
</tbody>
</table>

in the risky security, resulting in a smaller demand and hence a lower price. From another viewpoint, if Agent B is more risk adverse, it would demand a higher risk premium for the same amount of risk. Consequently, the price of the security must be lowered.

4.3 The Price Adjustment Process

In Section 2.2, we highlighted the role of \( \gamma \) in the price adjustment process. The value of \( \gamma \) has to be small enough. Otherwise, the price adjustment might be too big so that the equilibrium price will be missed. If \( \gamma \) is sufficiently small, then its value merely affects how fast the tâtonnement pricing process would converge.

To illustrate, let us consider two different values of \( \gamma \). If we set \( \gamma = 0.001 \), the price of mortality-linked security can be found in 13 iterations. In contrast, if we set \( \gamma = 0.005 \), the algorithm does not converge but goes into a dead loop. The price adjustment processes on the basis of these two values of \( \gamma \) are illustrated graphically in Figure 3.

4.4 The Risk-free Interest Rate

The risk-free interest rate plays two roles. It affects the coupons paid (the coupon rate is assumed to be \( r + 1.5\% \)) and also the rate at which the economic agents’ wealths are accumulated. In Table 4 we show the estimated prices of the illustrative mortality-linked security under different assumed risk-free interest rates.
Table 4 indicates that the equilibrium is not too sensitive to the risk-free interest rate. This is possibly because when the risk-free interest rate increases, both the coupon rate of mortality bond and the return on a risk-free investment would increase. As both investment vehicles become more attractive, the increase in the demand of the mortality bond would tend to be modest.

5 Allowing Trades after Time 0

At any time point, the optimal position of a mortality-linked security depends on the expectation of future mortality rates. As mortality experience is unfolded, the expectation of future mortality rates at $t > 0$ can be different from that at $t = 0$. As a result, at $t > 0$, the agents may be able to attain a higher expected terminal utility by adjusting their positions of the security.

The set-up presented in Section 2.3 does not permit the economic agents to trade the mortality-linked security after time 0. In this section we will generalize that set-
Table 4: Prices and numbers of unit traded when different risk-free interest rates are assumed.

\[
\begin{array}{|c|c|c|c|c|}
\hline
r & 0.03 & 0.04 & 0.05 & 0.06 \\
\hline
\text{Price} & $1.0319 & $1.0303 & $1.0285 & $1.0264 \\
\hline
\text{Units traded} & 1.86 & 1.88 & 1.90 & 1.92 \\
\hline
\end{array}
\]

up to allow the economic agents to sell or purchase the mortality-linked security at discrete time-steps before the security matures. This generalization may be treated as a sequential decision process.

5.1 Sequential Decision Processes

First, let us introduce a special type of sequential decision processes called Markov decision processes. In discrete Markov decision problems, decisions are made at \( t = 0, 1, \ldots, T - 1 \). At time \( t \), the system occupies a state. We denote the set of all possible states at time \( t \) by \( S_t \). If, at time \( t \), the system is in state \( s_t \in S_t \), the decision maker may choose an action \( a_t \) from \( A_{s_t} \), the set of allowable actions in state \( s_t \). At time \( t \), the system’s state at \( t + 1 \) is unknown and follows a certain probability distribution, which depends on \( s_t \) and \( a_t \).

As ‘Markov’ indicates, the decision depends on the current state \( s_t \) only. The decision is made according to a deterministic decision rule \( d_t \), which maps \( S_t \) to \( A_{s_t} \). The decision rule specifies an action selection procedure in each state \( s_t \) at time \( t \). As a result of choosing action \( a_t \in A_{s_t} \), the decision maker receives an immediate reward, \( r_t(s_t, a_t) \), which is a real-valued function of \( s_t \) and \( a_t \). At maturity (time \( T \)), the decision maker receives a terminal reward of \( r_T(s_T) \).

The decision rules to be used at all decision epochs compose a strategy \( \pi \). Specifically, a strategy \( \pi \) is a sequence of decision rules, that is, \( \pi = (d_0, d_1, \ldots, d_{T-1}) \), where \( d_t \in D_t \) for \( t = 0, 1, \ldots, T - 1 \).

Since the rewards received by the decision maker are not known before the strategy is implemented, the reward sequence is random. The decision maker’s objective is to

\[\text{In the literature, } \pi \text{ is more commonly referred to as a policy. However, to avoid any potential confusion with an insurance policy, we use the term ‘strategy’ instead.}\]
choose a strategy so that the corresponding random reward sequence is as ‘large’ as possible. This necessitates a method for finding an optimal strategy, which we will investigate in Section 5.2.

Some mortality-linked securities make payments that depend not only on the current but also the past values of the underlying mortality index. For such securities, a simple Markov decision process is not adequate, but we may use a less restrictive sequential decision process, which we now detail.

To explain this sequential decision process, we first represent the history of the system by the sequence of previous states and actions. The system’s history, which we denote by \( h_t = (s_0, a_0, \ldots, s_{t-1}, a_{t-1}, s_t) \), follows the recursion \( h_t = (h_{t-1}, a_{t-1}, s_t) \). We let \( H_t \) be the set of all possible histories. Note that \( H_0 = S_0 \), \( H_1 = S_0 \times A_{S_0} \times S_1 \), and \( H_t = S_0 \times A_{S_0} \times S_1 \times A_{S_1} \times \ldots \times S_t \).

Note also that the recursion \( H_t = H_{t-1} \times A_{S_{t-1}} \times S_t \) holds. Here, the decision is a function of \( h_t \) and the reward function is a function of \( h_t \) and \( a_t \). The allowable action set for each \( h_t \in H_t \) is \( A_{h_t} \). The decision process is no longer ‘Markov,’ because the decision rule \( d_t \) maps the entire history \( H_t \) to \( A_{H_t} \).

Nevertheless, the less restrictive sequential decision process can be transformed into a Markov decision process. Specifically, if we set a new state variable \( \hat{s}_t = h_t \), then the sequential decision process on the basis of the new state variable would have the same form as a Markov decision process. The set of all possible states at time \( t \) would become \( \hat{S}_t = H_t \). As such, the theorems and algorithms developed for Markov decision processes can still be applied to transformed sequential decision processes as well.

### 5.2 An Optimal Strategy

Bellman (1957) provides us a simple principle to find an optimal strategy.

**Bellman’s principle of optimality:** An optimal policy has the property that whatever the initial state and initial decision are, the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision.

---

6For sets \( C \) and \( D \), \( C \times D \) denotes the Cartesian product of \( C \) and \( D \); that is, \( C \times D = \{(c, d) : c \in C, d \in D\} \).
Suppose that the objective is to maximize the sum of rewards from time 0 to $T$, that is, $\sum_{t=0}^{T-1} r_t(s_t, a_t) + r_T(s_T)$. For a Markov sequential decision process, the principle of optimality implies that $(a_0, a_1, \ldots, a_{T-1})$ is the optimal actions if

$$a_t = \arg \sup_{a \in A_t} \{ r_t(s_t, a) + \mathbb{E}[v_{t+1}(s_{t+1}) | s_t, a] \},$$

where

$$v_t(s_t) = \begin{cases} \sup_{a \in A_t} \{ r_t(s_t, a) + \mathbb{E}[v_{t+1}(s_{t+1}) | s_t, a] \}, & t = 0, 1, \ldots, T - 1, \\ r_T(s_T), & t = T. \end{cases}$$

When $A_s$ is finite, we can replace the ‘sup’ in the equation above by ‘max.’ A mathemathical definition of the principle of optimality and a detailed proof of the results above can be found in Puterman (2005).

According to the results above, we work backward from $t = T$ to find the optimal actions. At time $T$, we calculate all possible values of $v_T(s_T)$. We then proceed to time $T - 1$. For each possible state $s_{T-1} \in S_{T-1}$, we obtain $v_{T-1}(s_{T-1})$ and the optimal action $a_{T-1}$ by maximizing the expected value of $v_T$. Having completed the calculations for time $T - 1$, we proceed to times $T - 2, \ldots, 0$ in order. In general, the optimal action $a_t$ and $v_t(s_t)$ at $t = 0, 1, \ldots, T - 1$ can be expressed as

$$a_t = \arg \sup_{a \in A_t} \{ r_t(s_t, a) + \mathbb{E}[v_{t+1}(s_{t+1}) | s_t, a] \} \quad (5.1)$$

and

$$v_t(s_t) = r_t(s_t, a_t) + \mathbb{E}[v_{t+1}(s_{t+1}) | s_t, a_t], \quad (5.2)$$

respectively. The above procedure can be conveniently summarized by the following algorithm:

**Algorithm 2**

1. Evaluate $v_T(s_T)$ for each $s_T \in S_T$. Set $t = T$.
2. Find the optimal action $a_{t-1}$ and $v_{t-1}(s_{t-1})$ by equations (5.1) and (5.2). Set $t = t - 1$.
3. Repeat Step 2. Stop when $t = 0$. 

23
Algorithm 2 can also be applied to the less restrictive sequential decision process in which the distribution of \( s_{t+1} \) depends not only on \( s_t \) but also the history of the states and actions. Specifically, we replace \( s_t \) in the procedure above by \( \hat{s}_t = h_t \) and the evaluation is done for each \( \hat{s}_t \in \hat{S}_t \). The amount of information we include in the state variable determines the computation power we need to solve the problem. It is therefore more computationally demanding to work on the less restrictive sequential decision process.

5.3 The Generalized Pricing Process

Let \( \theta^A_t \) be Agent A’s short position in the mortality-linked security immediately before trade at time \( t \). Denote by \( a^A_t \) the action taken by Agent A at time \( t \); that is, \( a^A_t \) is the number of units of the security sold at time \( t \). It is easy to see that \( \theta^A_0 = 0 \) and \( \theta^A_t = \theta^A_{t-1} + a^A_{t-1} \).

We define similar notation for Agent B. We let \( \theta^B_t \) be Agent B’s long position in the security immediately before trade at time \( t \), and let \( a^B_t \) be the number of units of the security purchased at time \( t \). We have \( \theta^B_0 = 0 \) and \( \theta^B_t = \theta^B_{t-1} + a^B_{t-1} \).

If the time-\( t \) payment from the security depends only on the mortality index \( q_t \) at time \( t \), we may set the state variable as \( s_t = (q_t, \theta_t) \). However, more generally, if the payment at time \( t \) also depends on values of the index before time \( t \), then we have to use \( \hat{s}_t = h_t \) as the state variable.

Again we assume that the mortality-linked security and the bank account are the only investment choices. Let \( P_t(\hat{s}_t) \) be the price of the mortality-linked security at time \( t \) and at state \( \hat{s}_t \). Suppose that the price process \( P_0(\hat{s}_0), P_1(\hat{s}_1), \ldots, P_{T-1}(\hat{s}_{T-1}) \) is known. Then the wealth processes for the two economic agents can be written as follows:

Agent A

\[
W^A_0 = \omega^A \\
W^A_1 = (W^A_0 + a^A_0 P_0(\hat{s}_0))e^r - (\theta^A_0 + a^A_0)g_1(Q_1) - f_1(Q_1) \\
W^A_2 = (W^A_1 + a^A_1 P_1(\hat{s}_1))e^r - (\theta^A_1 + a^A_1)g_2(Q_2) - f_2(Q_2) \\
\vdots \\
W^A_T = (W^A_{T-1} + a^A_{T-1} P_{T-1}(\hat{s}_{T-1}))e^r - (\theta^A_{T-1} + a^A_{T-1})g_T(Q_T) - f_T(Q_T) \quad (5.3)
\]
respectively. Note that a process into Algorithm 2, the actions taken by Agents A and B are not likely to agree ever, the price process is not known at the outset. If we plug an arbitrary price the optimal actions for both agents can be found if we know the price process. How-

which means the optimal action

Its optimal action at time

Taking Agent A as an example, its optimal actions are given by

As before, we assume that, given a price process, the agents choose their actions by maximizing their expected terminal utility. We can model this with a sequential decision process by setting \( r_t(s_t, a_t) \) to 0 for \( t = 0, 1, \ldots, T - 1 \) and \( r_T(s_T) \) to the terminal utility of the agent in state \( s_T \). We assume that agents have exponential utility functions \( U^A(x) = 1 - e^{-kAx} \) and \( U^B(x) = 1 - e^{-kBx} \).

We identify \( v_t(s_t) \) in equation (5.2) for Agents A and B by \( v_t^A(s_t) \) and \( v_t^B(s_t) \), respectively. Note that \( v_t^A(s_T) = U^A(W_T^A) \) and \( v_t^B(s_T) = U^B(W_T^B) \).

Taking Agent A as an example, its optimal actions are given by

Its optimal action at time \( T - 1 \) is

which means the optimal action \( a_{T-1}^A \) depends on the past information only through \( Q_{T-1} \) and \( \theta_{T-1}^A \). If we write down the equations for \( t = T - 1, T - 2, \ldots, 0 \), we will see that the optimal action \( a_t^A \) only depends on \( Q_t \) and \( \theta_t^A \) as well. Therefore, we may reduce the the state variable to \( \hat{s}_t = (Q_t, \theta_t^A) \), because that would reduce the content of the state variable and hence the computational effort needed.

With Algorithm 2 and the wealth processes specified by equations (5.3) and (5.4), the optimal actions for both agents can be found if we know the price process. However, the price process is not known at the outset. If we plug an arbitrary price process into Algorithm 2, the actions taken by Agents A and B are not likely to agree with each other and the market is not likely to clear. To make the market clear,
we have to adjust the price process, and this can be accomplished by a tâtonnement approach, which we introduced and used in Section 2. The tâtonnement approach can be combined with a sequential decision process to solve the pricing problem.

The basic idea is to find the equilibrium price at each state successively. We begin with time \( T - 1 \). For each possible state \( \hat{s}_{T-1} \in \hat{S}_{T-1} \), the market clearing price \( P_{T-1}(\hat{s}_{T-1}) \) can be found by using Algorithm 1 and the optimality criterion of maximizing the expected terminal utility. At the same time, we obtain the actions \( a^A_{T-1} \) and \( a^B_{T-1} \) and the values of \( v^A_{T-1}(s_{T-1}) \) and \( v^B_{T-1}(s_{T-1}) \). Then, we repeat the procedure for times \( T - 2, T - 3, \ldots, 0 \) in order. Since there is only one state at time 0, we would be able to obtain a unique time-0 price and the corresponding optimal actions for both agents. The above procedure can be summarized by the following algorithm:

**Algorithm 3**

1. Evaluate \( v^A_T(\hat{s}_T) \) and \( v^B_T(\hat{s}_T) \). Set \( t = T \).

2. For each possible state \( \hat{s}_{t-1} \in \hat{S}_{t-1} \) at time \( t - 1 \), use Algorithm 1 and equations (5.1) and (5.2) to find \( P_{t-1}(\hat{s}_{t-1}), a^A_{t-1}, a^B_{t-1}, v^A_{t-1}(\hat{s}_{t-1}) \), and \( v^B_{t-1}(\hat{s}_{t-1}) \). Set \( t = t - 1 \).

3. Repeat Step 2. Stop when \( t = 0 \). The time-0 price of the mortality-linked security is given by \( P_0(\hat{s}_0) \).

In the Appendix, we consider an alternative set-up that is based on the agents’ lifetime utilities rather than their terminal utilities. We show in the Appendix that if exponential utility functions are assumed, then the tâtonnement equilibrium resulting from this alternative set-up is exactly the same as that from the set-up that is based on terminal utilities. This property holds no matter if trades after time 0 are permitted.

### 5.4 A Multinomial Mortality Tree

In the sequential decision process, the system occupies a state \( s_t \in S_t \) at time \( t \). Optimal actions are determined for each state. For computational reasons, we need to keep the state space \( S_t \) discrete and finite. This means that the stochastic mortality
Time = \( t \) 

\[ \tilde{\kappa}_{t+1}(3) = \tilde{\kappa}_t + \mu + 3\sigma \]

\[ \tilde{\kappa}_{t+1}(2) = \tilde{\kappa}_t + \mu + 2\sigma \]

\[ \tilde{\kappa}_{t+1}(1) = \tilde{\kappa}_t + \mu + \sigma \]

\[ \tilde{\kappa}_{t+1}(0) = \tilde{\kappa}_t + \mu \]

\[ \tilde{\kappa}_{t+1}(-1) = \tilde{\kappa}_t + \mu - \sigma \]

\[ \tilde{\kappa}_{t+1}(-2) = \tilde{\kappa}_t + \mu - 2\sigma \]

\[ \tilde{\kappa}_{t+1}(-3) = \tilde{\kappa}_t + \mu - 3\sigma \]

\[ \vdots \]

Figure 4: The discretized state space for \( \tilde{\kappa}_{t+1} \) given \( \tilde{\kappa}_t \).

model in Section 3.1 cannot be applied directly here, as it allows the mortality index \( q_t \) to take any non-negative real value.

To solve this problem, we construct in this section a mortality tree that is based on the stochastic mortality model in Section 3.1. The tree, which models the evolution of \( \kappa_t \) over time, is composed of two smaller multinomial trees, one for \( \tilde{\kappa}_t \) and the other for \( N_tY_t \). Given \( \kappa_t \), the mortality rates \( m_{x,t} \) and hence the mortality index \( q_t \) for \( t > 0 \) can be calculated straightforwardly.

First of all, note that the sequence \( \{ \tilde{\kappa}_{t+1} - \tilde{\kappa}_t \} \) is independent and identically distributed with mean \( \mu \) and variance \( \sigma^2 \). We discretize the state space for \( \tilde{\kappa}_{t+1} \) by assuming that in state \( i \), where \( i \in \{ \ldots, -2, -1, 0, 1, 2, \ldots \} \), the value of \( \tilde{\kappa}_{t+1} \) given \( \tilde{\kappa}_t \) is \( \tilde{\kappa}_{t+1}(i) = \tilde{\kappa}_t + \mu + i\sigma \). The discretization is illustrated diagrammatically in Figure 4.

We then assign a probability mass to each state. Specifically, the probability mass assigned to state \( j \) is the probability that, given \( \tilde{\kappa}_t \), \( \tilde{\kappa}_{t+1} \) lies within the interval of \( [\tilde{\kappa}_t + \mu + (i - 0.5)\sigma, \tilde{\kappa}_t + \mu + (i + 0.5)\sigma] \). It follows that the probability mass assigned to state \( i \) is

\[
\Pr[\tilde{\kappa}_{t+1} = \tilde{\kappa}_{t+1}(i) | \tilde{\kappa}_t] = \Phi(i + 0.5) - \Phi(i - 0.5).
\]

We then turn to \( N_{t+1}Y_{t+1} \). Note that the sequence \( \{N_{t+1}Y_{t+1} \} \) is also independent and identically distributed. If \( N_{t+1} = 0 \), we have \( N_{t+1}Y_{t+1} = 0 \); if \( N_{t+1} = 1 \), we have \( N_{t+1}Y_{t+1} = Y_{t+1} \), which follows a normal distribution with mean \( m \) and variance \( s^2 \).
\[
\begin{align*}
\text{Time} &= t + 1 \\
N_{t+1} &= 0 \\
&: \\
N_{t+1} &= 1, Y_{t+1}(3) = m + 3s \\
N_{t+1} &= 1, Y_{t+1}(2) = m + 2s \\
N_{t+1} &= 1, Y_{t+1}(1) = m + s \\
N_{t+1} &= 1, Y_{t+1}(0) = m \\
N_{t+1} &= 1, Y_{t+1}(-1) = m - s \\
N_{t+1} &= 1, Y_{t+1}(-2) = m - 2s \\
N_{t+1} &= 1, Y_{t+1}(-3) = m - 3s \\
&: 
\end{align*}
\]

Figure 5: The discretized state space for \( N_{t+1}Y_{t+1} \).

We discretize the state space for \( Y_{t+1} \) by assuming that in state \( j \), where \( j \in \{\ldots, -2, -1, 0, 1, 2, \ldots\} \), the value of \( Y_{t+1} \) is \( Y_{t+1}(j) = m + js \). The probability mass assigned to state \( j \) is the probability that \( Y_{t+1} \) lies within the interval of \([m + (j - 0.5)s, m + (j + 0.5)s]\). Hence, the state space for \( N_{t+1}Y_{t+1} \) contains a state that corresponds to \( N_{t+1} = 0 \) and states that correspond to \( N_{t+1} = 1 \) and \( Y_{t+1} = Y_{t+1}(j) \), where \( j \in \{\ldots, -2, -1, 0, 1, 2, \ldots\} \). The probability mass assigned to the state \( N_{t+1} = 0 \) is

\[ \Pr[N_{t+1} = 0] = 1 - p, \]

whereas the probability mass assigned to the state \( N_{t+1} = 1 \) and \( Y_{t+1} = Y_{t+1}(j) \) is

\[ \Pr[N_{t+1} = 1, Y_{t+1} = Y_{t+1}(j)] = p(\Phi(j + 0.5) - \Phi(j - 0.5)). \]

The discretization of \( N_{t+1}Y_{t+1} \) is illustrated diagrammatically in Figure 5.

Combining the trees in Figures 4 and 5, we obtain a tree for \( \kappa_{t+1} \). If \( \tilde{\kappa}_{t+1} = \tilde{\kappa}_{t+1}(i) \), \( N_{t+1} = 1 \) and \( Y_{t+1} = Y_{t+1}(j) \), then we have \( \kappa_{t+1} = \tilde{\kappa}_{t+1}(i) + Y_{t+1}(j) \). The probability mass assigned to this state is

\[
\Pr[\kappa_{t+1} = \tilde{\kappa}_{t+1}(i) + Y_{t+1}(j), N_{t+1} = 1|\tilde{\kappa}_i] = p(\Phi(i + 0.5) - \Phi(i - 0.5))(\Phi(j + 0.5) - \Phi(j - 0.5)).
\]

On the other hand, if \( \tilde{\kappa}_{t+1} = \tilde{\kappa}_{t+1}(i) \) and \( N_{t+1} = 0 \), then we have \( \kappa_{t+1} = \tilde{\kappa}_{t+1}(i) \). The
probability mass assigned to this state is
\[
\Pr [\kappa_{t+1} = \tilde{\kappa}_{t+1}(i), N_{t+1} = 0|\tilde{\kappa}_t] = (1 - p)(\Phi(i + 0.5) - \Phi(i - 0.5)).
\]

The computational burden can be reduced by using a small (finite) number of states at each time step. Specifically, at each time step, we set the maximum and minimum states for \(\tilde{\kappa}_{t+1}\) to \(M\) and \(-M\), and the maximum and minimum states for \(Y_{t+1}\) to \(N\) and \(-N\), respectively. The probability masses beyond the truncation points \(M\), \(-M\), \(N\) and \(-N\) are assigned to the respective truncation points. This means that the total number of states at each time step is \((2M + 1)(2N + 2)\).

### 5.5 An Example

Let us revisit the illustrative mortality-linked security in Section 3.2. We now study how the price of the security may change if we allow trades between the economic agents after time 0. As in Section 3.2, it is assumed here that \(r = 3\%\), \(k^A = 1\), and \(k^B = 0.5\). The tâtonnement pricing process is implemented with a sequential decision process and the multinomial mortality tree in Section 5.4.

For computational reasons, the agents’ positions in the mortality-linked security can only take a finite number of values. In this example, we require \(\theta^A_t, \theta^B_t \in \{0, 0.05, 0.1, \ldots, 3.9, 3.95\}\) for all \(t\). Since \(\theta^A_t = \theta^A_{t-1} + a^A_{t-1}\), \(a^A_t\) must lie within the interval \([-3.95, 3.95]\) for all \(t\). The same interval also applies to \(a^B_t\).

Using 56 states (with \(M = 3\) and \(N = 3\)), the estimated price of the mortality-linked security is 1.0343. At this price, 2.1 units of the security would be traded. Figure 6 shows the demand and supply curves that are derived from the generalized pricing process. We observe that the supply is 0 when the price is below $1.0268, indicating that Agent A is not willing to sell the security for any price lower than $1.0268. Then the supply curve is strictly increasing until the price reaches $1.0359, after which the supply remains constant at 3.95 units. The demand curve is constant at a level of 3.95 units. The upper limit of 3.95 units is because we require \(\theta^A_t \in \{0, 0.05, 0.1, \ldots, 3.9, 3.95\}\) in our calculations. The demand

---

7In Section 3.2, we demonstrated that if trades beyond time 0 are not allowed, then 1.86 units of the illustrative security will be traded at time 0. Of course, when we allow trades after inception, a different quantity will be traded at time 0, but this quantity should not be too far from 1.86. Using 1.86 as a starting point, we believe that the range of \([0, 3.95]\) should be wide enough to encompass the equilibrium quantity in the case when trades beyond time 0 are permitted.
curve has an opposite trend. It is decreasing until the price reaches $1.0359, beyond which the demand is 0.

We then examine how the choice of $M$ and $N$ would affect the equilibrium. First, we set $N = 2$ and examine the impact of $M$. In Table 5 we show the estimated prices and numbers of unit traded for different choices of $M$ with $N$ fixed at 2. As we increase $M$, the values do not vary significantly, and we obtain a convergence at $M = 3$. Note that we need more computational power if we choose a larger $M$.

Next, we set $M = 3$ and examine the impact of $N$. In Table 6, we show the estimated prices and numbers of unit traded for different choices of $N$ with $M$ fixed at 3. As we increase $N$, the values do not vary significantly, and we obtain a convergence at $N = 3$. Again, more computational power would be required if a larger $N$ is used.

<table>
<thead>
<tr>
<th>$M$</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Price</td>
<td>1.0354</td>
<td>1.0352</td>
<td>1.0352</td>
</tr>
<tr>
<td>Units traded at time 0</td>
<td>2.2</td>
<td>2.25</td>
<td>2.25</td>
</tr>
</tbody>
</table>

Table 5: Estimated prices for different choices of $M$ when $N = 2$. 

Figure 6: Demand and supply curves at time 0.
### Table 6: Estimated prices for different choices of $N$ when $M = 3.$

<table>
<thead>
<tr>
<th>$N$</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Price</td>
<td>1.0352</td>
<td>1.0343</td>
<td>1.0343</td>
</tr>
<tr>
<td>Units traded at time 0</td>
<td>2.25</td>
<td>2.1</td>
<td>2.1</td>
</tr>
</tbody>
</table>

6 Conclusions and Further Research

In this paper we proposed an economic method for pricing mortality-linked securities. We discussed two versions of the method. The first version, which we described in Section 2, is simple and straightforward to implement. It is highly suitable for today’s market in which hedgers and investors may not find the liquidity to unwind their positions in a mortality-linked security. The second version, which was detailed in Section 5, permits the counterparties to unwind their positions in discrete time steps before maturity. Nevertheless, it requires more computational resources to implement, particularly if we divide the time-to-maturity into a large number of time steps.

The method we propose does not take actual market prices as given, therefore sparing us from the problems related to a lack of market price data. The advantage of not requiring market prices as input is particularly important when we price long-term longevity securities. It is not clear if annuity prices offer an adequate starting point, as it is difficult, if not impossible, to infer a pure longevity risk premium from an annuity contract. As of this writing, the BNP/EIB bond is the only long-term longevity security with pricing information available in the public domain. However, the bond did not actually trade, so the reliability of its announced price is quite questionable.

To date, most long-term longevity securities traded are bespoke securities. By a bespoke longevity security we mean the payoffs from the security are based on the actual number of survivors in the hedger’s portfolio. An example is the longevity swap agreed between Babcock International and Credit Suisse in 2009. Some financial analysts believes the dominance of bespoke securities will continue in the longevity risk market over the next year or two.\(^8\) Our pricing framework is ideal for pricing bespoke deals, as the payoff functions, $f_t$ and $g_t$, and other parameters can be adapted readily to suit the actual situations of the counterparties. Further, it does not require

---

the pricing information of other mortality-linked securities, which are most likely based on different reference populations.

For pension plans and annuity providers, the reason for trading mortality-linked securities is to hedge their longevity risk exposures. In an incomplete market like the current longevity risk market, a perfect hedge cannot be formed, but an approximate hedging strategy may be constructed on the basis of a hedging objective. For example, one may construct a hedging strategy to stabilize the variability of net cash flows over a certain period of time (Cairns et al., 2008; Coughlan, 2009; Li and Hardy, 2011). If one’s hedging objective is to maximize the expected utility at a certain future time, then the quantity \( \theta^4 \) in Section 2 can be viewed as the corresponding static hedging strategy. Further, if trades are permitted after time 0, then we may regard the actions \( a^A_0, a^A_1, \ldots, a^A_{T-1} \) in Section 5 as the corresponding dynamic hedging strategy.

In using the tâtonnement approach, we require the absolute risk aversion parameters for the agents involved. The absolute risk aversion for a firm depends heavily on the firm’s characteristics. For instance, Cozzolino and Kleinman (1982) suggest that the absolute risk aversion for an insurance company is inversely related to the sum of its capital funds and new premium income. In our numerical illustrations, both agents (the life insurer and the hedge fund) are hypothetical, and therefore it is impossible to estimate their risk aversion parameters. Note also that, other than the risk aversion parameters, we require information such as the insurer’s liability structure in order to derive a realistic price.

In practice, when the identities of the parties involved in the trade are known, we can estimate their risk aversion parameters. Cox et al. (2010) propose a method for estimating the absolute risk aversion of an annuity provider from the prices of the annuities it sells. We believe that a similar approach can also be used to estimate the risk aversion parameter for a life insurer. Other information required may also be found from, for example, the annual reports of the parties involved. The study of a specific trade is beyond the scope of this paper. In future research, it is warranted to conduct a case study of a public known deal, and compare the results from our pricing framework with the actual price and number of units traded.

Our work drills down into the very fundamental economic concepts: demand and supply. The pricing framework we proposed yields a pair of demand and supply curves (Figures 2 and 6), from which we can predict if there will be any trade between the counterparties. We demonstrated empirically that, for the mortality-linked security
we consider, a tâtonnement equilibrium exists and the price of the security is unique. Another possible avenue of future research is to examine the necessary conditions for the existence and uniqueness of an equilibrium in the tâtonnement process for pricing mortality-linked securities. The study of such conditions may give us a better understanding of some unsuccessful longevity securities, including the BNP/EIB longevity bond, which was announced in 2004 and withdrawn for redesign in the following year, and Goldman Sachs’ QxX index swap, which was launched in 2007 and withdrawn in late 2009.

Acknowledgments

The authors acknowledge the financial support from the Natural Science and Engineering Research Council of Canada.

References


**Appendix**

In this appendix, we consider an alternative set-up that is based on the agents’ lifetime utilities.

We let $c^A_t$, $t = 0, 1, \ldots, T$, be Agent A’s consumption at time $t$. No constraint is imposed on the amounts of consumption. We define the lifetime utility for Agent A by

$$
\sum_{t=0}^T e^{-\rho t} U^A(c^A_t) + e^{-\rho T} U^A(W^A_T),
$$

where $\rho$ is the rate at which future utilities are discounted, $W^A_T$ is the terminal wealth, and $e^{-\rho T} U^A(W^A_T)$ is the bequest valuation function.

We assume again exponential utility functions. For now, we do not permit trades after time 0. If each agent’s goal is to maximize its expected lifetime utility, then the optimization problem for Agent A can be formulated as

$$
\max_{\theta^A, \theta^A_0, c^A_1, \ldots, c^A_T} \mathbb{E} \left[ \sum_{t=0}^T e^{-\rho t} U^A(c^A_t) + e^{-\rho T} U^A(W^A_T) \right],
$$

36
and that for Agent B can be formulated in a similar fashion.

The equilibrium resulting from this alternative set-up is identical to that from the set-up described in the main text (which is based on the terminal utility only). To see why, we first examine the relation between $W^*_A$ and $W^A$, the terminal wealth of Agent A in the set-up described in the main text. We have

$$W^*_A = (\omega^A + \theta^A P) e^r T - \sum_{t=0}^{T} c^A_t e^{r(T-t)} - \theta^A \sum_{t=1}^{T} g_t(Q_t) e^{r(T-t)} - \sum_{t=1}^{T} f_t(Q_t) e^{r(T-t)},$$

and

$$W^A = (\omega^A + \theta^A P) e^r T - \theta^A \sum_{t=1}^{T} g_t(Q_t) e^{r(T-t)} - \sum_{t=1}^{T} f_t(Q_t) e^{r(T-t)},$$

which implies that $W^*_A = W^A - \sum_{t=0}^{T} c^A_t e^{r(T-t)}$. Then, the supply from Agent A in the alternative set-up can be written as

$$\text{argsup}_{\theta^A} \mathbb{E} \left[ \sum_{t=0}^{T} e^{-\rho t} U^A(c^A_t) + e^{-\rho T} U^A(W^*_A) \right] = \text{argsup}_{\theta^A} \mathbb{E} \left[ \sum_{t=0}^{T} e^{-\rho t} U^A(c^A_t) + e^{-\rho T} (1 - e^{-k^A(W^A - \sum_{t=0}^{T} c^A_t e^{r(T-t)})}) \right] = \text{argsup}_{\theta^A} \mathbb{E} \left[ \sum_{t=0}^{T} e^{-\rho t} U^A(c^A_t) - e^{-\rho T} e^{k^A \sum_{t=0}^{T} c^A_t e^{r(T-t)}} e^{-k^A W^A_t} \right] = \text{argsup}_{\theta^A} \left\{ \sum_{t=0}^{T} e^{-\rho t} U^A(c^A_t) + e^{-\rho T} e^{k^A \sum_{t=0}^{T} c^A_t e^{r(T-t)}} \mathbb{E} \left[ -e^{-k^A W^A_t} \right] \right\} = \text{argsup}_{\theta^A} \mathbb{E} \left[ U^A(W^A) \right],$$

which is exactly the same as Agent A’s supply in the set-up described in the main text.\(^9\) Similarly, we can prove that Agent B’s demands in both set-ups are the same. As a result, both set-ups yield the same tâtonnement equilibrium.

This interesting result also holds when we permit trades after time 0. In this case, if each agent’s goal is to maximize its expected lifetime utility, then the optimization

\(^9\)The second last step in the above calculations follows from the assumption that $c^A_t$, $t = 0, 1, \ldots, T$, are not constrained. This assumption implies that for any $t$, $c^A_t$ does not depend on $\theta^A$. 

37
problem for Agent A can be formulated as
\[
\max_{c_0^A, c_1^A, \ldots, c_T^A, a_0^A, a_1^A, \ldots, a_{T-1}^A} \mathbb{E} \left[ \sum_{t=0}^{T} e^{-\rho t} U^A(c_t^A) + e^{-\rho T} U^A(W_T^A) \right],
\]
and that for Agent B can be formulated in a similar fashion. When trades are allowed after time 0, we have
\[
W_T^A = \omega^A e^r T + \sum_{t=0}^{T-1} a_t^A P_t(s_t) e^{r(T-t)} - \sum_{t=1}^{T} \theta_t^A g_t(Q_t) e^{r(T-t)} - \sum_{t=1}^{T} f_t(Q_t) e^{r(T-t)},
\]
and
\[
W_T^{A^*} = \omega^A e^r T + \sum_{t=0}^{T-1} a_t^A P_t(s_t) e^{r(T-t)} - \sum_{t=0}^{T} c_t^A e^{r(T-t)} - \sum_{t=1}^{T} \theta_t^A g_t(Q_t) e^{r(T-t)} - \sum_{t=1}^{T} f_t(Q_t) e^{r(T-t)},
\]
which also implies \( W_T^A - \sum_{t=0}^{T} c_t^A e^{r(T-t)} \). Given a price process, the optimal actions for Agent A can be written as
\[
\arg\sup_{a_0^A, a_1^A, \ldots, a_{T-1}^A} \mathbb{E} \left[ \sum_{t=0}^{T} e^{-\rho t} U^A(c_t^A) + e^{-\rho T} U^A(W_T^A) \right].
\]
Using the fact that \( W_T^A - \sum_{t=0}^{T} c_t^A e^{r(T-t)} \), we can show easily that the optimal actions can also be expressed as
\[
\arg\sup_{a_0^A, a_1^A, \ldots, a_{T-1}^A} \mathbb{E} \left[ U^A(W_T^A) \right],
\]
which are exactly the optimal actions resulting from the set-up described in the main text. As Agent A’s optimal actions are unchanged, its supply is also unchanged. Similarly, we can prove that Agent B’s demands in both set-ups are the same. As a result, both set-ups yield the same tâtonnement equilibrium, even if we permit trades after time 0.

We emphasize that, however, if other utility functions are assumed and/or consumptions at different time points are subject to some constraints, then this property may no longer hold. The optimization entailed in the alternative set-up would then become a lot more complex.