Economic Pricing of Mortality-Linked Securities

by

Rui Zhou

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Actuarial Science

Waterloo, Ontario, Canada, 2012

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Rui Zhou

Abstract

In previous research on pricing mortality-linked securities, the no-arbitrage approach is often used. However, this method, which takes market prices as given, is difficult to implement in today's embryonic market where there are few traded securities. In particular, with limited market price data, identifying a risk neutral measure requires strong assumptions. In this thesis, we approach the pricing problem from a different angle by considering economic methods. We propose pricing approaches in both competitive market and non-competitive market.

In the competitive market, we treat the pricing work as a Walrasian tâtonnement process, in which prices are determined through a gradual calibration of supply and demand. Such a pricing framework provides with us a pair of supply and demand curves. From these curves we can tell if there will be any trade between the counterparties, and if there will, at what price the mortality-linked security will be traded. This method does not require the market prices of other mortality-linked securities as input. This can spare us from the problems associated with the lack of market price data.

We extend the pricing framework to incorporate population basis risk, which arises when a pension plan relies on standardized instruments to hedge its longevity risk exposure. This extension allows us to obtain the price and trading quantity of mortality-linked securities in the presence of population basis risk. The resulting supply and demand curves help us understand how population basis risk would affect the behaviors of agents. We apply the method to a hypothetical longevity bond, using real mortality data from different populations. Our illustrations show that, interestingly, population basis risk can affect the price of a mortality-linked security in different directions, depending on the properties of the populations involved.

We have also examined the impact of transitory mortality jumps on trading in a competitive market. Mortality dynamics are subject to jumps, which are due to events such as the Spanish flu in 1918. Such jumps can have a significant impact on prices of mortality-linked securities, and therefore should be taken into account in modeling. Although several single-population mortality models with jump effects

have been developed, they are not adequate for trades in which population basis risk exists. We first develop a two-population mortality model with transitory jump effects, and then we use the proposed mortality model to examine how mortality jumps may affect the supply and demand of mortality-linked securities.

Finally, we model the pricing process in a non-competitive market as a bargaining game. Nash's bargaining solution is applied to obtain a unique trading contract. With no requirement of a competitive market, this approach is more appropriate for the current mortality-linked security market. We compare this approach with the other proposed pricing method. It is found that both pricing methods lead to Pareto optimal outcomes.

Acknowledgements

I would like to take this opportunity to thank all the people who have in no small way made this thesis possible.

First and foremost, I deeply thank my supervisors, Dr. Johnny Li and Dr. Ken Seng Tan, who introduced me to actuarial science research and taught me how to become a good researcher over the years. Their advice and encouragement were always important guiding lights towards my personal and professional development.

Special thanks to other members of my thesis committee, Dr. Phelim Boyle, Dr. Adam Kolkiewicz, Dr. Richard MacMinn and Dr. Ken Vetzal, for giving their precious time and suggestions to my thesis. The discussions with them have been inspirational and fruitful.

Finally, my sincere thanks to my husband, Yougun Han, whose support and love enabled me to complete this work, and to my parents who have always believed in me and encouraged me to pursue my doctoral dream.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Background

The payouts from life insurers and pension providers depend on mortality rates, which are difficult to predict. One source of future uncertainty is the random fluctuations around the underlying survival distribution. According to the law of large numbers, this uncertainty can be diversified away by selling a large number of policies. The other source arises from the occurrence of unexpected events, such as wars, pandemics or medical breakthroughs. This risk is systematic, as it usually affects most of the policies in the same way. Therefore, the risk cannot be diversified and we require other methods to mitigate it. In the thesis, we focus on this risk.

When a war or pandemic breaks out, mortality rates are higher than expected and life insurers pay more life insurance benefits and thus suffer a loss. This is called mortality risk. Mortality risk features low frequency and high severity. When medical breakthrough happens, mortality rates are lower than expected, which means annuity and pension providers pay longer periods of annuity or retirement benefits. This is called longevity risk.

To hedge against the systematic uncertainty in mortality rates, several methods can be employed. Firstly, insurers may retain the risk and rely on a natural hedge formed with their books of life insurances and annuities. This might work, because mortality movements have opposite effects on these two blocks of business. However, natural hedging can be ineffective and cost prohibitive in many circumstances as Cox and Lin (2007) point out. For example, the maturities of the two blocks of business may differ significantly, prohibiting an insurer to form an effective natural hedge. It might also be impractical for an insurer to adjust its business composition to support an effective natural hedge.

Secondly, the insurer may transfer the risk to reinsurers utilizing traditional life reinsurance on a quota share or surplus basis. Unfortunately, reinsurers have capacity constraints. Taking too much risk impairs the solvency of reinsurers, and increases their possibilities of default. To protect insurers, credit derivatives, such as credit default swaps, might be written on the renisurers. However, as the probability of reinsurers' default increases, credit derivatives become more expensive. The cost of reinsurance becomes less affordable or prohibitively high. Therefore, credit derivatives may not completely solve the capacity constraints problem.

Thirdly, the insurers may transfer their risk to investors in capital market through securitization. Investors may long mortality or longevity risk exposures to diversity their portfolios or to purely earn risk premium. Investors can also be insurers bearing the opposite risk and acquiring the missing line of business to optimize its natural hedge. Securities traded for these purposes are called mortality-linked securities.

Mortality-linked securities are financial derivatives with payments linked to certain mortality or longevity indexes. They allow market participants to either take or hedge exposure to the longevity and mortality experience of a given population of individuals. Mortality-linked securities can be divided into two categories, mortality derivatives and longevity derivatives, which differ in their purposes. Mortality derivatives are usually short-term contracts that are designed to hedge against catastrophic mortality risk for life insurers. In contrast, longevity derivatives are usually long-term contracts with payments reflecting unexpected mortality improvements. They provide protection for annuity and pension providers against an unexpected increase in their policyholders' life expectancy. They may have different payout structures. For example, a q-forward has only one single payout, while a longevity bond has multiple payouts over time.

The first mortality derivative was issued by Swiss Re in 2003 through a Special Purpose Vehicle (SPV) called Vita Capital. The term of the bond is three years and the principal payment is related to the mortality index that is based on a weighted average of mortality of the populations of the US, UK, France, Italy and Switzerland. The principal is at risk, should the mortality index exceed 130% of the 2002 level. No principal will be paid, should the mortality exceed 150% of the 2002 level. Since mortality has been improving consistently over recent decades, a situation that breaches the low end of this range (130-150% of 2002) is very rare. In other words, the investors have great chances to obtain full principal payment. This mortality bond was very successful and Swiss Re issued a series of Vita Capital after that.

At the same time, the market of longevity derivative has been taking off slowly. The first longevity bond, designed in 2004 for UK life insurance companies and pension funds, was developed by BNP-Paribas and European Investment Bank (EIB). It is a 25-year bond providing coupons linked to a survival index, which is based on the mortality experience of a cohort of males in England and Wales aged 65 in 2003. Unfortunately, this new product has never reached the market. Nevertheless, longevity derivatives still receive strong interest from pension sectors. Credit Suisse First Boston (CSFB) announced the release of the Credit Suisse Longevity Index in 2005, the first index designed specifically to enable the structuring and settlement of longevity risk transfer instruments. In 2007, JP Morgan launched the LifeMetrics Index, an "international index designed to benchmark and trade longevity risk". In 2008, we saw the first transaction for longevity risk management executed by Lucida plc and Canada Life in the UK. In this transaction, Lucida plc entered a mortality forward rate contract called a "q-forward", the payoff of which was linked to the LifeMetrics longevity index for England and Wales. In 2010, Swiss Re Kortis launched the first successful Longevity Bond. This bond provides coverage to Swiss Re against the increase in the difference between mortality improvement for older UK males (England and Wales), ages 75-85, and younger US males, ages 55-65. Buyers of the bond receive fixed quarterly coupons and principal repayment at maturity. The principal repayment will be reduced when the difference between mortality improvements exceeds a predetermined value. Recently, the Life and Longevity Markets Association (LLMA) has been set up to promote the development of a liquid traded market in longevity and mortality related risk.

1.2 Issues in the Current Market

Although mortality-linked securities have received great interest from investors, insurers and pensioners, the market is still embryonic and growing slowly. Few products are available in the market and few transactions have taken place. Illiquidity has been the concern for potential players in the market and impedes the growth of the market. One typical example is the failure of BNP/EIB longevity bond in 2004. To improve liquidity, consistent standards, methodologies and benchmarks should be established for the market.

To date, most long-term longevity securities traded are bespoke (or indemnity trigger) 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. Thus, bespoke deals have very low liquidity. Securities linked to standardized indexes have potentially higher liquidity. The LifeMetrics Index launched by JP Morgan is one of such standardized indexes designed to develop a liquid market for mortality and longevity hedging. Standardized indexes may be favored by investors because they involve no moral hazard, and are more transparent and liquid. However, hedgers who use standardized securities are subject to population basis risk, since the mortality experience of the standardized index and the underwritten population of insurers or pensioners are different. Hedge effectiveness can be significantly compromised in some circumstances. Therefore, there is a need to conduct a comprehensive study of population basis risk involved in standardized hedges. Some work has been done in this area. For example, Coughlan et al. (2011) proposed a framework for analyzing hedge effectiveness and basis risk.

Pricing mortality-linked securities poses another challenge to the development of this market. Most researchers use no-arbitrage method to price mortality-linked securities. For example, Cairns et al. (2006) obtained risk neutral measures by adding market prices of risk parameters into a two factor mortality model; Lin and Cox (2005) applied the Wang transform; Li and Ng (2011) used a method called canonical valuation to identify a risk neutral measure. However, these methods require prices of other mortality-linked securities. In today's embryonic market, these prices are often unavailable. Even if some market price data are available, finding a risk neutral measure is still not trivial, because in an incomplete market, one needs to pick a risk-neutral measure out of many possible risk-neutral measures.

In addition, the accuracy of future mortality forecasting also concerns the market. Mortality model directly affects the pricing results of mortality-linked securities. If investors and hedgers cannot agree on it, the liquidity will certainly be jeopardized. Therefore, it is necessary to develop a mortality model suitable for the mortality-linked security market. In the past, mortality models are mostly designed for one population, such as the Lee-Carter model proposed by Lee and Carter (1992) and the CBD model proposed by Cairns et al. (2006). However, when standardized index is used in mortality-linked security, it is very possible that the population associated with hedger's exposure is different from that associated with the index. If a one-population model is used for each population separately, the correlation between the two populations cannot be modeled. The ignorance of the correlation will underestimate hedging effectiveness significantly and mislead hedgers. Two-population or multi-population mortality models that take into account correlation between populations are necessary in this case. Some multi-population models have been proposed recently. Carter and Lee (1992) introduced the joint-k model for two populations. Li and Lee (2005) proposed the augmented common factor model. Li and Hardy (2010) discussed the co-integrated Lee-Carter model. Cairns et al. (2011a) built a model based on two age-period-cohort models. Nevertheless, there is still room for improvement. Current multi-population models do not incorporate mortality jumps, which are caused by disruptive events, for example, wars and epidemics. They contribute significantly to the variations of mortality rates. Examples of mortality jumps are discussed in Chapter 4. It is important not to ignore mortality jumps in modeling and forecasting, because otherwise we can seriously understate the uncertainty surrounding a central mortality projection.

1.3 Existing Pricing Methods

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 will 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, λ_1 and λ_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 (2009) 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.

1.4 Objectives and Outlines of the Thesis

This thesis establishes economic pricing approaches for mortality-linked securities. An advantage of the economic approaches is that they do not require market prices of other mortality-linked securities. We consider the pricing in both competitive and non-competitive markets. The effect of introducing population basis risk and transitory mortality jumps on the trading of mortality-linked securities is studied in a competitive market. A comparison of security pricing in the two markets is provided and analyzed.

Chapter 2 develops an approach to price mortality-linked security in a competitive market. The method approaches the pricing problem from the fundamental economic concepts, supply and demand. The pricing works as a Walrasian tâtonnement process. Specifically, an imaginary auctioneer matches supply and demand by gradually adjusting the price of the security being traded. Supply and demand from the economic agents in the market are determined through maximizing their own expected utility at a certain future time. The pricing framework is first set up for one-period mortality-linked securities and is then extended to multi-period mortality-linked securities by utilizing a Markov decision process. Both settings are illustrated with a hypothetical mortality-linked security and mortality data from the US population.

Chapter 3 extends the tâtonnement approach to allow a mismatch between the population associated with the hedger's risk exposure and that of the security being priced. Combining the extension with the two-population age-period-cohort mortality model proposed by Cairns et al. (2011a), we examine the effect of population basis risk on the price and trades of a hypothetic longevity bond. The major driving forces of the behavior of hedgers and investors are identified. They are then used to explain how the effect of population basis risk is formed. The hedging strategy implied by the tâtonnement approach is also examined.

Chapter 4 proposes a two-population mortality model with transitory jump effects and studies the impact of incorporating mortality jumps on the trading of mortality risk in a competitive market. The proposed model takes mortality jumps into account. The incorporation of jumps allows us to better estimate the probability of having a catastrophic mortality deterioration. This is particularly important for pricing securities for hedging extreme mortality risk. The impact of mortality jumps on trading is examined through a numerical illustration. The pair of demand and supply curves provided by the pricing framework is studied to help us understand the effect of introducing mortality jumps on the behaviors of the counterparties.

Chapter 5 models a trade of mortality-linked securities by a Nash bargaining game. Compared to the pricing framework proposed in Chapter 2 and Chapter 3, it is more suitable for the current mortality/longevity risk market, since it does not require the assumption of a competitive market. We numerically compare the pricing result from the Nash bargaining game with that from the competitive equilibrium. A common property of these two solutions, Pareto optimality, is also investigated to gain further insights.

Chapter 6 concludes and discusses plans for future research.

Chapter 2

Pricing in a Competitive Market

2.1 Introduction

A fundamental question in the study of mortality-linked securities is how to place a value on them. Risk-neutral methods, which are often used in past research, require market prices of other similar securities, making them difficult to implement in today's embryonic market.

In this chapter, 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 idea of tâtonnement in an exchange economy was first proposed by Walras (1874).[1] A Walrasian tâtonnement process assumes that there exists a fictitious Walrasian auctioneer who matches supply and demand from different economic agents in a market with perfect information and no transaction costs. 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 de-

[1] This seminal work of Walras (1874) was translated to English by William Jaffe in 1954. The translated version is cited as Walras (1954).

mand. 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 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 was developed under the assumption of a competitive market in Walras (1874). 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[2] , 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 utility 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 chapter is organized as follows. In Section 2.2, we introduce the idea of a tâtonnement process. Given this idea, we set up our pricing framework

[2]In a competitive market, there should exist many agents on both supply and demand sides. We assume that there is a group of homogeneous agents on each side. We then use one agent to represent each group, and model the trade between the two representative agents.

in a single-period setting, and then extend it to a multi-period one. In Section 2.3, 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 2.4 we perform sensitivity tests on different assumptions we have made in the pricing process. In Section 2.5, 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. In Section 2.6, we study how the results would be different if the agents maximize their expected life time utilities instead of their expected terminal utilities. Finally, in Section 2.7, we conclude this chapter.

2.2 A Tâtonnement Approach

2.2.1 The Idea

A Walrasian tâtonnement process assumes that there exists a fictitious Walrasian auctioneer who matches supply and demand from different economic agents in a market with perfect information and no transaction costs. 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 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.

2.2.2 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. In this connection, 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 buys (or 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 payout may also be related with q_0, the mortality index of the reference population at time 0. Since it is known at time 0, we suppress it in the notation for brevity. The mortality-linked security and the life contingent liability are associated with the mortality experience of the same reference population. The other economic agent, Agent B, is an investor who might issue (or purchase) the mortality-linked security, possibly for earning a risk premium. We assume that Agent A and Agent B have homogeneous beliefs about the stochastic evolution of the mortality index.This assumption is made throughout this thesis.

The quantities Agent A and B are willing to purchase or sell at time 0 are θ^A and θ^B, respectively. Following the specification of a tâtonnement process, we suppose that there exists an imaginary auctioneer who cries a arbitrary price, say P, at the beginning. Given this price, Agents A and B then decide values for θ^A and θ^B , based on a certain criterion. θ^A and θ^B can be either positive or negative. A positive quantity means that the agent purchases the security, while a negative quantity means the agent sells the security. In this chapter, we assume that Agent A issues the mortality-linked security. Thus, $\theta^A \leq 0$ and $\theta^B \geq 0$. We also assume that the agents will choose a supply or demand of the security that will maximize their expected terminal utilities.

Let ω^A and ω^B be the initial wealths of Agents A and B, respectively. It is assumed that the wealth of each agent can only be used be invested in either the mortality-linked security or a bank account which yields a continuously compounding risk-free interest rate of r per annum. In the real world, the agents often have many more different types of assets to choose from. It will be interesting to add more asset types and analyze the diversification effect of mortality-linked securities for investors. However, the main purpose of this chapter is to develop pricing framework. Therefore, we consider the simplest case of only two asset types. 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 the proceeds and its initial wealth together 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 sold by the other agent.

As mentioned earlier, at time 0, each agent chooses a supply or demand of the security that maximizes its expected terminal utility. In terms of the notation defined above, given a price P, the chosen supply and demand, $\hat{\theta}^A$ and $\hat{\theta}^B$, can be formulated as follows:

$$\hat{\theta}^A = \underset{\theta^A}{\operatorname{argsup}}\, \mathbb{E}[U^A((\omega^A - \theta^A P)e^r + \theta^A g(q_1) - f(q_1))] \tag{2.1}$$

$$\hat{\theta}^B = \underset{\theta^B}{\operatorname{argsup}}\, \mathbb{E}[U^B((\omega^B - \theta^B P)e^r + \theta^B g(q_1))] \tag{2.2}$$

Note that $\hat{\theta}^A$ and $\hat{\theta}^B$ are functions of the price P. We suppress the argument of these functions for simplicity. Trades only take place at a price P at which $\hat{\theta}^A + \hat{\theta}^B = 0$. We call this price equilibrium price, and denote it by P^*.

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., $\hat{\theta}^A + \hat{\theta}^B > 0$), and vice versa. Mathematically, the $(i+1)$th update of the price can be expressed as

$$P^{(i+1)} = P^{(i)} + d_i, \quad i = 0, 1, \ldots, \tag{2.3}$$

where $P^{(0)}$ is the initial guess of the price, and d_i is the adjustment function that always has the same sign as the excess demand, $\hat{\theta}^A + \hat{\theta}^B$.

In our calculations, it is assumed that the auctioneer adjusts the price in a way linear to the excess demand. Specifically, we assume that $P^{(i+1)} = P^{(i)} + \vartheta |P^{(i)}|(\hat{\theta}^A + \hat{\theta}^B)$, where ϑ 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 ϑ has to be chosen carefully. If ϑ is too large, the changes in P, $\hat{\theta}^A$, and $\hat{\theta}^B$ in each iteration tend to be large. This may lead us to missing the equilibrium. In contrast, if ϑ 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 ϑ. We will revisit this problem in Section 2.4.

It is interesting to note that we are essentially solving the equation $Z(P) = \hat{\theta}^A + \hat{\theta}^B = 0$ 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

$$d_i = \left(-\frac{\partial \hat{\theta}^A}{\partial P}_{|P=P^{(i)}} - \frac{\partial \hat{\theta}^B}{\partial P}_{|P=P^{(i)}} \right)^{-1} (\hat{\theta}^A + \hat{\theta}^B).$$

Since we have $\frac{\partial \hat{\theta}^A}{\partial P} \leq 0$ and $\frac{\partial \hat{\theta}^B}{\partial P} \leq 0$, d_i always has the same sign as $\hat{\theta}^A + \hat{\theta}^B$.[3] 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 2.5, as that $\hat{\theta}^A$ and $\hat{\theta}^B$ are confined to a set of discrete values.

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, $\hat{\theta}^B$, and supply, $\hat{\theta}^A$, on the basis of the current estimate of the price and the optimizing criteria specified by equations (2.1) and (2.2).
3. Stop if $|\hat{\theta}^A + \hat{\theta}^B|$ is less than a tolerance level, say 10^{-4}. Otherwise, adjust the price using equation (2.3).
4. Repeat Steps 2 and 3.

We set the convergence tolerance value to 10^{-4}. Smaller tolerance values have also been experimented and the difference they make on the pricing is neglectable. Therefore, 10^{-4} is chosen considering the convergence speed.

The equilibrium price P^* is the price at which the algorithm terminates. There are two possible situations. We may obtain $P^* > 0$ and $\hat{\theta}^A + \hat{\theta}^B = 0$, which means

[3]The law of demand and supply implies that $\frac{\partial \hat{\theta}^A}{\partial P} \leq 0$ and $\frac{\partial \hat{\theta}^B}{\partial P} \leq 0$. To have Newton's method work, we require that the partial derivatives are not both zero.

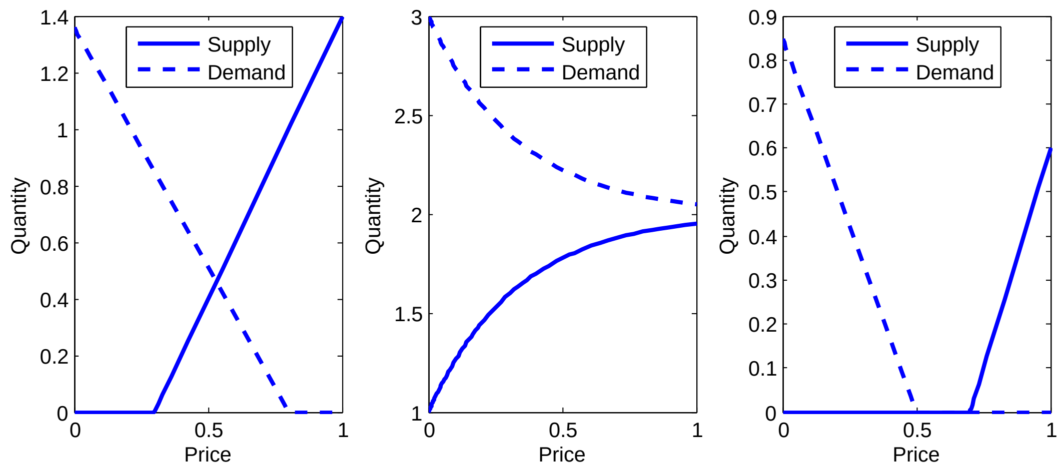


Figure 2.1: Possible situations when Algorithm 1 converges.

that the market attains equilibrium and trade happens between the economic agents (see Figure 2.1, the left panel). 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 2.1.

2.2.3 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 $Q_t = (q_0, q_1, \ldots, q_t)$, 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, per unit of which pays an amount of $g_t(Q_t)$ at time t, where g_t is a deterministic function of Q_t.[4] Agent B is 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 θ^A and the demand from Agent B is θ^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 we present in Section 2.5.

We keep all other assumptions in the single-period set-up. Let W_t^A and W_t^B 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

$$\begin{aligned}
W_0^A &= \omega^A \\
W_1^A &= (W_0^A - \theta^A P)e^r + \theta^A g_1(Q_1) - f_1(Q_1) \\
W_2^A &= W_1^A e^r + \theta^A g_2(Q_2) - f_2(Q_2) \\
\vdots \quad &\vdots \quad \vdots \\
W_T^A &= W_{T-1}^A e^r + \theta^A g_T(Q_T) - f_T(Q_T)
\end{aligned} \tag{2.4}$$

Agent B

$$\begin{aligned}
W_0^B &= \omega^B \\
W_1^B &= (W_0^B - \theta^B P)e^r + \theta^B g_1(Q_1) \\
W_2^B &= W_1^B e^r + \theta^B g_2(Q_2) \\
\vdots \quad &\vdots \quad \vdots \\
W_T^B &= W_{T-1}^B e^r + \theta^B g_T(Q_T)
\end{aligned} \tag{2.5}$$

Given a price P, each agent choose a demand or supply that maximizes its ex-

[4]We assume that f_t and g_t are functions of Q_t rather than q_t because the payouts can be path dependent.

pected terminal utility, as specified below,

$$\hat{\theta}^A = \underset{\theta^A}{\text{argsup}}\, \mathbb{E}[U^A(W_T^A)] \tag{2.6}$$

$$\hat{\theta}^B = \underset{\theta^B}{\text{argsup}}\, \mathbb{E}[U^B(W_T^B)] \tag{2.7}$$

Of course, in general, the initial guess of the price will not lead to $\hat{\theta}^A + \hat{\theta}^B = 0$. We may adjust the price by using Algorithm 1, with equations (2.1) and (2.2) replaced by (2.6) and (2.7), respectively.

2.3 An Illustration

2.3.1 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

$$\ln(m_{x,t}) = \beta_x^{(0)} + \beta_x^{(1)}\kappa_t + \epsilon_{x,t}, \tag{2.8}$$

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, κ_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 κ_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 μ and σ 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 κ_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 overestimating 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 (2.8), but permits transitory jumps in the evolution of κ_t. In particular, it models the time-varying factor κ_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 effect, while the second component, denoted by $N_t Y_t$, is designated for measuring jump effects.

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

$$N_t = \begin{cases} 1, & \text{with probability } p, \\ 0, & \text{with probability } 1-p, \end{cases} \tag{2.9}$$

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 κ_t can be expressed as follows:

$$\begin{cases} \tilde{\kappa}_{t+1} & = & \tilde{\kappa}_t + \mu + \sigma Z_{t+1}, \\ \kappa_{t+1} & = & \tilde{\kappa}_{t+1} + N_{t+1} Y_{t+1}, \end{cases} \tag{2.10}$$

where μ and σ 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 (2.10) are displayed in Table 2.1. The fitted values of $\beta_x^{(0)}$ and $\beta_x^{(1)}$ can be found on p.734 of Chen and Cox (2009).

Parameter	Estimate
μ	-0.2173
σ	0.3733
m	0.8393
s	1.4316
p	0.0436

Table 2.1: Estimates of parameters in the Lee-Carter model with transitory jump effects. (Source: 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 obtained. The extended Lee-Carter model, with parameters shown in Table 2.1, will be used in the tâtonnement pricing process.

2.3.2 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.[5] 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

[5] Source: NCHS report GMWK293R (http://www.cdc.gov/nchs/nvss/mortality/gmwk293r.htm).

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 loss_t is defined by

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

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 κ_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. Note that the assumption of exponential utility functions ensures a downward sloping demand curve[6]. 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.

[6]In the theory of consumer choice, price effect is a sum of income effect and substitution effect. Substitution effect means that the rate of consumption falls as the price of the good rises. It is always negative. Income effect means that as the wealth of the individual rises, demand increases, shifting the demand curve higher at all rates of consumption. Under the exponential utility function, there is no income effect. This will be shown later in the sensitivity test for initial wealths. Adding the substitution effect and income effect together, price effect is negative, and thus produces a downward sloping demand curve.

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 $\hat{\theta}^B$ and supply $\hat{\theta}^A$. This allows us to plot a curve of $\hat{\theta}^B$ against P (the demand curve) and a curve of $\hat{\theta}^A$ against P (the supply curve). The resulting demand and supply curves are shown in Figure 2.2. 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.

2.3.3 The Choice of Utility Function

Exponential utility functions are used in this thesis. Though power utility and other utility functions exhibiting a decreasing absolute risk aversion is considered more plausible, exponential utility is particularly convenient for many calculations. For example, it allows the agents' wealth to be negative. In contrast, power utility functions require both agents to have positive wealths at all times. This condition is not satisfied, because with uncertain future mortality, the payouts from mortality-linked security and life-contingent liability are also uncertain. Later in this chapter, we will also see that exponential utility functions facilitate the optimization problem.

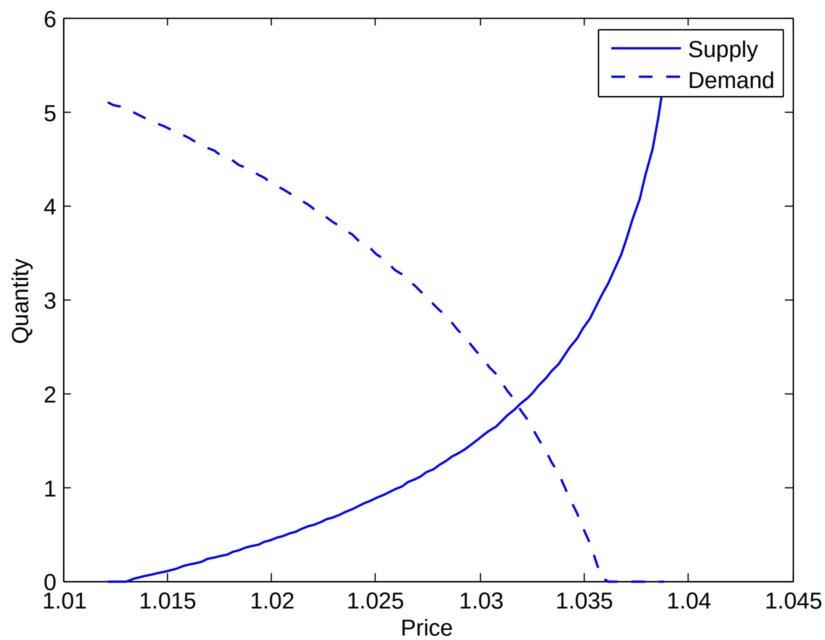


Figure 2.2: Supply and demand curves at time 0.

Linear utility functions are more tractable than exponential utility functions. However, they should not be used here, because the resulting equilibrium would not be meaningful. Using linear utility functions implies that the agents are risk-neutral. Mathematically, if linear utility function is assumed, we have

$$\begin{aligned} & \mathbb{E}[U^A(W_T^A)] \\ = & \mathbb{E}[W_T^A] \\ = & \theta^A \mathbb{E}\left[\sum_{t=1}^{T} g_t(Q_t)e^{-rt} - P\right]. \end{aligned}$$

The expected terminal utility function is a linear function of θ^A. Assuming Agent A is the supplier, i.e. $\theta^A \leq 0$, we have

$$\hat{\theta}^A = \begin{cases} 0, & \text{if } \; P < \mathbb{E}[\sum_{t=1}^{T} g_t(Q_t)e^{-rt}]; \\ -\infty, & \text{if } \; P > \mathbb{E}[\sum_{t=1}^{T} g_t(Q_t)e^{-rt}]; \\ (-\infty, 0], & \text{if } \; P = \mathbb{E}[\sum_{t=1}^{T} g_t(Q_t)e^{-rt}]. \end{cases}$$

Similarly, we can obtain that

$$\hat{\theta}^B = \begin{cases} \infty, & \text{if} \quad P < \mathbb{E}[\sum_{t=1}^T g_t(Q_t)e^{-rt}]; \\ 0, & \text{if} \quad P > \mathbb{E}[\sum_{t=1}^T g_t(Q_t)e^{-rt}]; \\ [0,\infty), & \text{if} \quad P = \mathbb{E}[\sum_{t=1}^T g_t(Q_t)e^{-rt}]. \end{cases}$$

The equilibrium can only be achieved at $P^* = \mathbb{E}[\sum_{t=1}^T g_t(Q_t)e^{-rt}]$. However, the trading at P^* makes no difference on the expected terminal wealths of the agents. This market will end up with no trade.

2.3.4 Initial Price Selection

In theory, the initial price $P^{(0)}$ can be an arbitrary value. However, an arbitrary $P^{(0)}$ may lead the price adjustment process to a dead end. Examples of this case are shown in the following discussion. To make sure that $P^{(0)}$ lies within a reasonable range, the lower bound and upper bound of price need to be determined.

Let v_L and v_H respectively be the accumulated values of the life contingent liabilities and the payouts from one unit of mortality-linked security at a future time T. Both v_L and v_H are positive real numbers. The terminal wealths for Agents A and B can be expressed as follows:

$$\begin{aligned} W_T^A &= \omega^A e^{rT} - v_L + \theta^A(v_H - Pe^{rT}), \\ W_T^B &= \omega^B e^{rT} + \theta^B(v_H - Pe^{rT}). \end{aligned}$$

Agent A and Agent B make decisions in order to maximize their own expected terminal utilities. Recall that we work under the assumption of exponential utility functions for both agents. Through simple calculations, we can rewrite equation (2.6) and equation (2.7) as below,

$$\hat{\theta}^A = \operatorname*{argsup}_{\theta^A} \mathbb{E}[U^A(W_T^A)] = \operatorname*{arginf}_{\theta^A} \mathbb{E}\left[e^{-k^A[-v_L+\theta^A(v_H-Pe^{rT})]}\right] \tag{2.11}$$

$$\hat{\theta}^B = \operatorname*{argsup}_{\theta^B} \mathbb{E}[U^B(W_T^B)] = \operatorname*{arginf}_{\theta^B} \mathbb{E}\left[e^{-k^B\theta^B(v_H-Pe^{rT})}\right] \tag{2.12}$$

The expectations involved are evaluated by Monte Carlo simulations. We simulate 50 000 mortality paths and thus there are 50 000 simulated values for v_H and v_L.

If Pe^{rT} is smaller than all simulated values of v_H, $v_H - Pe^{rT}$ in equation (2.11) and (2.12) is always positive. In this case, the terminal utilities for Agent A and B are increasing functions of θ^A and θ^B. The expected terminal utilities for Agent A and Agent B are maximized when θ^A and θ^B take their maximum values respectively. If Agent A is the supplier and Agent B is the demander, the optimal supply and demand are $\hat{\theta}^A = 0$ and $\hat{\theta}^B = \infty$. If Agent A is the demander and Agent B is the supplier, the optimal demand and supply are $\hat{\theta}^A = \infty$ and $\hat{\theta}^B = 0$. No trade will take place if Pe^{rT} is less than or equal to the minimum value of all simulated v_H. Therefore,

$$P^* > \frac{\min\text{ (simulated } v_H)}{e^{rT}} \geq \frac{\min\ (v_H)}{e^{rT}}.$$

Similarly, if Pe^{rT} is larger than all the simulated values of v_H, then $v_H - Pe^{rT}$ is always negative. The expected terminal utilities for Agent A and Agent B are maximized when θ^A and θ^B take their minimum values respectively. If Agent A is the supplier and Agent B is the demander, the optimal supply and demand are $\hat{\theta}^A = -\infty$ and $\hat{\theta}^B = 0$. If Agent A is the demander and Agent B is the supplier, the optimal demand and supply are $\hat{\theta}^A = 0$ and $\hat{\theta}^B = -\infty$. Therefore, no trade will happen if Pe^{rT} is larger than or equal to the maximum value of all simulated v_H. Therefore,

$$P^* < \frac{\max\text{ (simulated } v_H)}{e^{rT}} \leq \frac{\max\ (v_H)}{e^{rT}}.$$

Note that v_H, the accumulated value of mortality-linked security payouts, is often bounded. Its upper and lower bounds may be attained. As the number of simulated paths increase, the minimum and maximum of the simulated values of v_H may be very close or even equal to the theoretical minimum and maximum of v_H.

The lower bound and upper bound can also be understood intuitionally. If the price of a security is lower than the present value of its all possible future payments, there exists an arbitrage opportunity because agents in the market can earn free money by longing the security. Therefore, the demand will be infinitely large and no one is willing to supply. If the price of a security is larger than the present value of its all possible future payments, there also exists an arbitrage opportunity because agents in the market can earn free money by shorting the security. Therefore, the supply will be infinitely large and no one is willing to purchase.

Consider the case that $P^{(0)}$ is a value lower than the lower bound. Assuming Agent A is the supplier and Agent B is the demander, we have $\hat{\theta}^A = 0$ and $\hat{\theta}^B = \infty$, given $P^{(0)}$. Using equation 2.3, we will obtain $P^{(1)} = \infty, P^{(2)} = -\infty, P^{(3)} = \infty, \ldots$. No convergence will be achieved. Therefore, $P^{(0)}$ needs to be higher than the lower bound to ensure the convergence. Due to the same reason, $P^{(0)}$ needs to be smaller than the upper bound. In fact, when $P^{(i)}$ is out of the bounds for any $i = 0, 1, 2, \ldots$, there is no convergence. As a result, we shall use a small enough ϑ that will not cause $P^{(i)}$ to be out of the bounds in the implementation of Algorithm 1.

2.3.5 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 (2.10) individually to obtain the following jump mortality process in a risk-neutral probability measure:

$$\begin{cases} \tilde{\kappa}^*_{t+1} &= \tilde{\kappa}^*_t + \mu + \sigma Z^*_{t+1}, \\ \kappa^*_{t+1} &= \tilde{\kappa}^*_t + \mu + N^*_{t+1} Y^*_{t+1}, \end{cases}$$

λ_1	5.1449	0	1.5000
λ_2	0	3.4808	1.5000
λ_3	0	0	1.5000
Price	\$0.1318	\$0.9384	\$0.4976

Table 2.2: Prices of the illustrative mortality-linked security implied by different market prices of risk. (The values of λ_1, λ_2 and λ_3 are obtained from Chen and Cox (2009).)

where $Z_t^* \sim \mathrm{N}(\lambda_1, 1)$, $Y_t^* \sim \mathrm{N}(m + \lambda_2 s, s^2)$,

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

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

In the above, the unknown constants λ_1, λ_2, and λ_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 λ_1, λ_2, and λ_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 λ_1 when $\lambda_2 = \lambda_3 = 0$ and the value of λ_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 2.2 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 2.2. 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 2.2 clearly suggest the estimated price of the security is highly sensitive to the assumed values of λ_1, λ_2, and λ_3. For instance, if we assume $\lambda_2 = \lambda_3 = 0$, the estimated price would be as small as \$0.1318. The 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 2.2, 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 λ_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 λ_3 is also associated with jumps. However, this set of market prices of risk is not provided by Chen and Cox (2009).

2.4 Sensitivity Tests/Comparative Statics

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.

2.4.1 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. This convenient property of exponential utility functions was proven by Pratt (1964). Here, we verify that this property also holds in our set-up.

Proposition 1. *If exponential utility functions are assumed, ω_A and ω_B have no effect on the estimated price.*

Proof. We aim to prove that ω_A and ω_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

$$G_1(\theta^A, Q_1, Q_2, \dots, Q_T) = -\theta^A P e^{rT} + \sum_{i=0}^{T-1} e^{(T-i-1)r}[\theta^A g_{i+1}(Q_{i+1}) - f_{i+1}(Q_{i+1})]$$

and

$$G_2(\theta^B, Q_1, Q_2, \dots, Q_T) = -\theta^B P e^{rT} + \sum_{i=0}^{T-1} e^{(T-i-1)r}\theta^B g_{i+1}(Q_{i+1}).$$

Given a price P, the trading quantity that maximizes the terminal utility of Agent A is given by

$$\begin{aligned}
\hat{\theta}^A &= \operatorname*{argsup}_{\theta^A} \mathbb{E}[U^A(W_T^A)] \\
&= \operatorname*{argsup}_{\theta^A} \mathbb{E}[e^{-k^A W_T^A}] \\
&= \operatorname*{argsup}_{\theta^A} \mathbb{E}[e^{-k^A W_0^A e^{rT}} e^{-k^A G_1(\theta^A, Q_1, Q_2, \dots, Q_T)}] \\
&= \operatorname*{argsup}_{\theta^A} \mathbb{E}[e^{-k^A G_1(\theta^A, Q_1, Q_2, \dots, Q_T)}],
\end{aligned} \tag{2.13}$$

which is free of W_0^A and hence ω^A. Similarly, given P, the trading quantity that maximizes the terminal utility of Agent B is

$$\begin{aligned}
\hat{\theta}^B &= \operatorname*{argsup}_{\theta^B} \mathbb{E}[U^B(W_T^B)] \\
&= \operatorname*{argsup}_{\theta^B} \mathbb{E}[e^{-k^B G_2(\theta^B, Q_1, Q_2, \dots, Q_T)}],
\end{aligned}$$

k^A	k^B	Price	Units traded
1	0.25	\$1.0337	2.31
1	0.5	\$1.0319	1.86
1	1	\$1.0289	1.35
1	2	\$1.0250	0.87
0.5	0.5	\$1.0331	1.41
2	0.5	\$1.0306	2.23
4	0.5	\$1.0299	2.44

Table 2.3: Prices and numbers of unit traded when different risk aversion parameters are assumed.

which is free of W_0^B and hence ω^B. Since both initial wealths have no effect on the demand and supply curves, they do not affect the estimated price of the security. $\square$

2.4.2 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. In order to examine the impact of k^B on the price and trading quantity, we consider four cases in which k^A is fixed at 1 and the base value of k^B is multiplied by $\frac{1}{2}$, 1, 2 and 4, respectively. The comparison of the four cases helps us understand the impact of k^B. We use a similar approach to examine the impact of k^A on the trading. The estimated prices and the corresponding numbers of unit traded are displayed in Table 2.3.

Recall that a larger risk aversion parameter means that the agent is more risk adverse. From Table 2.3 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

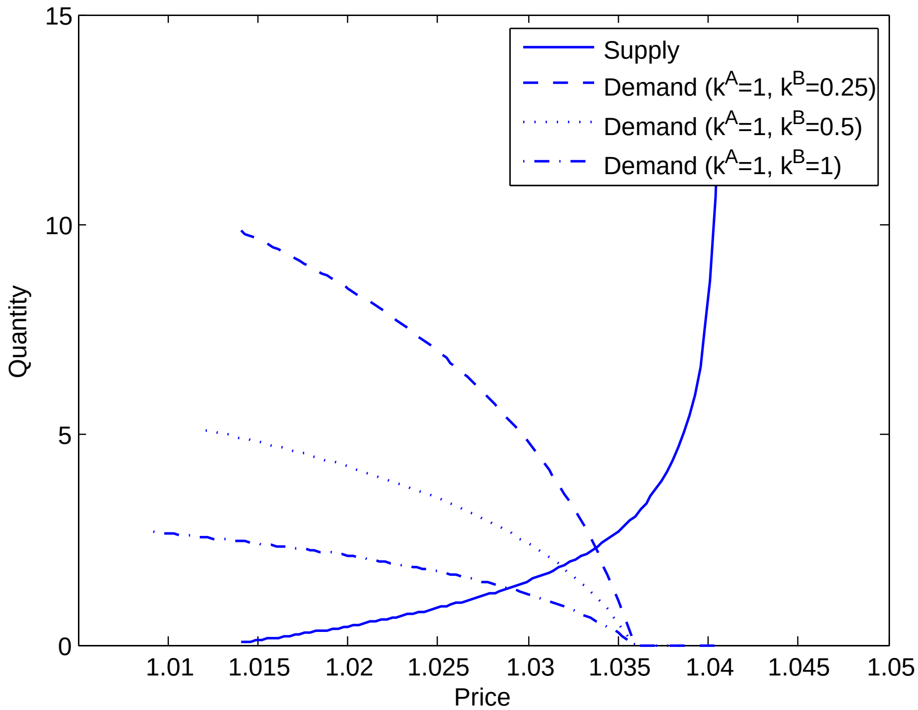


Figure 2.3: Supply and demand curves under different values of k^A.

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 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. Figure 2.3 and 2.4 show how the supply and demand curves move when we change the value of risk aversion parameters. In Figure 2.3, we keep k^A unchanged. Therefore, the supply curve does not move. Similarly, the demand curve does not move in Figure 2.4 because k^B is unchanged.These two graphical illustrations are in line with the results shown in Table 2.3.

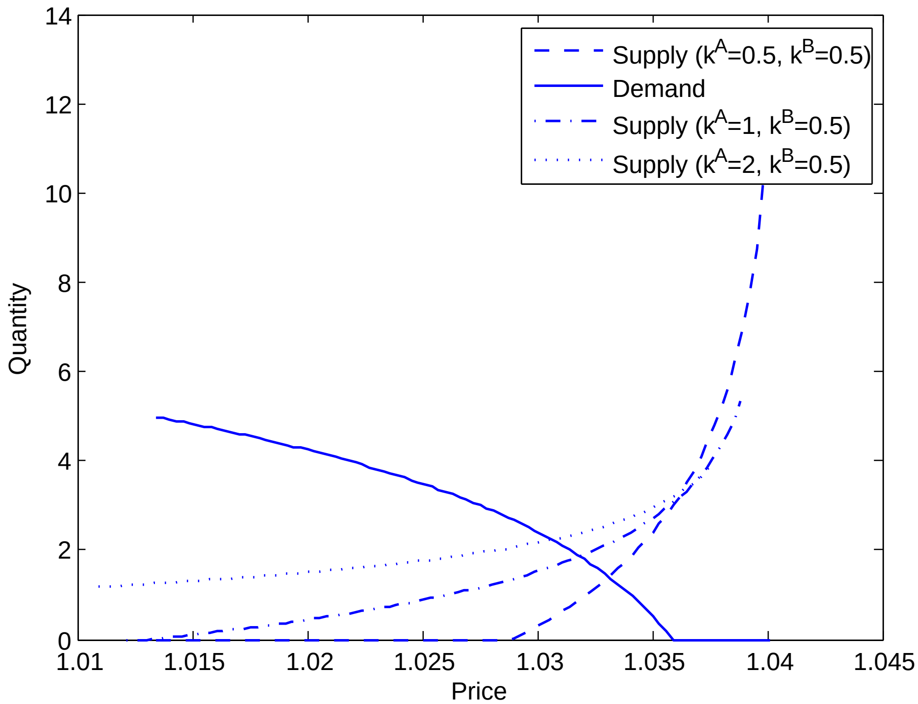


Figure 2.4: Supply and demand curves under different values of k^B.

2.4.3 The Price Adjustment Process

In Section 2.2.2, we highlighted the role of ϑ in the price adjustment process. The value of ϑ has to be small enough. Otherwise, the price adjustment might be too big so that the equilibrium price will be missed. If ϑ 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 ϑ. If we set $\vartheta = 0.001$, the price of mortality-linked security can be found in 13 iterations. In contrast, it we set $\vartheta = 0.005$, the algorithm does not converge but goes into a dead loop. The price adjustment processes on the basis of these two values of ϑ are illustrated graphically in Figure 2.5.

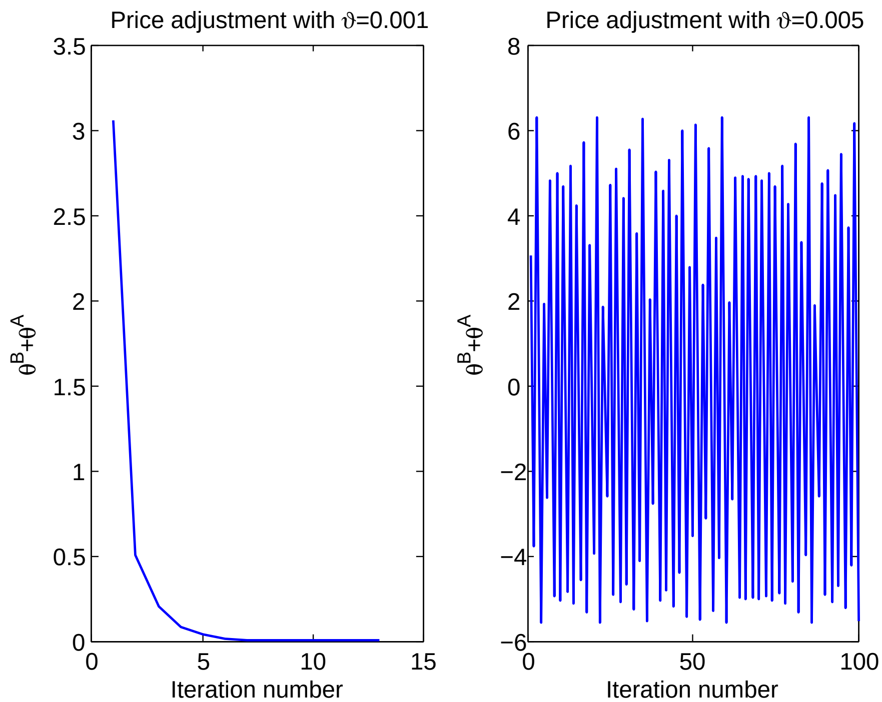


Figure 2.5: The price adjustment processes when $\vartheta = 0.001$ and $\vartheta = 0.005$.

2.4.4 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 2.4 we show the estimated prices of the illustrative mortality-linked security under different assumed risk-free interest rates.

Table 2.4 indicates that the equilibrium price and quantity traded are not quite 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. The supply and demand curves with assumptions of different risk-free interest rates

r	0.03	0.04	0.05	0.06
Price	\$1.0319	\$1.0303	\$1.0285	\$1.0264
Units traded	1.86	1.88	1.90	1.92

Table 2.4: Prices and numbers of unit traded when different risk-free interest rates are assumed.

are shown in Figure 2.6. This figure is consistent with the results in Table 2.4.

2.5 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 2.2.3 does not permit the economic agents to trade the mortality-linked security after time 0. In this section we will generalize that set-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.

2.5.1 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 $\mathcal{A}_{s_t}$, the set of allowable actions in state s_t. At time t,

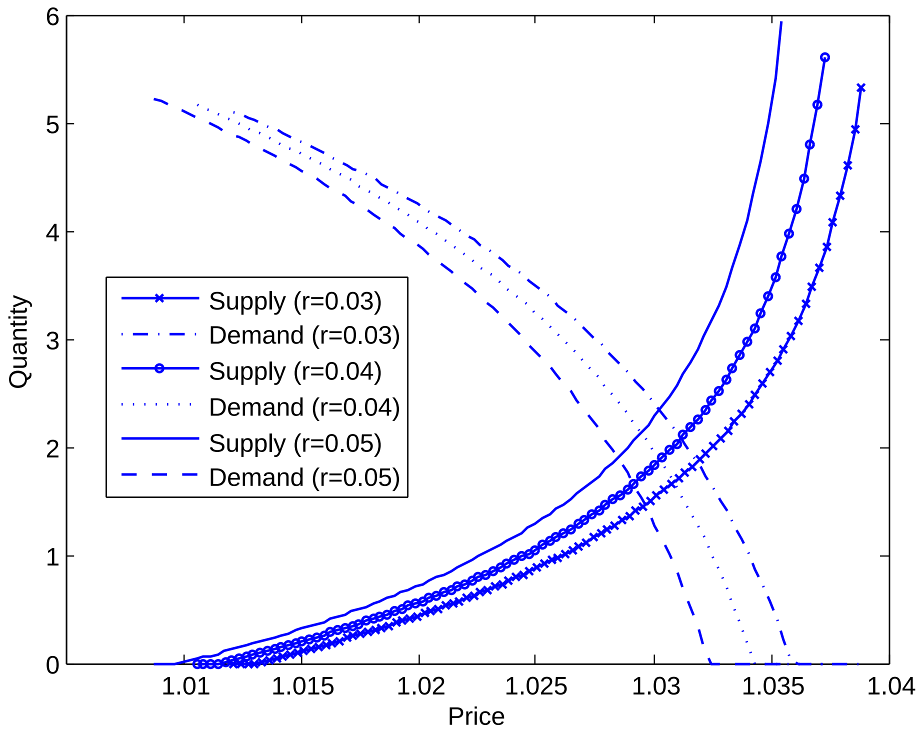


Figure 2.6: Supply and demand curves when different risk-free interest rates are assumed.

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, which specifies an action selection procedure in each state s_t at time t. As a result of choosing action $a_t \in \mathcal{A}_{s_t}$, 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 $r_T(s_T)$.

Since the rewards received by the decision maker are not known before any action is taken, the reward sequence is random. The decision maker's objective is to choose a sequence of actions so that the corresponding random reward sequence is as 'large' as possible. This necessitates a method for finding an optimal action sequence, which we will investigate in Section 2.5.2.

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 $H_0 = S_0, H_1 = S_0 \times \mathcal{A}_{S_0} \times S_1$, and $H_t = S_0 \times \mathcal{A}_{S_0} \times S_1 \times \mathcal{A}_{S_1} \times \ldots \times S_t$.[7] Note also that the recursion $H_t = H_{t-1} \times \mathcal{A}_{S_{t-1}} \times S_t$ 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 $\mathcal{A}_{h_t}$. The decision process is no longer 'Markov,' because the decision rule d_t maps the entire history H_t to $\mathcal{A}_{H_t}$.

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.

2.5.2 An Optimal Action Sequence

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

> *Bellman's principle of optimality: An optimal policy has the property that whatever the initial state and initial decision are, the remaining decisions*

[7] For sets C and D, $C \times D$ denotes the Cartesian product of C and D; that is, $C \times D = \{(c, d) : c \in C, d \in D)\}$.

must constitute an optimal policy with regard to the state resulting from the first decision.

Suppose that the objective is to maximize the sum of rewards from time 1 to T, that is, $\sum_{i=1}^{T-1} r_i(s_i, a_i) + r_T(s_T)$. For a Markov sequential decision process, the principle of optimality implies that $(a_0, a_1, \ldots, a_{T-1})$ is the optimal action sequence if

$$a_t = \underset{a \in \mathcal{A}_{s_t}}{\operatorname{argsup}} \{r_t(s_t, a) + \mathbb{E}[v_{t+1}(s_{t+1})|s_t, a]\},$$

where

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

When $\mathcal{A}_{S_t}$ is finite, we can replace the 'sup' in the equation above by 'max.' A mathematical 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

$$a_t = \underset{a \in \mathcal{A}_{s_t}}{\operatorname{argsup}} \{r_t(s_t, a) + \mathbb{E}[v_{t+1}(s_{t+1})|s_t, a]\} \tag{2.14}$$

and

$$v_t(s_t) = r_t(s_t, a_t) + \mathbb{E}[v_{t+1}(s_{t+1})|s_t, a_t], \tag{2.15}$$

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 (2.14) and (2.15). Set $t = t - 1$.

3. Repeat Step 2. Stop when $t = 0$.

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.

2.5.3 The Generalized Pricing Process

Let θ_t^A be Agent A's position in the mortality-linked security immediately before trade at time t. Denote by a_t^A the action taken by Agent A at time t; that is, $-a_t^A$ is the number of units of the security sold at time t. It is easy to see that $\theta_0^A = 0$ and $\theta_t^A = \theta_{t-1}^A + a_{t-1}^A$.

We define similar notation for Agent B. We let θ_t^B be Agent B's position in the security immediately before trade at time t, and let a_t^B be the number of units of the security purchased at time t. We have $\theta_0^B = 0$ and $\theta_t^B = \theta_{t-1}^B + a_{t-1}^B$.

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 depend 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

$$\begin{aligned}
W_0^A &= \omega^A \\
W_1^A &= (W_0^A - a_0^A P_0(\hat{s}_0))e^r + (\theta_0^A + a_0^A)g_1(Q_1) - f_1(Q_1) \\
W_2^A &= (W_1^A - a_1^A P_1(\hat{s}_1))e^r + (\theta_1^A + a_1^A)g_2(Q_2) - f_2(Q_2) \\
\vdots &\ \vdots\ \vdots \\
W_T^A &= (W_{T-1}^A - a_{T-1}^A P_{T-1}(\hat{s}_{T-1}))e^r + (\theta_{T-1}^A + a_{T-1}^A)g_T(Q_T) - f_T(Q_T)
\end{aligned} \tag{2.16}$$

Agent B

$$\begin{aligned}
W_0^B &= \omega^B \\
W_1^B &= (W_0^B - a_0^B P_0(\hat{s}_0))e^r + (\theta_0^B + a_0^B)g_1(Q_1) \\
W_2^B &= (W_1^B - a_1^B P_1(\hat{s}_1))e^r + (\theta_1^B + a_1^B)g_2(Q_2) \\
\vdots &\ \vdots\ \vdots \\
W_T^B &= (W_{T-1}^B - a_{T-1}^B P_{T-1}(\hat{s}_{T-1}))e^r + (\theta_{T-1}^B + a_{T-1}^B)g_T(Q_T)
\end{aligned} \tag{2.17}$$

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 $U^A(x) = 1 - e^{-k^A x}$ and $U^B(x) = 1 - e^{-k^B x}$.

We identify $v_t(s_t)$ in equation (2.15) 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 $v_T^B(s_T) = U^B(W_T^A)$.

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

$$\operatorname*{argsup}_{a_0, a_1, \ldots, a_{T-1}} \mathbb{E}[U^A(W_T^A)].$$

Its optimal action at time $T-1$ is

$$\begin{aligned} a_{T-1}^A &= \operatorname*{argsup}_{a \in A_{\hat{s}_{T-1}}} \mathbb{E}[U^A((W_{T-1}^A - aP_{T-1}(\hat{s}_{T-1}))e^r + (\theta_{T-1}^A + a)g_T(Q_T) - f_T(Q_T))|\hat{s}_{T-1}] \\ &= \operatorname*{arginf}_{a \in A_{\hat{s}_{T-1}}} \mathbb{E}[e^{-k^A(-aP_{T-1}(\hat{s}_{T-1})e^r + (\theta_{T-1}^A + a)g_T(Q_T) - f_T(Q_T))}|\hat{s}_{T-1}], \end{aligned}$$

which means the optimal action a_{T-1}^A depends on the past information only through Q_{T-1} and θ_{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 θ_t^A as well. Therefore, we may reduce the state variable to $\hat{s}_t = (Q_t, \theta_t^A)$, because that would reduce the content of the state variable and hence reduce the computational effort needed.

With Algorithm 2 and the wealth processes specified by equations (2.16) and (2.17), 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 2.2. 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_{T-1}^A and a_{T-1}^B and the values of $v_{T-1}^A(s_{T-1})$ and $v_{T-1}^B(s_{T-1})$. Then, we repeat the procedure for times $T-2$, $T-3$, ..., 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 $v_T^A(\hat{s}_T)$ and $v_T^B(\hat{s}_T)$. 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 (2.14) and (2.15) 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 $P_0(\hat{s}_0)$.

2.5.4 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 model in Section 2.3.1 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 2.3.1. The tree, which models the evolution of κ_t over time, is composed of two smaller multinomial trees, one for $\tilde{\kappa}_t$ and the other for $N_t Y_t$. Given κ_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 μ and variance σ^2. We discretize the state space for $\tilde{\kappa}_{t+1}$ by assuming that in state i, where $i \in \{\ldots,\text{-}2,\text{-}1,0,1,2,\ldots\}$, 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 2.7.

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

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

$$
\begin{array}{cc}
\text{Time} = t & \text{Time} = t+1 \\
& \vdots \\
& \tilde{\kappa}_{t+1}(3) = \tilde{\kappa}_t + \mu + 3\sigma \\
& \tilde{\kappa}_{t+1}(2) = \tilde{\kappa}_t + \mu + 2\sigma \\
& \tilde{\kappa}_{t+1}(1) = \tilde{\kappa}_t + \mu + \sigma \\
\tilde{\kappa}_t & \tilde{\kappa}_{t+1}(0) = \tilde{\kappa}_t + \mu \\
& \tilde{\kappa}_{t+1}(-1) = \tilde{\kappa}_t + \mu - \sigma \\
& \tilde{\kappa}_{t+1}(-2) = \tilde{\kappa}_t + \mu - 2\sigma \\
& \tilde{\kappa}_{t+1}(-3) = \tilde{\kappa}_t + \mu - 3\sigma \\
& \vdots
\end{array}
$$

Figure 2.7: The discretized state space for $\tilde{\kappa}_{t+1}$ given $\tilde{\kappa}_t$.

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.

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 2.8.

Combining the trees in Figures 2.7 and 2.8, we obtain a tree for κ_{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

$$
\begin{gathered}
\text{Time} = t+1 \\
N_{t+1} = 0 \\
\vdots \\
N_{t+1} = 1, Y_{t+1}(3) = m+3s \\
N_{t+1} = 1, Y_{t+1}(2) = m+2s \\
N_{t+1} = 1, Y_{t+1}(1) = m+s \\
N_{t+1} = 1, Y_{t+1}(0) = m \\
N_{t+1} = 1, Y_{t+1}(-1) = m-s \\
N_{t+1} = 1, Y_{t+1}(-2) = m-2s \\
N_{t+1} = 1, Y_{t+1}(-3) = m-3s \\
\vdots
\end{gathered}
$$

Figure 2.8: The discretized state space for $N_{t+1}Y_{t+1}$.

probability mass assigned to this state is

$$
\begin{aligned}
& \Pr\left[\kappa_{t+1} = \tilde{\kappa}_{t+1}(i) + Y_{t+1}(j), N_{t+1} = 1 | \tilde{\kappa}_t\right] \\
= \; & p(\Phi(i+0.5) - \Phi(i-0.5))(\Phi(j+0.5) - \Phi(j-0.5)).
\end{aligned}
$$

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

$$
\Pr\left[\kappa_{t+1} = \tilde{\kappa}_{t+1}(i), N_{t+1} = 0 | \tilde{\kappa}_t\right] = (1-p)(\Phi(i+0.5) - \Phi(i-0.5)).
$$

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 $(2M+1)(2N+2)$.

2.5.5 An Example

Let us revisit the illustrative mortality-linked security in Section 2.3.2. 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 2.3.2, 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 2.5.4.

For computational reasons, the agents' positions in the mortality-linked security can only take a finite number of values. In this example, we require $\theta_t^A, \theta_t^B \in \{0, 0.05, 0.1, \ldots, 3.9, 3.95\}$ for all t.[8] Since $\theta_t^A = \theta_{t-1}^A + a_{t-1}^A$, a_t^A must lie within the interval $(-3.95, 3.95)$ for all t. The same interval also applies to a_t^B.

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 2.9 shows the demand and supply curves at time 0. They 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 upper limit of 3.95 units is because we require $\theta_t^A \in \{0, 0.05, 0.1, \ldots, 3.9, 3.95\}$ in our calculations. The demand 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 2.5 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.

[8] In Section 2.3.2, 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.

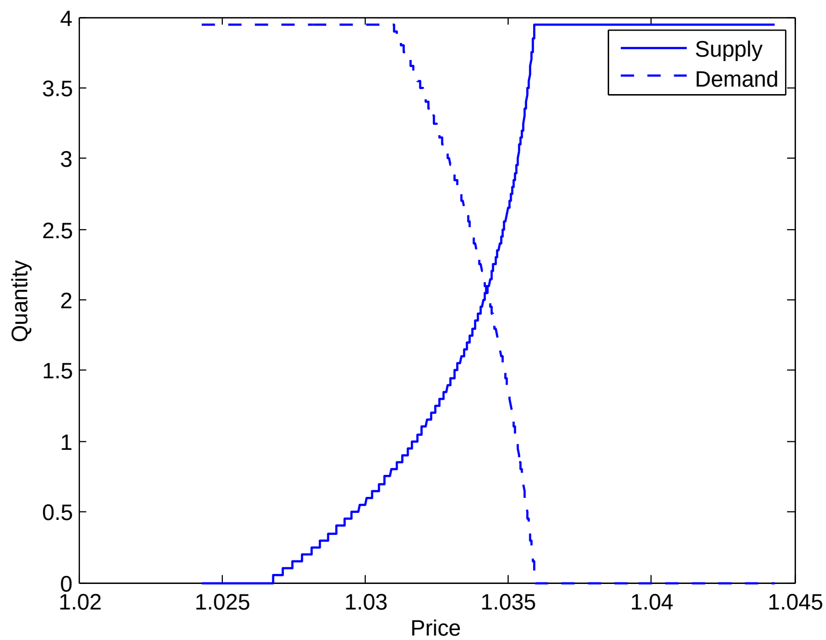


Figure 2.9: Demand and supply curves at time 0.

Next, we set $M = 3$ and examine the impact of N. In Table 2.6, 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.

2.6 Maximizing Expected Lifetime Utility

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.

Proposition 2. *Under the assumption of exponential utility functions, maximizing each agent's expected terminal utility and maximizing each agent's expected lifetime utility would lead to the same equilibrium price and quantity traded.*

M	2	3	4
Price	1.0354	1.0352	1.0352
Units traded at time 0	2.2	2.25	2.25

Table 2.5: Estimated prices for different choices of M when $N = 2$.

N	2	3	4
Price	1.0352	1.0343	1.0343
Units traded at time 0	2.25	2.1	2.1

Table 2.6: Estimated prices for different choices of N when $M = 3$.

Proof. We consider an alternative set-up that is based on the agents' lifetime utilities.

We let c_t^A, $t = 0, 1, \ldots, T$, 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

$$\sum_{t=0}^{T} e^{-\rho t} U^A(c_t^A) + e^{-\rho T} U^A(W_T^{A*}),$$

where ρ is the rate at which future utilities are discounted, W_T^{A*} is the terminal wealth, and $e^{-\rho T} U^A(W_T^{A*})$ 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

$$\sup_{\theta^A, c_0^A, c_1^A, \ldots, c_T^A} \mathbb{E}\left[\sum_{t=0}^{T} e^{-\rho t} U^A(c_t^A) + e^{-\rho T} U^A(W_T^{A*})\right],$$

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

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_T^{A*} and W_T^A, the terminal wealth of Agent A in the set-up described in the main text. We have

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

and

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

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

$$\begin{aligned}
& \operatorname*{argsup}_{\theta^A} \mathbb{E}\left[\sum_{t=0}^{T} e^{-\rho t} U^A(c_t^A) + e^{-\rho T} U^A(W_T^{A*})\right] \\
= \; & \operatorname*{argsup}_{\theta^A} \mathbb{E}\left[\sum_{t=0}^{T} e^{-\rho t} U^A(c_t^A) + e^{-\rho T}\left(1 - e^{-k^A(W_T^A - \sum_{t=0}^{T} c_t^A e^{r(T-t)})}\right)\right] \\
= \; & \operatorname*{argsup}_{\theta^A} \mathbb{E}\left[\sum_{t=0}^{T} e^{-\rho t} U^A(c_t^A) - e^{-\rho T} e^{k^A \sum_{t=0}^{T} c_t^A e^{r(T-t)}} e^{-k^A W_T^A}\right] \\
= \; & \operatorname*{argsup}_{\theta^A} \left\{\sum_{t=0}^{T} e^{-\rho t} U^A(c_t^A) + e^{-\rho T} e^{k^A \sum_{t=0}^{T} c_t^A e^{r(T-t)}} \mathbb{E}\left[-e^{-k^A W_T^A}\right]\right\} \\
= \; & \operatorname*{argsup}_{\theta^A} \mathbb{E}\left[-e^{-k^A W_T^A}\right] \\
= \; & \operatorname*{argsup}_{\theta^A} \mathbb{E}\left[U^A(W_T^A)\right],
\end{aligned}$$

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_t^A, $t = 0, 1, \ldots,$ T, are not constrained. This assumption implies that for any t, c_t^A does not depend on θ^A.

problem for Agent A can be formulated as

$$\sup_{c_0^A, c_1^A, \ldots, c_T^A, a_0^A, a_1^A, \ldots, a_{T-1}^A} \mathbb{E}\left[\sum_{t=0}^{T} e^{-\rho t} U^A(c_t^A) + e^{-\rho T} U^A(W_T^{A*})\right],$$

and that for Agent B can be formulated in a similar fashion. When trades are allowed after time 0, we have

$$W_T^A = \omega^A e^r T + \sum_{t=0}^{T-1} a_t^A P_t(\hat{s}_t) e^{r(T-t)} - \sum_{t=1}^{T} \theta_t^A g_t(Q_t) e^{r(T-t)} - \sum_{t=1}^{T} f_t(Q_t) e^{r(T-t)},$$

and

$$W_T^{A*} = \omega^A e^r T + \sum_{t=0}^{T-1} a_t^A P_t(\hat{s}_t) e^{r(T-t)} - \sum_{t=0}^{T} c_t^A e^{r(T-t)} - \theta^A \sum_{t=1}^{T} g_t(Q_t) e^{r(T-t)} - \sum_{t=1}^{T} f_t(Q_t) e^{r(T-t)},$$

which also implies $W_T^{A*} = W_T^A - \sum_{t=0}^{T} c_t^A e^{r(T-t)}$. Given a price process, the optimal actions for Agent A can be written as

$$\operatorname*{argsup}_{a_0^A, a_1^A, \ldots, a_{T-1}^A} \mathbb{E}\left[\sum_{t=0}^{T} e^{-\rho t} U^A(c_t^A) + e^{-\rho T} U^A(W_T^{A*})\right].$$

Using the fact that $W_T^{A*} = W_T^A - \sum_{t=0}^{T} c_t^A e^{r(T-t)}$, we can show easily that the optimal actions can also be expressed as

$$\operatorname*{argsup}_{a_0^A, a_1^A, \ldots, a_{T-1}^A} \mathbb{E}\left[U^A(W_T^A)\right],$$

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.

2.7 Concluding Remarks

In this chapter, 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 2.2, 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 2.5, 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.

Our work drills down into the very fundamental economic concepts: demand and supply. The pricing framework we contribute yields a pair of demand and supply curves (Figures 2.2 and 2.9), 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.

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.[10] 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. Moral hazard often appears in the bespoke deals, because the agents have asymmetric information. However, our pricing framework does not take moral hazard into account. It will be interesting to incorporate it in the framework and analyze its impact on the trading in the future research.

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, 2010). If one's hedging objective is to maximize the expected utility at a certain future time, then the quantity θ^A in Section 2.2 can be viewed as the corresponding static hedging strategy. Further, if trades are permitted after time 0, then we may regard the actions $a_0^A, a_1^A, \ldots, a_{T-1}^A$ in Section 2.5 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 other information, such as the insurer's liability structure, in order to derive a realistic price. As a price taker, the insurer decides supply/demand according to a given price and certain criterion. The liability structure of the insurer affects this decision, and thus affects the equilibrium.

[10]See "Pensions Ready for Longevity Solutions," *InsuranceERM*, May 2009.

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 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 chapter. 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.

In this chapter, we assumed that the two agents have homogeneous beliefs about future mortality dynamics. If this assumption is relaxed, our pricing framework can still be used. The only change is that each agent uses its own mortality model to determine supply or demand.

Chapter 3

Incorporating population basis Risk

3.1 Introduction

In Chapter 2, we introduced an economic approach for pricing mortality-linked securities. This new pricing method is developed from an idea called tâtonnement, which was first proposed by Walras (1874) to model trades in an exchange economy. It models the actual trade between the hedger and investor. It is therefore more transparent relative to standard no-arbitrage approaches, in which the price of a security is estimated by extrapolating prices of other similar securities available in the market. On top of the estimated price, the economic pricing method provides us with a pair of demand and supply curves, from which we can infer the quantity of a mortality-linked security to be traded in equilibrium. Further, by examining the response of the demand and supply curves to changes in, for example, the volatility of mortality rates, we can have a better idea about how and why the price of a security will change in different circumstances. Another appealing feature of the economic pricing method is that, as opposed to no arbitrage approaches, it works when there are no market prices of other mortality-linked securities[1]. This can spare us from the problems associated with the lack of market price data.

However, the pricing framework proposed in Chapter 2 assumes that the population of individuals associated with the hedger's risk exposure is identical to that associated with the security being priced. This simple set-up is appropriate for pricing bespoke mortality-linked securities, from which the payoffs are linked to the actual number of survivors in the hedger's own portfolio. Nevertheless, this set-up may not be adequate for valuing standardized mortality-linked instruments, which are usually based on broad population mortality indexes rather than the hedger's own mortality experience. More specifically, the mismatch in mortality movements would create what is referred to as population basis risk. In this chapter, we relax this assumption and extend the economic pricing method so that it can be applied in the presence of population basis risk. To achieve this objective, we integrate the two-population age-period-cohort mortality model proposed by Cairns et al. (2011a) into the economic pricing framework.

The existence of population basis risk may have a significant impact on the demand and/or supply of a mortality-linked security. The problem of population basis risk has recently been considered by some academics, including Cairns et al. (2011a), Coughlan et al. (2011), and Li and Hardy (2010). Their studies focus mainly on the measurement of basis risk, but have made no attempt to investigate how population basis risk may affect the trade and hence the price of a mortality-linked security. This chapter fills in this gap. Given the proposed extension, we can readily estimate how the time-0 price of a security would change if the hedger's exposure and the security under consideration are tied to different populations.

Besides prices, we also investigate the impact of population basis risk on the behaviors of hedgers and investors in the longevity risk market. For a deeper understanding of population basis risk, we drill down into its fundamental components:

[1] This pricing method produces a partial equilibrium, where the clearance on this market is obtained independently from prices and quantities in other markets. When the market prices of other mortality-linked securities are available, a general equilibrium that takes all the market prices into account should be used.

1. the difference in the magnitude of mortality rates between the two populations involved in the pricing framework;
2. the volatility of the mortality-linked cash flows associated with the hedger's liability relative to that associated with the security being priced;
3. the (imperfect) correlation between the mortality-linked cash flows associated with the hedger's liability and the security being priced.

We shall examine the influence of these three components of population basis risk on the trading of mortality-linked securities.

Another focus of this chapter is the hedging strategy implied by the tâtonnement pricing process. In an incomplete market like today's longevity risk market, a perfect hedge cannot be formed, but an approximate hedging strategy may be constructed according to a pre-defined hedging objective. Specifically, if the hedger's objective is to maximize its expected utility at a certain future time, then the quantity traded in the tâtonnement equilibrium can be regarded as the corresponding static hedging strategy. We shall examine how such a hedging strategy would change when population basis risk is taken into account. Using the two-population age-period-cohort mortality model, we also evaluate the effectiveness of the hedging strategy implied by the tâtonnement pricing process, with and without population basis risk.

We illustrate our ideas with a hypothetical mortality-linked security, which is designed to help pension funds and annuity providers hedge their exposures to longevity risk. The security is similar to the first ever longevity bond jointly announced by the European Investment Bank and BNP Paribas in November 2004. The illustrations are based on mortality data from UK male population and two of its sub-populations, Scottish male population and UK male insured lives. The data for the two national populations are obtained from the Human Mortality Database (2010), while the insured lives data are provided by the Continuous Mortality Investigation (CMI) Bureau of the Institute and Faculty of Actuaries.

The rest of this chapter is organized as follows. In Section 3.2, we extend the tâtonnement pricing process to incorporate population basis risk and describe two possible methods, one numerical and one analytic, for implementing the pricing process. In Section 3.3, we detail the two-population age-period-cohort mortality model and fit the model to the populations under consideration. We also provide a forecasting method for the two-population mortality model. In Section 3.4, we use the hypothetical longevity bond to illustrate the effect of population basis risk on security prices, static hedging strategies, and the behaviors of the counterparties involved. Finally, in Section 3.5, we conclude this chapter.

3.2 The Extended Tâtonnement Pricing Process

3.2.1 The Set-up

The tâtonnement approach proposed in Chapter 2 is limited to the case that the hedger's risk exposure and the security being priced must be linked to the same population of individuals. In this subsection, we relax this assumption and describe a tâtonnement process for modeling the trade of a mortality-linked security between two economic agents, Agents A and B, when population basis risk is involved.

Suppose that Agent A has life contingent liabilities that are due at times 1, 2, ..., T. The amount due at time t is $f_t(Q_t^L)$, which is a deterministic function of Q_t^L, where $Q_t^L = (q_0^L, q_1^L, \ldots, q_t^L)$ is a vector of mortality indexes up to and including time t. We allow f_t to be a function of Q_t^L rather than just q_t^L because the liability payouts can be path-dependent.

The index q_t^L contains information about the mortality of the population associated with Agent A's liability over the period of $t-1$ to t. Depending on the structure of the liability, q_t^L can be a scalar or a vector. For example, if the liability is the annuity payments to a single cohort of annuitants, then q_t^L would be a scalar, representing the mortality rate of the cohort from time $t-1$ to t. On the other hand, if the liability is associated with multiple cohorts of individuals, then q_t^L would be a vector, which contains mortality rates at various ages. At time 0, the values of q_t^L for $t > 0$ are not known and are governed by an underlying stochastic process.

To mitigate its exposure to mortality or longevity risk, Agent A sells (or purchases) a mortality-linked security maturing at time T. At time t, the security makes a payout of $g_t(Q_t^H)$, which is deterministic function of Q_t^H, where $Q_t^H = (q_0^H, q_1^H, \ldots, q_t^H)$ is a vector of mortality indexes up to and including time t. The index q_t^H contains information about the mortality of the population associated with the security over the period of $t-1$ to t. As with q_t^L, q_t^H can be a scalar or a vector, and its value is not known at time 0. We emphasize that, due to population basis risk, q_t^H and q_t^L are not necessarily the same.

Agent B is an investor who trades the mortality-linked security with Agent A, in order to earn a risk premium. At time t, Agent B receives (or pays) an amount of $g_t(Q_t^H)$ per unit of the mortality-linked security purchased (or sold).

Following the specification of a tâtonnement process, we suppose that there exists an imaginary auctioneer who cries an arbitrary price, P, at the beginning. Given this price, each agent will trade a quantity that will maximize its expected terminal utility.

Let ω^A and ω^B be the initial wealths of Agents A and B, respectively. We assume 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 sources of income or payout. We assume $r = 3\%$ in our numerical illustrations.

We let θ^A and θ^B respectively be the quantity that Agents A and B are willing to trade at time 0. As defined in Chapter 2, a positive quantity means that the agent purchases the security, while a negative quantity means the agent sells the security.

At time 0, Agent A trades θ^A units of the mortality-linked security and deposits the rest of its wealth into its bank account. At time $t = 1, 2, \ldots, T$, it receives (or pays) an amount of $|\theta^A g_t(Q_t^H)|$ for the mortality-linked security it traded, pays its life contingent liability $f_t(Q_t^L)$, and deposits the rest of its wealth into its bank account. Let W_t^A be the time-t wealth of Agent A. Then the wealth process for Agent A can be expressed as follows:

$$
\begin{aligned}
W_0^A &= \omega^A \\
W_1^A &= (W_0^A - \theta^A P)e^r + \theta^A g_1(Q_1^H) - f_1(Q_1^L) \\
W_2^A &= W_1^A e^r + \theta^A g_2(Q_2^H) - f_2(Q_2^L) \\
\vdots \; &\;\vdots\; \;\vdots \\
W_T^A &= W_{T-1}^A e^r + \theta^A g_T(Q_T^H) - f_T(Q_T^L).
\end{aligned}
$$

At time 0, Agent B trades θ^B units of the mortality-linked security and deposits the rest of its wealth into its bank account. At time $t = 1, 2, \ldots, T$, it receives (or pays) an amount of $|\theta^B g_t(Q_t^H)|$ for the mortality-linked security it traded, and deposits the rest of its wealth into its bank account. Let W_t^B be the time-t wealth of Agent B. Then the wealth process for Agent B can be expressed as follows:

$$
\begin{aligned}
W_0^B &= \omega^B \\
W_1^B &= (W_0^B - \theta^B P)e^r + \theta^B g_1(Q_1^H) \\
W_2^B &= W_1^B e^r + \theta^B g_2(Q_2^H) \\
\vdots \; &\;\vdots\; \;\vdots \\
W_T^B &= W_{T-1}^B e^r + \theta^B g_T(Q_T^H).
\end{aligned}
$$

We denote the utility functions for Agents A and B by U^A and U^B, respectively. As mentioned earlier, given a price P, each agent trades a quantity that maximizes its expected terminal utility, as specified below,

$$\hat{\theta}^A = \underset{\theta^A}{\operatorname{argsup}}\, \mathbb{E}[U^A(W_T^A)]; \tag{3.1}$$

$$\hat{\theta}^B = \operatorname{argsup}_{\theta^B} \mathbb{E}[U^B(W_T^B)]. \tag{3.2}$$

Note that both $\hat{\theta}^A$ and $\hat{\theta}^B$ are functions of the price P. In the pricing process, the price P is adjusted until the market has reached equilibrium, that is, $\hat{\theta}^A + \hat{\theta}^B = 0$. We use P^* to denote the resulting tâtonnement equilibrium price.

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. It is reasonable to assume that Agent A is more conservative than Agent B, because Agent A wants to hedge away its mortality or longevity risk exposure while Agent B is willing to take the risk in return of a risk premium. In our numerical examples, the assumed value of k for Agent A is $k^A = 1.0$ and for Agent B is $k^B = 0.5$.

It is interesting to note that, when an exponential utility function is assumed, the initial wealth of each agent has no effect on the estimated price and the quantity traded in the tâtonnement equilibrium. A proof of this property can be found Chapter 2.

3.2.2 Solving by a Numerical Procedure

The tâtonnement equilibrium can be solved numerically similarly with Chapter 2. We begin with a first guess of the price. If the market does not clear, we adjust the price. In particular, the price needs to be raised if demand exceeds supply (i.e., $\hat{\theta}^A + \hat{\theta}^B > 0$), and vice versa. Mathematically, the $(i+1)$th update of the price can be expressed as

$$\begin{aligned} P^{(i+1)} &= P^{(i)} + d_i, \quad i = 0, 1, \ldots, & (3.3) \\ d_i &= \vartheta |P^{(i)}| (\hat{\theta}^A + \hat{\theta}^B) & (3.4) \end{aligned}$$

where $P^{(0)}$ is the initial guess of the price, and d_i is a function that always has the same sign as the excess demand, $\hat{\theta}^A + \hat{\theta}^B$. ϑ is set to 0.001 after some experiments considering a suitable balance between speed and accuracy.

Summing up, a numerical solution to the tâtonnement equilibrium can be obtained with the algorithm below:

1. Guess a price $P^{(0)}$.

2. Determine $\hat{\theta}^A$ and $\hat{\theta}^B$, on the basis of the current estimate of the price and the optimizing criteria specified by equations (3.1) and (3.2).

3. Terminate the algorithm if $|\hat{\theta}^A + \hat{\theta}^B|$ is less than a tolerance level, 10^{-4}. Otherwise, adjust the price using equation (3.3).

4. Repeat Steps 2 and 3.

The expected terminal utility involved in the algorithm above can be calculated by Monte Carlo simulations, as detailed in Chapter 2.

3.2.3 Solving by an Approximate Analytic Formula

Alternatively, if we were to impose some assumptions on the contingent cash flows involved in the wealth processes, we can solve for the tâtonnement equilibrium analytically.

Let v_L and v_H respectively be the accumulated values of the life contingent liabilities and the payouts from the mortality-linked security at a future time T. Both v_L and v_H are positive real numbers.

As the initial wealths, ω_A and ω_B, has no effect on the tâtonnement equilibrium under the assumed utility function, we may set $\omega_A = \omega_B = 0$ without loss of generality. In this case, the terminal wealths for Agents A and B can be expressed as follows:

$$
\begin{aligned}
W_T^A &= -v_L + \theta^A (v_H - Pe^{rT}), \\
W_T^B &= \theta^B (v_H - Pe^{rT}).
\end{aligned}
$$

We can obtain an analytical formula for the tâtonnement equilibrium price if we assume that $(v_H, v_L)'$ follows a bivariate normal distribution. The validity of this assumption will be checked later in this chapter. In particular, we assume that $(v_H, v_L)'$ follows a bivariate normal distribution with mean vector $(\mu_H, \mu_L)'$ and variance-covariance matrix

$$
\Sigma = \begin{pmatrix} \sigma_H^2 & \rho\sigma_H\sigma_L \\ \rho\sigma_H\sigma_L & \sigma_L^2 \end{pmatrix},
$$

where $-1 \leq \rho \leq 1$ is the correlation coefficient between the random variables v_H and v_L. It is easy to show that, under this assumption, the terminal wealths W_T^A and W_T^A are normally distributed. Specifically,

$$\begin{aligned} W_T^A &\sim \mathrm{N}\left(-\mu_L + \theta^A(\mu_H - Pe^{rT}), \sigma_L^2 + (\theta^A\sigma_H)^2 - 2\rho\sigma_L\sigma_H\theta^A\right), \\ W_T^B &\sim \mathrm{N}\left(\theta^B(\mu_H - Pe^{rT}), (\theta^B\sigma_H)^2\right), \end{aligned}$$

where $\mathrm{N}(\mu, \sigma^2)$ denotes a normal distribution with mean μ and variance σ^2.

Recall that, given a price, each agent trades a quantity that maximizes its expected terminal utility. Assuming exponential utility functions, we have the following result:

Proposition 3. *Assume that $(v_H, v_L)'$ follows a bivariate normal distribution and that the utility functions for Agents A and B are exponential with respective parameters k^A and k^B. At a given price P, Agent A will trade*

$$\hat{\theta}^A = \frac{\rho\sigma_L\sigma_H k^A + \mu_H - Pe^{rT}}{k^A\sigma_H^2} \tag{3.5}$$

units of the security, while Agent B will trade

$$\hat{\theta}^B = \frac{\mu_H - Pe^{rT}}{k^B\sigma_H^2} \tag{3.6}$$

units of the security.

Proof. For Agent A, the trading quantity that maximizes its expected terminal utility given a price P is

$$\begin{aligned}
\hat{\theta}^A &= \underset{\theta^A}{\operatorname{argsup}}\ \mathbb{E}\left[U^A(W_T^A)\right] \\
&= \underset{\theta^A}{\operatorname{argsup}}\ \mathbb{E}\left[1 - e^{-k^A W_T^A}\right] \\
&= \underset{\theta^A}{\operatorname{arginf}}\ \mathbb{E}\left[e^{-k^A W_T^A}\right] \\
&= \underset{\theta^A}{\operatorname{arginf}}\ e^{-k^A\left(-\mu_L+\theta^A(\mu_H-Pe^{rT})\right)+\frac{1}{2}(k^A)^2\left(\sigma_L^2+{\theta^A}^2\sigma_H^2-2\rho\sigma_L\sigma_H\theta^A\right)} \\
&= \underset{\theta^A}{\operatorname{arginf}}\ e^{\frac{1}{2}\left(k^A\sigma_H\left(\theta_A-\frac{\rho\sigma_L\sigma_H k^A+\mu_H-Pe^{rT}}{k^A\sigma_H^2}\right)\right)^2} \\
&= \underset{\theta^A}{\operatorname{arginf}}\ \left(\theta^A - \frac{\rho\sigma_L\sigma_H k^A + \mu_H - Pe^{rT}}{k^A\sigma_H^2}\right)^2 \\
&= \frac{\rho\sigma_L\sigma_H k^A + \mu_H - Pe^{rT}}{k^A\sigma_H^2}.
\end{aligned}$$

In the above, the third step follows from the moment generating function of a normal distribution.

Similarly, for Agent B, the trading quantity that maximizes its expected terminal utility given a price P is

$$\begin{aligned}
\hat{\theta}^B &= \underset{\theta^B}{\operatorname{argsup}}\ \mathbb{E}\left[U^B(W_T^B)\right] \\
&= \underset{\theta^B}{\operatorname{argsup}}\ \mathbb{E}\left[1 - e^{-k^B W_T^B}\right] \\
&= \underset{\theta^B}{\operatorname{arginf}}\ \mathbb{E}\left[e^{-k^B W_T^B}\right] \\
&= \underset{\theta^B}{\operatorname{arginf}}\ e^{-k^B\theta^B\left(\mu_H-Pe^{rT}\right)+\frac{1}{2}(k^B\theta^B\sigma_H)^2} \\
&= \underset{\theta^B}{\operatorname{arginf}}\ e^{\frac{1}{2}\left(k^B\sigma_H\left(\theta^B-\frac{\mu_H-Pe^{rT}}{k^B\sigma_H^2}\right)\right)^2} \\
&= \underset{\theta^B}{\operatorname{arginf}}\ \left(\theta^B - \frac{\mu_H - Pe^{rT}}{k^B\sigma_H^2}\right)^2 \\
&= \frac{\mu_H - Pe^{rT}}{k^B\sigma_H^2}.
\end{aligned}$$

$\square$

Recall that in the tâtonnement equilibrium, we have $\hat{\theta}^A + \hat{\theta}^B = 0$. Substituting $\hat{\theta}^A$ and $\hat{\theta}^B$ into this equation, we can solve for the tâtonnement equilibrium price P^* readily.

Proposition 4. *Assume that $(v_H, v_L)'$ follows a bivariate normal distribution and that the utility functions for Agents A and B are exponential with respective parameters k^A and k^B. The tâtonnement equilibrium price of the security is given by*

$$P^* = \frac{(k^A + k^B)\mu_H + k^A k^B \rho \sigma_L \sigma_H}{(k^A + k^B)e^{rT}}. \tag{3.7}$$

It is not surprising that we are able to find an analytic solution under the assumption of normally distributed terminal wealths and exponential utility functions. Freud (1956) shows that the combination of these two assumptions can significantly simplify the maximization problem. In more detail, under these two assumptions, maximizing expected terminal utility is equivalent to maximizing a linear function of the mean and variance of terminal wealth.

To evaluate $\hat{\theta}^A$, $\hat{\theta}^B$, and P^*, we need the parameters in the bivariate normal distribution which $(v_H, v_L)'$ follows. These parameters can be estimated with the algorithm below:

1. simulate 10000 mortality paths from a stochastic mortality model;
2. for each of the simulated mortality paths, calculate $(v_H, v_L)'$;
3. take the sample mean vector of the simulated values of $(v_H, v_L)'$ as an estimate of $(\mu_H, \mu_L)'$;
4. take the sample variance-covariance matrix of the simulated values of $(v_H, v_L)'$ as an estimate of Σ.

The analytic solution requires a rather restrictive assumption on $(v_H, v_L)'$, but, as we will demonstrate in Section 4, it allows us to understand the impact of population basis risk on the tâtonnement equilibrium more easily.

3.3 Mortality Models

3.3.1 Data

Recall that the generalized pricing process involves two populations, one is associated with the hedger's exposure and the other is linked directly to the security being priced. To illustrate the pricing process, we consider the following two pairs of populations:

1. UK male population and Scottish male population.
2. UK male population and the population of UK male insured lives (thereafter called the CMI population for brevity).

In both pairs, the second population is a subset of the first one. We believe that such an arrangement would maximize the resemblance of our illustrations to actual trades, because, for example, in forming a longevity hedge with a standardized instrument, a UK pension plan would naturally choose one that is linked to the national population of UK, if that is available.

The data (death and exposure counts) for the two national populations are obtained from the Human Mortality Database (2010), while that for the insured lives are provided by the CMI Bureau of the Institute and Faculty of Actuaries. For all three populations, we consider a sample period of 1947 to 2005 and a sample age range of 60 to 89. Hence, the estimated models cover 88 years of birth (1858 to 1945).

In Figure 3.1 we compare the mortality of the two subpopulations with that of the UK population. We observe that the mortality of Scottish males is higher than that of UK males, while the mortality of CMI males is the opposite. It is not surprising that the CMI population has a lighter mortality, as people who are willing and able to buy insurance are generally wealthier and healthier. In Sections 3.4, we will demonstrate that the difference in the magnitude of mortality rates between the populations involved in the trade would have an effect on the tâtonnement equilibrium.

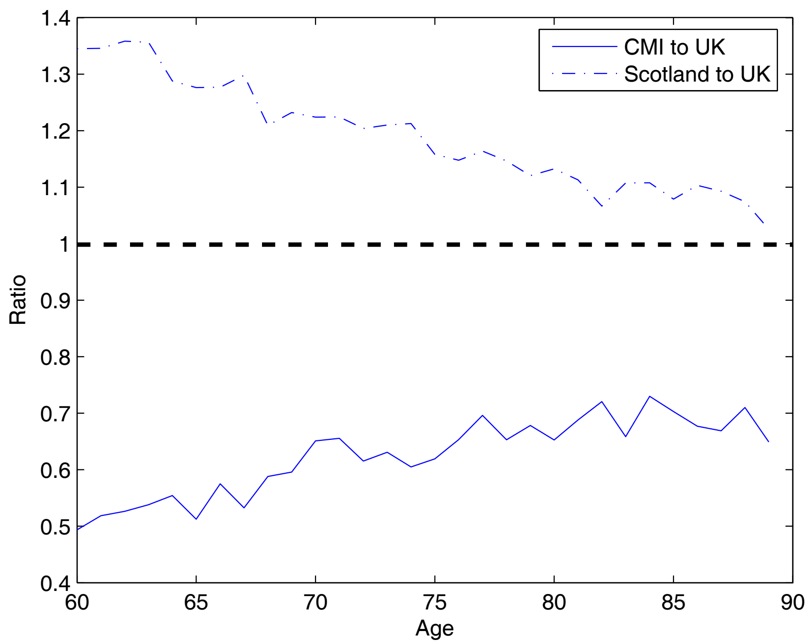


Figure 3.1: Ratios of central death rates in 2005: Scottish males to UK males; CMI males to UK males.

In addition, the difference in the volatility of mortality rates has also an impact on the equilibrium price and quantity traded. We will give a deeper account of volatility in Section 3.3.3 where we estimate the volatility parameters in the assumed model.

3.3.2 Model Specification

Previous studies have shown empirically that the mortality rates of a population and its subpopulation are correlated. For instance, Coughlan et al. (2011) found a stable long-term relationship between the mortality of English & Welsh population and UK insured lives. From a statistical viewpoint, it is more sensible to model the mortality of each pair of populations jointly rather than in isolation.

There are two other reasons for using a joint model. First, using two independent mortality models is likely to result in an increasing divergence in life expectancy in the long run, counter to the expected and observed trend towards convergence (see, e.g., Li and Lee (2005), White (2002), Wilson (2001)). Second, given that the populations involved in the trade are related, the use of two independent models will overstate the underlying population basis risk. This may lead us to overestimating the agents' reaction to population basis risk, and consequently misestimating the tâtonnement equilibrium.

We implement the generalized pricing process with the two-population age-period-cohort model proposed by Cairns et al. (2011a). The model is built from two classical age-period-cohort models, one for each population:

$$\ln(m_{x,t}^{(1)}) = \beta_x^{(1)} + \frac{1}{n_a}\kappa_t^{(1)} + \frac{1}{n_a}\gamma_{t-x}^{(1)}; \tag{3.8}$$

$$\ln(m_{x,t}^{(2)}) = \beta_x^{(2)} + \frac{1}{n_a}\kappa_t^{(2)} + \frac{1}{n_a}\gamma_{t-x}^{(2)}, \tag{3.9}$$

where $m_{x,t}^{(i)}$, $i = 1, 2$, is the central death rate at age x and in year t for population i, and n_a is a constant which equals the total number of ages in the sample age range. In the model, the base age-pattern of mortality for population i, $i = 1, 2$, is characterized by an age-specific parameter $\beta_x^{(i)}$. For each population, the variation of mortality over time is captured by two indexes: a period effect index $\kappa_t^{(i)}$ and a cohort effect index $\gamma_{t-x}^{(i)}$. Note that $t - x$ is the year of birth for an individual who is aged x in year t. Projections of future mortality can be made by extrapolating the indexes.

In both models we are estimating, population 1 refers to UK male population. This larger population is assumed to have a dominant effect on the mortality improvements for populations 1 and 2. Following Cairns et al. (2011a), we model $\kappa_t^{(1)}$ with a random walk with drift,

$$\kappa_t^{(1)} = \mu_\kappa + \kappa_{t-1}^{(1)} + Z_\kappa(t), \tag{3.10}$$

where μ_κ is a constant, and model $\gamma_{t-x}^{(1)}$ with a second order autoregressive model, AR(2), with a deterministic trend,

$$\gamma_c^{(1)} = \mu_\gamma + \phi_{\gamma,1}\gamma_{c-1}^{(1)} + \phi_{\gamma,2}\gamma_{c-2}^{(1)} + \delta_\gamma c + Z_\gamma(c), \tag{3.11}$$

where $c = t - x$, and μ_γ, $\phi_{\gamma,1}$, $\phi_{\gamma,2}$, and δ_γ are constants.

The indexes for population 2 may deviate from those for population 1. However, to ensure that the resulting forecasts are biologically reasonable, there is a need to avoid a divergence of death rates in the two populations over time. The conditions for non-divergence are:

1. $\Delta_\kappa(t) = \kappa_t^{(1)} - \kappa_t^{(2)}$ is mean-reverting;
2. $\Delta_\gamma(c) = \gamma_c^{(1)} - \gamma_c^{(2)}$ is mean-reverting.

Following Cairns et al. (2011a), we model $\Delta_\kappa(t)$ and $\Delta_\gamma(c)$ with an AR(1) process and an AR(2) process, respectively. That is,

$$\Delta_\kappa(t) = \mu_{\Delta_\kappa} + \phi_{\Delta_\kappa}\Delta_\kappa(t-1) + Z_{\Delta_\kappa}(t), \tag{3.12}$$

and

$$\Delta_\gamma(c) = \mu_{\Delta_\gamma} + \phi_{\Delta_{\gamma,1}}\Delta_\gamma(c-1) + \phi_{\Delta_{\gamma,2}}\Delta_\gamma(c-2) + Z_{\Delta_\gamma}(c), \tag{3.13}$$

where μ_{Δ_κ}, ϕ_{Δ_κ}, μ_{Δ_γ}, $\phi_{\Delta_{\gamma,1}}$, and $\phi_{\Delta_{\gamma,2}}$ are constants. These two time-series processes are stationary, ensuring that both $\Delta_\kappa(t)$ and $\Delta_\gamma(c)$ will revert to their long-term means.

Other than mean-reversions, as indicated in the empirical results of Coughlan et al. (2011), we might also expect to see some correlations between the year-on-year changes in both the period and cohort effects. For instance, a flu epidemic may have similar effects on populations 1 and 2. To incorporate these potential correlations, we treat both $(Z_\kappa(t), Z_{\Delta_\kappa}(t))'$ and $(Z_\gamma(c), Z_{\Delta_\gamma}(c))'$ as zero-mean bivariate normal random vectors, with variance-covariance matrices V_κ and V_γ, respectively.

3.3.3 Fitting the Model

In Cairns et al. (2011a), the two-population model is fitted by a single-stage estimation procedure based on Markov Chain Monte Carlo (MCMC). The procedure estimates the parameters in the age-period-cohort models, equations (3.8) and (3.9), and the time-series processes, equations (3.10) to (3.13), simultaneously. According to Cairns et al. (2011a), this single-stage Bayesian approach can handle missing data easily, and can provide measures of parameter risk.

While we fully agree with its advantages, the single-stage approach is not particularly suitable for our applications. A problem of this approach is that it will yield different parameter estimates for UK males, when the other population involved in the model is different. In our analysis, however, we require both models under consideration to generate a consistent mortality forecast for UK males. To circumvent this problem, we use a two-stage estimation procedure, which has been considered extensively by researchers such as Cairns et al. (2006), Lee and Carter (1992) and Li et al. (2009).

In the first stage, we estimate the parameters in the age-period-cohort models by the method of maximum likelihood. Let us define $D^{(i)}_{x,t}$ by the number of deaths in population i at age x from in year t, and $E^{(i)}_{x,t}$ by the corresponding exposures to the risk of death. To construct the likelihood function, we treat $D^{(i)}_{x,t}$ as independent Poisson responses, that is,

$$D^{(i)}_{x,t} \sim \text{Poisson}(E^{(i)}_{x,t}\, m^{(i)}_{x,t}),$$

where the product $E^{(i)}_{x,t}\, m^{(i)}_{x,t}$ is the expected number of deaths in population i at age x and in year t. This gives the following log-likelihood:

$$l(i) = \sum_{x,t} \left(D^{(i)}_{x,t} \ln(m^{(i)}_{x,t}) - E^{(i)}_{x,t}\, m^{(i)}_{x,t} \right) + e,$$

where e is a constant.

Parameter estimates for population i can be obtained by maximizing the likelihood function, $l(i)$. The maximization can be accomplished by an iterative Newton-Raphson method, in which parameters are updated one at a time.

The age-period-cohort model has an identifiability problem. To stipulate parameter uniqueness, three constraints are needed for each population. The constraints

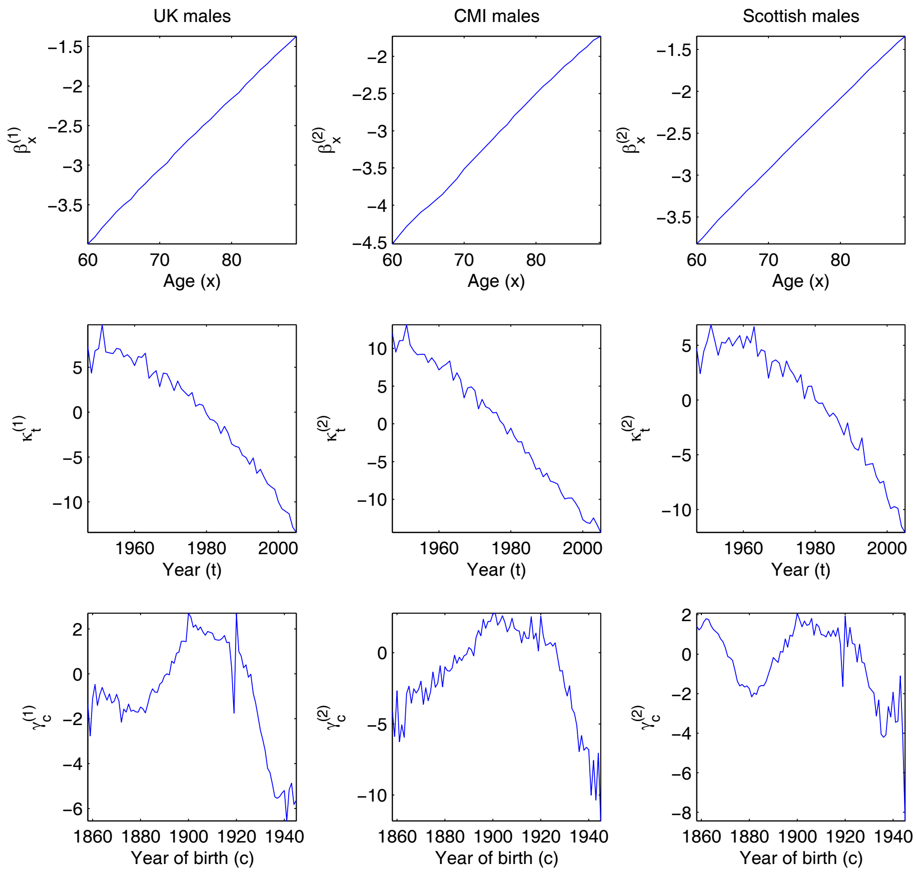


Figure 3.2: Estimated parameters in the age-period-cohort models.

used here are the same as those used by Cairns et al. (2009). They are applied at the end of each iteration of the Newton-Raphson algorithm.

The first stage estimation gives us estimates of $\beta_x^{(i)}$, $\kappa_t^{(i)}$, and $\gamma_{t-x}^{(i)}$ for $i = 1, 2$, $t = 1947, \ldots, 2005$, and $x = 60, \ldots, 89$. The estimates are displayed graphically in Figure 3.2.

In the second stage, the parameters in the time-series processes are estimated. We first fit equations (3.10) to (3.13) as if $\kappa_t^{(1)}$ is independent of $\Delta_\kappa(t)$ and $\gamma^{(1)}(c)$ is independent of $\Delta_\gamma(c)$. This would give estimates of all parameters in the time-series

Parameters	UK and Scottish males	UK and CMI males
μ_κ	-0.3564	-0.3564
μ_γ	-0.0098	-0.0098
δ_γ	-0.0427	-0.0427
$\phi_{\gamma,1}$	0.6232	0.6232
$\phi_{\gamma,2}$	0.3430	0.3430
μ_{Δ_κ}	0.0152	-0.0295
ϕ_{Δ_κ}	0.8822	0.9692
μ_{Δ_γ}	-0.0660	0.1155
$\phi_{\Delta_{\gamma,1}}$	0.7953	0.2060
$\phi_{\Delta_{\gamma,1}}$	0.0674	0.7177
$V_\kappa(1,1)$	1.1287	1.1287
$V_\kappa(1,2)$	0.1409	0.2543
$V_\kappa(2,2)$	0.3317	0.4053
$V_\gamma(1,1)$	0.4844	0.4844
$V_\gamma(1,2)$	0.1258	0.1218
$V_\gamma(2,2)$	0.5731	0.8220

Table 3.1: Estimates of the parameters in the time-series processes for $\kappa_t^{(1)}$, $\gamma_c^{(1)}$, $\Delta_\kappa(t)$, and $\Delta_\gamma(c)$.

processes, except the variance-covariance matrices V_κ and V_γ. We then estimate V_κ with the sample variance-covariance matrix for the residuals in fitting κ_t and $\Delta_\kappa(t)$. Similarly, we estimate V_γ with the sample variance-covariance matrix for the residuals in fitting γ_c and $\Delta_\gamma(c)$. The estimates of the parameters in the time-series processes are shown in the Table 3.1.

From the estimates of V_κ and V_γ, we can calculate the volatilities of the year-on-year innovations for the period and cohort effects associated with each of the three populations. The results, which we summarize in Table 3.2, indicate that Scottish

Population	Period effect	Cohort effect
UK males	1.062	0.696
Scottish males	1.086	0.897
CMI males	1.013	1.031

Table 3.2: Volatilities of the year-on-year innovations for the period and cohort effect indexes.

males have the most volatile period effect index, while CMI males have the most volatile cohort effect index. As we will see in Section 3.4, the knowledge on the sources of volatility can help us better understand the tâtonnement equilibria formed with different pairs of populations.

3.3.4 Simulating Mortality Paths

A naive way to simulate mortality paths for the three populations is to first simulate from the mortality model for UK and Scottish males, and then simulate from the mortality model for UK and CMI males. However, this method would produce two different collections of sample paths for UK males. The inconsistency would create two (slightly) different supply (or demand) curves for a security, even though all pricing parameters remain unchanged.

To ensure a consistent collection of mortality paths is produced for UK males, we use the following procedure to conduct the necessary simulations:

1. generate 10 000 sample paths for $Z_\kappa(t)$ on the basis of its marginal distribution, $\mathrm{N}(0, V_\kappa(1,1))$;

2. generate 10 000 sample paths for $Z_\gamma(c)$ on the basis of its marginal distribution, $\mathrm{N}(0, V_\gamma(1,1))$;

3. based on the simulated values of $Z_\kappa(t)$ and $Z_\gamma(c)$ in Steps 1 and 2, obtain 10 000 sample mortality paths for UK males;

4. for each value of $Z_\kappa(t)$ in Step 1, find the distribution of $Z_{\Delta_\kappa}(t)|Z_\kappa(t)$ for Scottish males, and simulate a value of $Z_{\Delta_\kappa}(t)$ from the conditional distribution;

5. for each value of $Z_\gamma(c)$ in Step 2, find the distribution of $Z_{\Delta_\gamma}(c)|Z_\gamma(c)$ for Scottish males, and simulate a value of $Z_{\Delta_\gamma}(c)$ from the conditional distribution;

6. based on the 10 000 simulated paths of $Z_\kappa(t)$, $Z_{\Delta_\kappa}(t)$, $Z_\gamma(c)$, and $Z_{\Delta_\gamma}(c)$, obtain 10 000 sample mortality paths for Scottish males;

7. repeat Steps 4 to 6 for CMI males to obtain 10 000 sample paths for CMI males.

Following from the assumption that $(Z_\kappa(t), Z_{\Delta_\kappa}(t))'$ follows a bivariate normal distribution with a zero mean vector and a variance-covariance matrix of V_κ, we have

$$Z_{\Delta_\kappa}(t)|Z_\kappa(t) \sim \mathrm{N}\left(Z_\kappa(t)\frac{V_\kappa(1,2)}{V_\kappa(1,1)}, \frac{\det(V_\kappa)}{V_\kappa(1,1)}\right), \tag{3.14}$$

where $\det(v)$ denotes the determinant of a matrix v. Similarly, we have

$$Z_{\Delta_\gamma}(c)|Z_\gamma(c) \sim \mathrm{N}\left(Z_\gamma(c)\frac{V_\gamma(1,2)}{V_\gamma(1,1)}, \frac{\det(V_\gamma)}{V_\gamma(1,1)}\right). \tag{3.15}$$

A proof for the above is provided in the Appendix A.

Samples of the simulated mortality paths are displayed in Figure 3.3. The mortality paths enable us to evaluate expressions (3.1) and (3.2), which are involved in the numerical procedure for solving the tâtonnement equilibrium. They also allow us to calculate the sample mean and sample variance-covariance matrix for $(v_H, v_L)'$, which are involved in the approximate analytic formula for the tâtonnement equilibrium price of the security.

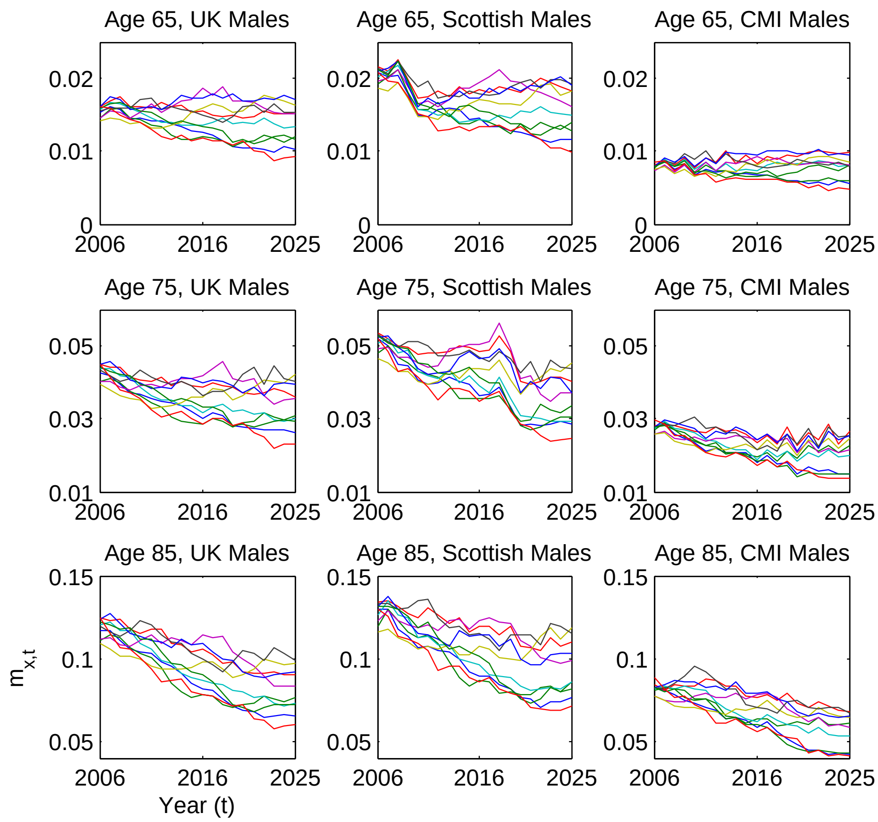


Figure 3.3: Sample Paths of $m_{x,t}$, $x = 65, 75, 85$, $t = 2006, \ldots, 2025$, simulated from the two-population age-period-cohort models.

3.4 An Illustration

3.4.1 Specification of the Security

In this section, we illustrate the generalized tâtonnement pricing process with a hypothetical security. The security we consider is a 25-year annuity bond (a bond without principal repayment), which is similar to the longevity bond jointly announced by the European Investment Bank and BNP Paribas in November 2004.[2]

[2]More details on this bond are given in Blake et al. (2006, 2008).

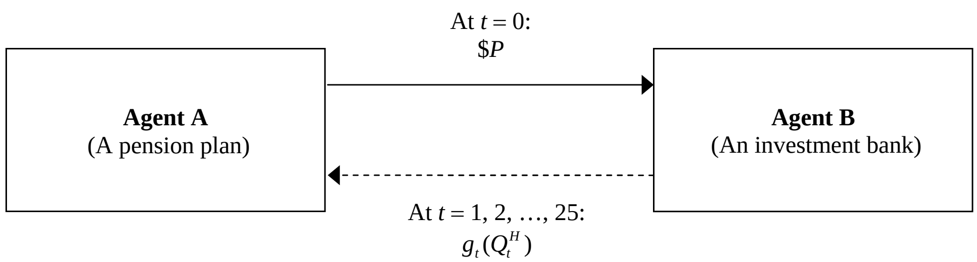


Figure 3.4: Cash flows involved in the hypothetical longevity bond.

The coupon payments are linked to the realized mortality rates of individuals in the reference population who are aged 65 at time 0 (the beginning of year 2006). They can be specified by using the notation defined in Section 5.2. Specifically, we set the mortality index q_t^H to the time-t value of the central death rate for the cohort of individuals. Let $Q_t^H = (q_1^H, \ldots, q_t^H)$ be the vector of mortality indexes up to and including time t. The coupon payment at time $t = 1, 2, \ldots, 25$ is given by

$$g_t(Q_t^H) = \prod_{i=1}^{t}(1 - q_i^H),$$

which is the (approximate) realized survival rate to time t.[3]

If the realized survival rates are larger than expected, then pension payments, on the whole, will last longer, which means larger pension liabilities. Therefore, a pension plan wishing to hedge longevity risk may purchase the longevity bond, which will pay out to the pension plan, at time $t = 1, 2, \ldots, 25$, an amount that increases with the realized survival rate to offset the correspondingly higher value of pension liabilities.

In the absence of credit risk, the cash flows involved are simple to specify (see Figure 3.4). The investor makes an initial payment of $\$P$ (i.e., the issue price) and receives in return an annual mortality-dependent payment of $g_t(Q_t^H)$ in each year t for 25 years.

[3]It is approximate because it is based on central death rates rather than death probabilities.

3.4.2 The Trade

We assume that the longevity bond is sold by Agent B, which could possibly be an investment bank attempting to earn a longevity risk premium. The bond is sold to Agent A, a pension plan provider having an exposure to longevity risk.

We assume that Agent A is obligated to pay each of its pensioner an amount of \$0.01 at the end of each year. The pension payment ceases if the pensioner dies or reaches age 90, whichever occurs earlier. At time 0 (the beginning of year 2006), the plan contains 1000 pensioners, who are distributed over the age range of 65 to 89. The age distribution of the pensioners is shown in Figure 3.5. For simplicity, we assume that the plan is closed, that is, there are no new entrants to the plan.

It is obvious that Agent A's financial obligation is linked to the mortality of its pensioners. In particular, it is linked to an index q_t^L, which contains the realized death rates of 25 cohorts of pensioners (with years of birth ranging from 1916 to 1940) at time t. We permit population basis risk in our pricing process so that q_t^L and q_t^H are not necessarily associated with the same population.

If q_t^L and q_t^H are positively correlated, then an increase in the pension liability will be accompanied with an increase in the coupon payments from the longevity bond. Hence, by purchasing the bond, Agent A can reduce its risk exposure.

As mentioned earlier, our illustrations are based on two different pairs of populations. Each pair is composed of UK males and one of its subpopulation, which is either Scottish males or CMI males. For each pair of populations, the following three cases are examined:

- **Case 1**
  Both the pension liability and the longevity bond are linked to the mortality of UK males.[4] There is no basis risk involved in this case.

[4]By saying the pension liability is linked to a certain population, we mean the realized mortality rates for members in the pension plan and individuals in that population are the same.

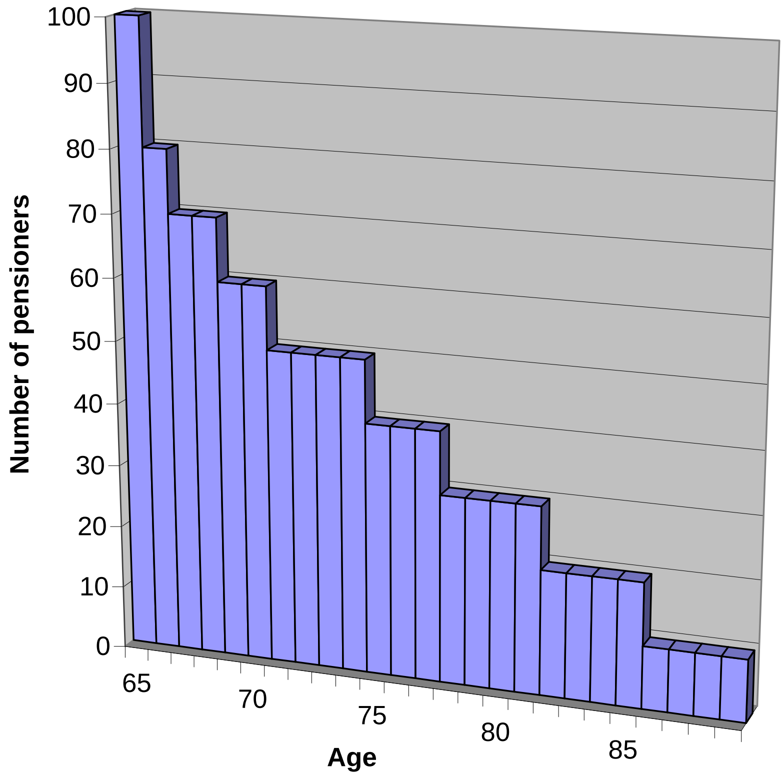


Figure 3.5: Age distribution of the pensioners in Agent A's plan.

- **Case 2**

  The bond is linked to the mortality of UK males, while the pension liability is linked to the subpopulation. Basis risk is involved in this case.

- **Case 3**

  Both the pension liability and the longevity bond are linked to the mortality of the subpopulation. There is no basis risk involved in this case.

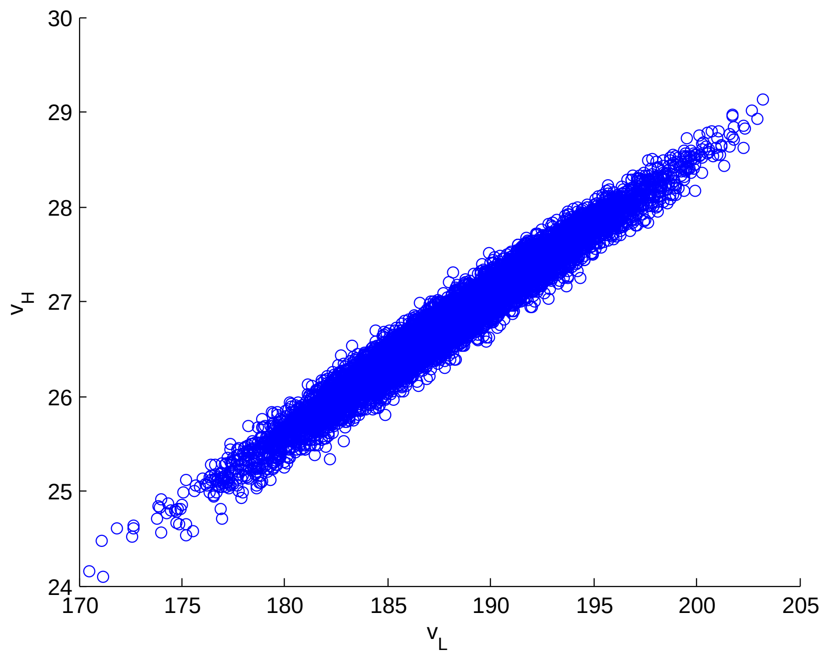


Figure 3.6: Scatter plot of $(v_H, v_L)'$, Case 1.

3.4.3 Validating the Bivariate Normal Approximation

Before describing the details about pricing, we validate the bivariate normal approximation of $(v_H, v_L)'$. Consider Case 1, in which both the pension liability and the longevity bond are linked to the mortality of UK males. Using the 10 000 simulated mortality paths, we calculate $(v_H, v_L)'$ corresponding to each path. Its values are shown on the scatter plot in Figure 3.6. Figure 3.6 shows that all the simulated values of $(v_H, v_L)'$ fall on a line that is almost linear. The pattern of the dots slopes from lower left to upper right. This suggests that v_H and v_L have high positive correlation.

To examine the normality of v_H and v_L, we draw the Q-Q plots for them respectively. In Figure 3.7, the upper panel presents the Q-Q plot of v_H and the lower panel presents the Q-Q plot of v_L. Both Q-Q plots show that the left tail and the right tail have quartiles lower than standard normal distribution. This pattern suggests that v_H and v_L are slightly left skewed.

The Q-Q plots only examine the normality of marginal distributions. We still need to test the multivariate normality for $(v_H, v_L)'$. Henze-Zirkler's multivariate normality test is used for this purpose here. According to Henze-Wagner (1997), the Henze-Zirkler test is based on a nonnegative function that measures the distance between the characteristic function of the multivariate normal distribution and the empirical characteristic function. If the data come from a multivariate normal distribution, the test statistic is approximately lognormally distributed. It proceeds to calculate and lognormalize the mean and variance. Finally, the p-value is estimated.

Due to the limited computer memory, we are only able to test 8 000 sets of simulated values. The p-value is 0.1354, which means we cannot reject the null hypothesis that $(v_H, v_L)'$ follows bivariate normal distribution at a significance level of 5%. The approximation is therefore reasonable.

3.4.4 UK and Scottish Males

Let us suppose here that the subpopulation is Scottish males. For each of the three cases, we calculate the exact tâtonnement equilibrium price of the longevity bond by using the algorithm presented in Section 3.2.2. The estimated prices are displayed in Table 3.3. Also shown in Table 3.3 are the quantities traded in each case.

To know how the prices are formed, we need to examine the demand and supply curves, which can be derived by evaluating expressions (3.1) and (3.2) at different price levels.

The upper panel of Figure 3.8 depicts the demand and supply curves for Cases 1 and 2. The supply curves for both cases are the same, but the demand curve shifts downwards when the population to which the pension liability is linked is changed from UK males to Scottish males. This leads to a reduction in the price of the security.

The lower panel of Figure 3.8 depicts the demand and supply curves for Cases 2 and 3. As we change the population to which the longevity bond is linked, the demand curve shifts downwards, but the supply curve shifts upwards. This results in a reduction in the price of the security.

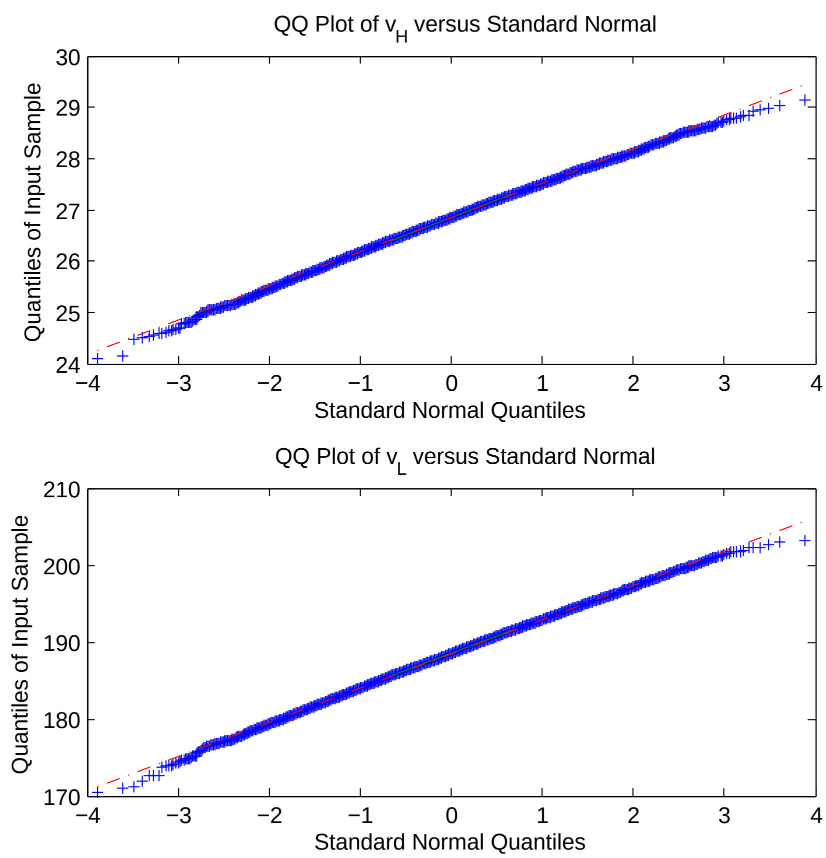


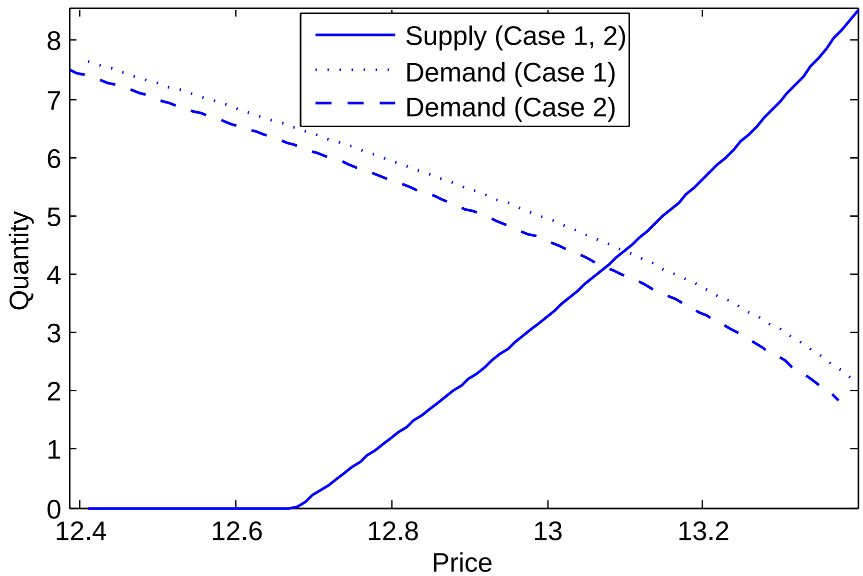
Figure 3.7: Q-Q plots of v_H and v_L, Case 1

What constitutes the shifts in the supply and demand curves? The analytic formulas in Section 3.2.3 may help us answer this question. In the current application, Agent B is the supplier, so we have $\hat{\theta}^B \leq 0$ and hence $\hat{\theta}^A \geq 0$. Using equations (3.5) and (3.6), at a price P, the (approximate) demand of the security from Agent A is

$$|\hat{\theta}^A| = \max\left(\frac{\rho\sigma_L\sigma_H k^A + \mu_H - Pe^{rT}}{k^A\sigma_H^2}, 0\right), \tag{3.16}$$

while the (approximate) supply of the security from Agent B is

$$|\hat{\theta}^B| = \max\left(\frac{Pe^{rT} - \mu_H}{\sigma_H^2 k^B}, 0\right). \tag{3.17}$$



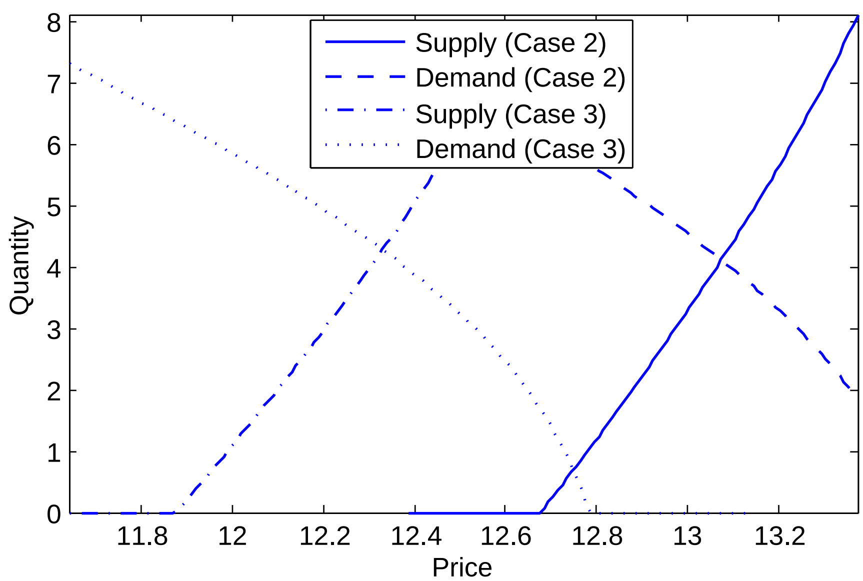


Figure 3.8: Demand and supply curves for the hypothetical longevity bond, UK and Scottish males.

	Case 1	Case 2	Case 3
Populations			
Longevity bond	UK males	UK males	Scottish males
Pension liability	UK males	Scottish males	Scottish males
Prices and units traded in equilibrium			
Time-0 price	\$13.0989	\$13.0753	\$12.3295
Units traded	4.3981	4.1245	4.2883
Parameters in the distribution of $(v_l, v_h)'$			
μ_H	26.8379	26.8379	25.1402
σ_H	0.6684	0.6684	0.6981
σ_L	4.4571	4.5728	4.5728
ρ	0.9804	0.9237	0.9807
Effectiveness of the static hedge			
Variance reduction	85.8%	75.0%	85.6%

Table 3.3: Results of the pricing process, UK and Scottish males.

We observe that both $|\hat{\theta}^A|$ and $|\hat{\theta}^B|$ are linear functions of price P when they are greater than zero. Their slopes and intercepts are determined by parameters μ_H, σ_H, σ_L, and ρ. The effects of these parameters on the supply and demand of the security are summarized in Table 3.4. In the table, '↑' means an increase, '↓' means an decrease, and '−' means there is no change.

The relations are rather intuitive. First, consider the supply, $|\hat{\theta}^B|$, from Agent B. Recall that v_H is the accumulated value of the payouts from the longevity bond at maturity, and that μ_H is its expectation. Therefore, the difference $Pe^{rT} - \mu_H$ is the reward to Agent B for accepting a longevity risk exposure. As a result, when μ_H increases, the reward to Agent B is smaller, and hence it will supply less. On the other hand, when σ_H increases, the longevity bond becomes more risky. If the reward is held constant, Agent B must supply less.

Next, we consider the demand, $|\hat{\theta}^A|$, from Agent A. It is quite obvious that Agent

	$\mu_H \uparrow$	$\sigma_H \uparrow$	$\sigma_L \uparrow$	$\rho \uparrow$
Demand, $\lvert\theta^A\rvert$	$\uparrow$	$\downarrow$	$\uparrow$	$\uparrow$
Supply, $\lvert\theta^B\rvert$	$\downarrow$	$\downarrow$	–	–

Table 3.4: The effects of μ_H, σ_H, σ_L, and ρ on the supply and demand of the hypothetical longevity bond.

A will demand more if its pension liability is more volatile (i.e., σ_L is larger). It will also demand more if less compensation to Agent B is needed (i.e., μ_H is higher). Moreover, the demand from Agent A is dependent on ρ, which is a positive number in this application. When ρ increases (becomes closer to one), the bond becomes a more effective hedging instrument, and therefore Agent A will demand more. Finally, Agent A will demand less when σ_H is large relative to σ_L, as in this situation fewer units of the bond will be needed for the same amount of risk reduction.

In Table 3.3 we show the estimates μ_H, σ_H, σ_L, and ρ when the pension liability and the longevity bond are linked to the mortality of either UK or Scottish males.

Recall that the longevity bond is related only to the cohort born in 1940 and that the pension liability is related only to cohorts born between 1916 and 1940. These years of birth are covered by the data sample, so when we project v_H and v_L, future values of the cohort indexes are not required. Consequently, the values of σ_H and σ_L depend entirely on the variability in the period indexes. As the period index for Scottish males is more volatile (see Table 3.2), the value of σ_H in Case 3 is higher than that in Cases 1 and 2, and the value of σ_L in Cases 2 and 3 is higher than that in Case 1.

Recall also that the payouts from the longevity bond are proportional to the realized survival rates. Since Scottish males have heavier mortality than UK males (see Figure 3.1), the value of μ_H in Case 3 is smaller than that in Cases 1 and 2.

Finally, ρ is the lowest (closest to zero) in Case 2, since in this case the longevity bond and the pension liability are linked to different populations.

Using equation (3.7), the approximate prices in Cases 1, 2 and 3 are 13.1372, 13.1219 and 12.3683, respectively. They are quite close to the corresponding prices obtained by the numerical procedure, indicating that the bivariate normal approximation is reasonable.

Now, let us revisit the supply and demand curves in Figure 3.8. The supply curves for Cases 1 and 2 are the same, as μ_H and σ_H remains unchanged when we alter the population to which the pension liability is linked. Nevertheless, the change in the hedger's population would lead to an increase in σ_L and a decrease in ρ. These changes, according to Table 3.4, have offsetting effects on the demand. In this example, the effect of ρ outweighs that of σ_L, and therefore the demand curve shifts downwards.

Similar arguments can explain the shifts in the curves when we move from Case 2 to Case 3. As the bond's reference population is changed from UK males to Scottish males, μ_H decreases, σ_H increases, and ρ increases. According to Table 3.4, a lower μ_H and a higher σ_H have opposite effects on the supply. In this example, the effect of μ_H is more significant and therefore the supply curve shifts upwards. On the other hand, the changes in μ_H and σ_H will exert pressure on the demand. Although a higher ρ will bring the demand up, its effect is not as great as the combined effect of μ_H and σ_H. Overall, the demand curve shifts downwards.

Recall that if the hedger's objective is to maximize its expected utility at a certain future time, then the quantity traded in the tâtonnement equilibrium can be regarded as the corresponding static hedging strategy. From Table 3.3 we observe that, in the presence of population basis risk, the hedger tends to use fewer longevity bonds to static hedge its risk exposure.

We also evaluate the effectiveness of this static hedging strategy over a horizon of 25 years. The measure we use is the reduction in the variance of Agent A's terminal wealth, that is, the wealth at $t = 25$.[5] From Table 3.3 we observe that the static

[5] The terminal wealth is the accumulated value of the initial wealth and all subsequent net cash flows over the hedging horizon. In some other studies, for example, Li and Hardy (2010), the evaluation of hedge effectiveness is based on the present value of the cash flows instead.

	Case 1	Case 2	Case 3
Populations			
Longevity bond	UK males	UK males	CMI males
Pension liability	UK males	CMI males	CMI males
Prices and units traded in equilibrium			
Time-0 price	\$13.0989	\$12.9500	\$14.5588
Units traded	4.3981	2.7394	4.7396
Parameters in the distribution of $(v_l, v_h)'$			
μ_H	26.8379	26.8379	30.3435
σ_H	0.6684	0.6684	0.4744
σ_L	4.4571	3.4242	3.4242
ρ	0.9804	0.8517	0.9749
Effectiveness of the static hedge			
Variance reduction	85.8%	62.5%	84.9%

Table 3.5: Results of the pricing process, UK and CMI males.

hedge is fairly effective, even though it is composed of one single instrument only. In Cases 1 and 3, more than 80% of the variance is elimiated. In Case 2, the population basis risk involved brings down the hedge effectiveness to about 63%.

3.4.5 UK and CMI Males

In this subsection, we repeat the same analysis by assuming that the subpopulation is CMI males. Again we use the algorithm in Section 3.2.2 to solve for the tâtonnement equilibrium in each of the three cases. The prices and the quantities traded are summarized in Table 3.5.

The upper panel of Figure 3.9 shows the demand and supply curves for Cases 1 and 2. The supply curves for both cases are the same, but the demand curve shifts downwards when the population to which the pension liability is linked is changed from UK males to CMI males. This leads to a reduction in the price of the security. Note that the reduction is more dramatic than that observed in the previous example.

The lower panel of Figure 3.9 shows the demand and supply curves for Cases 2 and 3. As we change the population to which the longevity bond is linked, the demand curve shifts upwards, while the supply curve shifts downwards. This results in an increase in the price of the security. It is noteworthy that the changes here are exactly opposite to the corresponding changes in the previous example.

Again we use the analytic formulas in Section 3.2.3 to explain the changes. In Table 3.5 we show the estimates μ_H, σ_H, σ_L, and ρ when the pension liability and the longevity bond are linked to the mortality of either UK or CMI males.

As the period index for CMI males is less volatile (see Table 3.2), the value of σ_H in Case 3 is smaller than that in Cases 1 and 2, and the value of σ_L in Cases 2 and 3 is smaller than that in Case 1. Also, since CMI males have lighter mortality than UK males (see Figure 3.1), the value of μ_H in Case 3 is higher than that in Cases 1 and 2. Further, the difference between the hedger's population and the bond's reference population makes ρ in Case 2 the lowest (closest to zero) among the three cases.

Using equation (3.7), the approximate prices in Cases 1, 2 and 3 are 13.1372, 12.9945 and 14.5909, respectively. By comparing them with the corresponding values in Table 3.5, we find that the bivariate normal approximation is also reasonable in this example.

With the information above, the explanation to the changes becomes straightforward. First, let us focus on Cases 1 and 2. As all parameters in equation (3.17) are unaffected, there is no change to the supply curve. However, the change in the hedger's population will reduce σ_L and ρ. Both reductions will both exert pressure on the demand, causing the demand curve to shift downwards.

Next, we focus on Cases 2 and 3. As the bond's reference population is changed from UK males to CMI males, μ_H increases, σ_H decreases, and ρ increases. According to Table 3.4, these changes will all lead to a higher demand, bringing the demand curve up. On the other hand, the changes in μ_H and σ_H have opposite effects on the

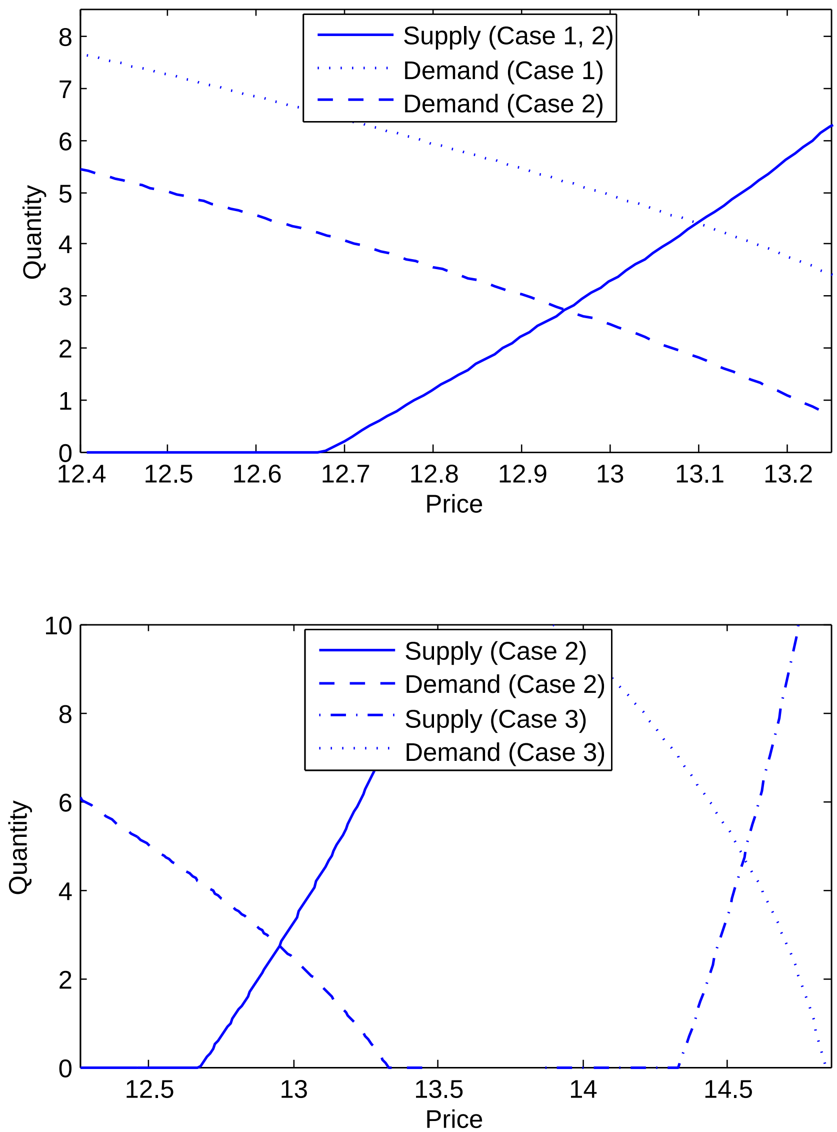


Figure 3.9: Demand and supply curves for the hypothetical longevity bond, UK and CMI males.

supply. In this example, the effect of μ_H is more significant, and therefore the supply curve shifts downwards.

As in the previous example, when population basis risk exists (Case 2), the hedger uses fewer longevity bonds to statically hedge its risk exposure, and the effectiveness of the hedge becomes lower.

3.5 Concluding Remarks

While hedgers may prefer bespoke securities, investors and intermediaries may favor more standardized instruments which are easier to analyze and more conductive to the development of liquidity. Recently, the Life and Longevity Markets Association (LLMA) has been setup to promote the development of a liquid traded market in longevity and mortality related risk. Part of LLMA's work is to develop standardized longevity indexes, upon which securities with a secondary market can be written.

The use of standardized instruments, as we mentioned at the outset, would expose hedgers to population basis risk. According to Coughlan (2010), the lack of knowledge in population basis risk is a major challenge to market development. In this chapter, we addressed this issue by introducing a tâtonnement process for pricing mortality-linked securities when population basis risk exists. The proposed method is highly transparent, allowing us to understand how population basis risk would affect security prices and hedging strategies.

From the solution to the tâtonnement pricing process, we can conclude that, on top of correlations, the magnitude and volatility of mortality rates play a critical role in determining the impact of population basis risk on security prices. As these factors are different among different pairs of populations, the effect of the risk varies. Our illustrations have shown empirically that, interestingly, population basis risk can affect the price of a mortality-linked security in different directions, depending on the properties of the populations involved.

It might be difficult to conduct a similar analysis with a no-arbitrage approach. Technically speaking, we could construct a risk-adjusted two-population mortality model by introducing market prices of risk to the stochastic factors, $\kappa_t^{(1)}$, $\gamma^{(1)}(c)$, $\Delta_\kappa(t)$, and $\Delta_\gamma(c)$. However, in today's market where market price data are very limited, it is difficult, if not impossible, to estimate that many market prices of risk, which are associated with multiple reference populations. Even if we have such a model, a no-arbitrage approach does not give us information such as demand and supply, which would help us better understand how population basis risk and its components are involved in the pricing process.

One reason for using standardized instruments is that they potentially have better liquidity. In our presentation, liquidity was not given much attention, as we assumed that the agents will hold their positions until the security matures. This assumption can be relaxed by employing a sequential decision process, which is detailed in Chapter 2. This extension allows both agents to unwind their positions at discrete time points before maturity. However, as decisions (optimizations) are made at multiple time points, this extension demands significantly more computational resources, adding an extra challenge in the pricing process.

Chapter 4

The Impact of Mortality Jumps on Trading

4.1 Introduction

As discussed in Chapter 3, the trading of mortality or longevity risk often involves two populations, one of which is associated with the hedger's portfolio, and the other of which is associated with the hedging instrument. Taking the mortality bond issued by Swiss Re in December 2003 as an example, the index to which the bond is linked is based on the realized mortality rates for some national populations, but the exposure of the hedger (Swiss Re) is associated with some insured lives. To adequately model trades involving more than one populations, a multi-population mortality model is necessary.

Multi-population mortality models take account of the potential correlations across different populations, and more importantly, they are structured in such a way that the resulting forecasts are biologically reasonable. In particular, they ensure that the forecasted life expectancies of two related populations do not diverge over the long run. Other than pricing, multi-population mortality models enable us to evaluate population basis risk, which arises from the difference in mortality experience between the hedger's population and the population associated with the hedging instrument.

In recent years, a few two-population mortality models have been proposed. These include the joint-k model (Carter and Lee, 1992), the augmented common factor model (Li and Lee, 2005), the co-integrated Lee-Carter model (Li and Hardy, 2011), the gravity model (Dowd et al., 2011), and the model proposed by Cairns et al. (2011a). These models, which are adapted from well-known single-population models, differ in the way in which the relation between the period/cohort effect indexes of the two populations is modeled. For example, Cairns et al. (2011a) fit a classical age-period-cohort model to each of the two populations in question, and use stationary time-series processes to model the difference between the period effect indexes and that between the cohort effect indexes.

Another factor that should be taken into account is mortality jumps, which are due to interruptive events such as the Spanish flu epidemic in 1918. It is important not to ignore mortality jumps in modeling, because otherwise we can seriously understate the uncertainty surrounding a central mortality projection. The incorporation of jumps is particularly important when pricing securities for hedging extreme mortality risk, because this allows us to better estimate the probability of having a catastrophic mortality deterioration. Models incorporating mortality jumps include those proposed by Biffs (2005), Chen and Cox (2009), Cox et al. (2010) and Deng et al. (2012).

Nevertheless, to date, there is no two-population model that incorporates mortality jumps. In this chapter, we fill this gap by developing a two-population mortality model with transitory jump effects. This extension is not straightforward, because the correlations between jump times and jump severities of the two populations in question have to be carefully modeled. Moreover, to construct the likelihood function on which parameter estimation is based, a careful conditioning on the jump counts is required. The jump model we develop allows practitioners to more accurately estimate prices of mortality-linked securities and the population basis risk in index-based mortality/longevity hedges.

Based on the proposed mortality model, we examine the impact of mortality jumps on the trading of mortality/longevity risk. We consider the pricing framework proposed in Chapter 3. This pricing framework allows the trade of a mortality-linked security between two counterparties, whose portfolios can be related to different populations. Besides the estimated price, this pricing framework provides us with a pair of demand and supply curves, which helps us better understand the effect of introducing mortality jumps on the behaviors of the counterparties. We find that, interestingly, the inclusion of mortality jumps does not always increase the estimated price of a mortality securitization. The effect on the estimated price is very much dependent on the structure of the security.

The rest of the paper is organized as follows. In Section 4.2, we review the basic two-population mortality model on which our proposed extensions are built. In Section 4.3, we incorporate concurrent transitory mortality jumps into the basic two-population model, and illustrate this extension with data from US total and US male populations. In Section 4.4, we further extend the model to permit nonconcurrent mortality jumps, and illustrate this extension with data from Swedish male and Finnish male populations. In Section 4.5, we examine the impact of mortality jumps on the trading of mortality-linked securities through the economic pricing framework. Finally, we conclude this chapter in Section 4.6.

4.2 A Two-Population Model without Jump Effects

We construct a two-population mortality model from two classical Lee-Carter models, one for each population. This basic model does not take mortality jumps into account. Thereafter, we call it the no-jump model for simplicity.

The no-jump model can be expressed as follows:

$$\ln(m_{x,t}^{(i)}) = \alpha_x^{(i)} + \beta_x^{(i)}\kappa_t^{(i)}, \quad i = 1, 2, \tag{4.1}$$

where $m_{x,t}^{(i)}$ denotes the central death rate for population i at age x and in year t, $\kappa_t^{(i)}$ is the period effect index for population i in year t, $\alpha_x^{(i)}$ measures the average level of mortality for population i at age x, and $\beta_x^{(i)}$ measures the sensitivity to the period effect index $\kappa_t^{(i)}$ for population i at age x.

It is reasonable to expect the death rates for two related population not to diverge in the long run (see, e.g., White, 2002; Wilson, 2001; United Nations, 1998). The necessary conditions for non-divergence are:

1. $\beta_x^{(1)} = \beta_x^{(2)}$ for all x;
2. $\kappa_t^{(1)}$ and $\kappa_t^{(2)}$ do not diverge over the long run.

Accordingly, we set $\beta_x^{(1)} = \beta_x^{(2)} = \beta_x$, and model the difference between $\kappa_t^{(1)}$ and $\kappa_t^{(2)}$ with a stationary first order autoregressive process. In more detail, the dynamics of $\kappa_t^{(1)}$ and $\kappa_t^{(2)}$ are specified as follows:

$$\begin{aligned}
\kappa_{t+1}^{(1)} &= \kappa_t^{(1)} + \mu_\kappa + Z_\kappa(t+1), \\
\Delta_\kappa(t) &= \kappa_t^{(1)} - \kappa_t^{(2)}, \\
\Delta_\kappa(t+1) &= \mu_{\Delta_\kappa} + \phi_{\Delta_\kappa}\Delta_\kappa(t) + Z_{\Delta_\kappa}(t+1),
\end{aligned}$$

where $|\phi_{\Delta_\kappa}| < 1$ and $\{(Z_\kappa(t), Z_{\Delta_\kappa}(t))'\}$ is a sequence of independent and identically distributed (i.i.d.) bivariate normal random vectors with mean zero and variance-covariance matrix v_Z. The specification above implies that $\Delta_\kappa(t)$ will revert to a constant $\mu_{\Delta_\kappa}/(1-\phi_{\Delta_\kappa})$ over the long run.

We use a two-stage approach to fit the no-jump model. In the first stage, we estimate parameters $\alpha_x^{(1)}$, $\alpha_x^{(2)}$, β_x, $\kappa_t^{(1)}$ and $\kappa_t^{(2)}$ by the method of maximum likelihood. Let $[x_0, x_1]$ and $[t_0, t_1]$ be the sample age range and sample period, respectively. Assuming Poisson death counts, the log-likelihood function can be expressed as

$$\sum_{x=x_0}^{x_1}\sum_{t=t_0}^{t_1}\sum_{i=1}^{2}\Big(D_{x,t}^{(i)}(\alpha_x^{(i)} + \beta_x\kappa_t^{(i)}) - E_{x,t}^{(i)}\exp(\alpha_x^{(i)} + \beta_x\kappa_t^{(i)})\Big) + c,$$

where c is a constant that does not depend on the model parameters, $D_{x,t}^{(i)}$ is the observed number of deaths for population i at age x and in year t, and $E_{x,t}^{(i)}$ is the corresponding number of persons at risk. The log-likelihood function can be maximized by an iterative Newton-Raphson method, in which parameters are updated one at a time.

In the second stage, we estimate the parameters in the time-series processes for $\kappa_t^{(1)}$ and $\Delta_\kappa(t)$. We let

$$S_t = (\kappa_{t+1}^{(1)} - \kappa_t^{(1)}, \Delta_\kappa(t+1) - \phi_{\Delta_\kappa}\Delta_\kappa(t))'.$$

Since $\{(Z_\kappa(t), Z_{\Delta_\kappa}(t))'\}$ is a sequence of i.i.d. bivariate normal random vectors, $\{S_t\}$ is also a sequence of i.i.d. bivariate normal random vectors. The mean and variance-covariance matrix of S_t are $(\mu_\kappa, \mu_{\Delta_\kappa})'$ and V_Z, respectively. It follows that the log-likelihood function for the processes can be expressed as

$$\begin{aligned}
& \ln f\left(S_1, S_2, S_3, \ldots, S_{T-1} | \mu_\kappa, \mu_{\Delta_\kappa}, \phi_{\Delta_\kappa}, V_Z\right) \\
= & \sum_{i=t_0}^{t_1-1} \ln f\left(S_i | \mu_\kappa, \mu_{\Delta_\kappa}, \phi_{\Delta_\kappa}, V_Z\right) \\
= & \sum_{i=t_0}^{t_1-1} \ln \operatorname{bvnpdf}(S_i, (\mu_\kappa, \mu_{\Delta_\kappa})', V_Z),
\end{aligned}$$

where $\operatorname{bvnpdf}(s, \mu, v)$ is the probability density function of a bivariate random vector with mean μ and variance-covariance v. The estimates of μ_κ, μ_{Δ_κ}, ϕ_{Δ_κ} and V_Z can be obtained by maximizing the log-likelihood function.

Alternatively, we can fit the entire model by using the one-stage Bayesian approach considered by Cairns et al. (2011a). As Cairns et al. (2011a) mentioned, this one-stage method have several advantages, for example, it can effectively handle missing data. We stay with the two-stage approach, mainly because it allows us to save some computational effort. Specifically, it avoids the need to re-estimate parameters in equation (4.1) when mortality jumps are introduced to the time-series processes for $\kappa_t^{(1)}$ and $\Delta_\kappa(t)$ later in this paper.

4.3 Concurrent Transitory Jumps

4.3.1 Mortality Data

In this section, we use US total and US male populations to illustrate the necessity and appropriateness of incorporating concurrent transitory jumps. Note that the latter population is a subset of the former. For both populations, we consider a sample period of 1900 to 2006 and a sample age range of 25 to 84. The data, which are obtained from the Nation Center for Health Statistics, are provided by decennial age groups.[1]

4.3.2 An Outlier Analysis

In Figure 4.1 we show the estimated values of the period effect indexes $\kappa_1^{(1)}$ (for US total population) and $\kappa_1^{(2)}$ (for US male population). In both indexes, we observe several jumps, most notably the one that corresponds to year 1918. To better understand the jumps in the period effect indexes, we herein perform a statistical outlier analysis.

We consider two types of outliers:

1. Additive Outliers
   An additive outlier (AO) affects only one single observation.
2. Temporary Changes
   A temporary change (TC) affects a series at a given time, and its effect decays at an exponential rate.

There are other types of outliers (see, e.g., Chen and Tiao, 1990; Tsay, 1988). However, since the focus of this paper is on short-term catastrophic mortality risk, we consider outliers that have short-term effects only.

[1] The required data (death and exposure counts) for years 1900 to 1998 are available at http://www.cdc.gov/nchs/nvss/mortality_historical_data.htm, while those for years 1999 to 2006 are available from CDC WONDER at http://wonder.cdc.gov/cmf-ICD10.html.

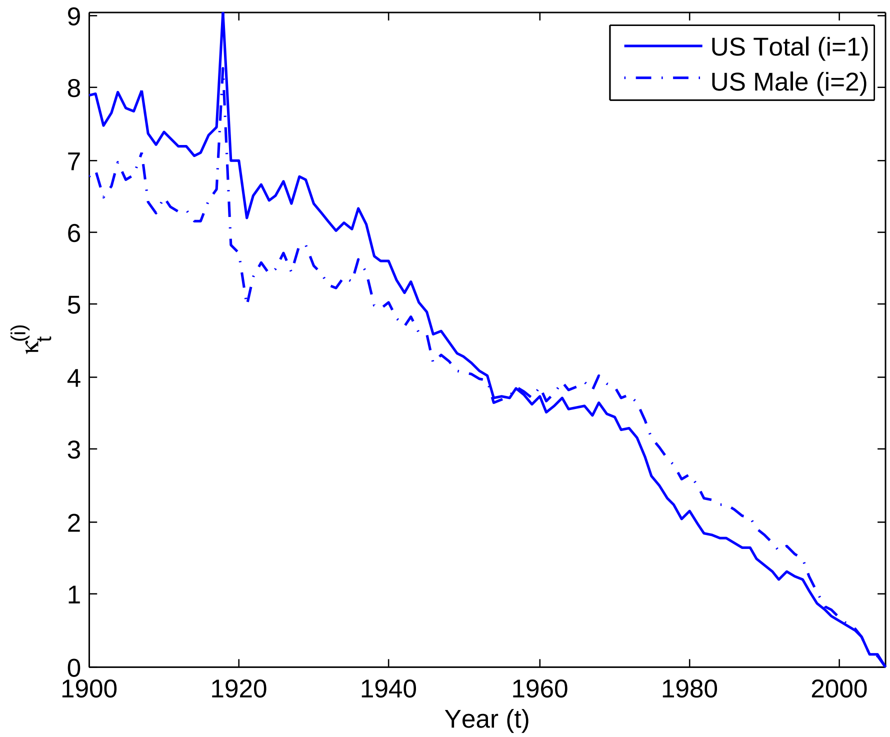


Figure 4.1: Estimates of the period indexes for US total and US male populations.

We use Chen and Liu's (1993) procedure to identify outliers in the period effect indexes $\kappa_1^{(1)}$ and $\kappa_1^{(2)}$. The procedure can be implemented with standard statistical software for time-series analysis, such as AUTOBOX, SAS/ETS and SCA.

In Table 4.1 we show the outliers identified in the period effect indexes for the two populations. Note that a negative outlier stands for an improvement in mortality, whereas a positive one means a deterioration. The identified outliers are broadly in line with the results produced by Li and Chan (2007).[2] We refer interested readers to Li and Chan (2007) for a discussion of the events that may have caused the outliers

[2] Li and Chan (2007) performed an outlier analysis on the Lee-Carter period effect index for US total population. They considered two additional types of outliers and used data up to year 2000 only.

Population	Year	Magnitude	Type
US Total	1907	0.431	AO
	1918	1.828	AO
	1921	-0.764	TC
	1927	-0.356	AO
	1936	0.406	TC
US Male	1907	0.503	AO
	1918	2.056	AO
	1921	-0.737	TC
	1927	-0.337	AO
	1936	0.419	TC

Table 4.1: Outliers detected in the period effect indexes for US total and US male populations.

found.

Five outliers are detected for both populations. Moreover, the locations of the detected outliers are the same for both populations. The results of this outlier analysis suggest that if the two populations being modeled are closely related to each other, we may assume that the two populations have identical frequency and timing of jumps. Such an assumption can make the resulting model more parsimonious and easier to estimate.

4.3.3 Modeling Concurrent Transitory Jumps

We now present a two-population model that permits concurrent mortality jumps. This model has the same structure as equation (4.1), but the period effect indexes are modeled in a different manner.

We let $\kappa_t^{(i)}$, $i = 1, 2$, be the observed period effect index for population i at time t. As seen from Figure 4.1, the process $\{\kappa_t^{(i)}\}$, $i = 1, 2$, is subject to transitory jumps.

We decompose $\kappa_t^{(i)}$ into a sum of two components, $\hat{\kappa}_t^{(i)} + N_t Y_t^{(i)}$. The first component, $\hat{\kappa}_t^{(i)}$, is time-t value of an unobserved period effect index that is free of jumps, and the second term, $N_t Y_t^{(i)}$, represents the jump effect at time t. Specifically, N_t denotes the jump count at time t, while $Y_t^{(i)}$ denotes the jump severity for population i at time t.

As in the no-jump model, we assume that $\hat{\kappa}_t^{(1)}$ follows a random walk with drift and that the difference $\hat{\Delta}_\kappa(t) = \hat{\kappa}_t^{(1)} - \hat{\kappa}_t^{(2)}$ follows a stationary first order autoregressive process. The latter assumption ensures that mortality rates for the two populations do not diverge over the long run.

Summing up, the period effect indexes are modeled by the following set of equations:

$$\begin{aligned}
\hat{\kappa}_{t+1}^{(1)} &= \hat{\kappa}_t^{(1)} + \mu_\kappa + Z_\kappa(t+1), \\
\kappa_{t+1}^{(1)} &= \hat{\kappa}_{t+1}^{(1)} + N_{t+1} Y_{t+1}^{(1)}, \\
\hat{\Delta}_\kappa(t) &= \hat{\kappa}_t^{(1)} - \hat{\kappa}_t^{(2)}, \\
\hat{\Delta}_\kappa(t+1) &= \mu_{\Delta_\kappa} + \phi_{\Delta_\kappa} \hat{\Delta}_\kappa(t) + Z_{\Delta_\kappa}(t+1), \\
\kappa_{t+1}^{(2)} &= \hat{\kappa}_{t+1}^{(2)} + N_{t+1} Y_{t+1}^{(2)}.
\end{aligned}$$

The error terms $Z_\kappa(t)$ and $Z_{\Delta_\kappa}(t)$ jointly follow a zero-mean bivariate normal distribution with a variance-covariance matrix V_Z. They have no serial dependence and are independent of the jump counts and jump severities.

We assume that there is at most one jump in a year and that the jump frequency p is constant over time. It follows that N_t, the jump count at time t, follows a binomial distribution with $\Pr(N_t = 1) = p$ and $\Pr(N_t = 0) = 1 - p$. The jump counts are not serially correlated.

We allow both positive and negative jumps. The jump severities $Y_t^{(1)}$ and $Y_t^{(2)}$ can be different from each other, but they are correlated. Specifically, we assume that the jump severity vector $(Y_t^{(1)}, Y_t^{(2)})'$ follows a bivariate normal distribution with mean μ_Y and variance-covariance matrix V_Y. We further assume that the jump severities are not serially correlated and that they are independent of the jump counts.

As the introduction of transitory jumps does not affect equation (4.1), there is no need to re-estimate the parameters in equation (4.1). We do, however, need to estimate the parameters in the processes for the period effect indexes. This can be accomplished by the method of maximum likelihood. The derivation of the likelihood function for the period effect processes is presented in Appendix B.

We fit the concurrent-jump model to US total and US males populations. To evaluate the benefit from introducing concurrent jumps, we also fit the no-jump model to this pair of populations. Estimated parameters in the concurrent-jump model and the no-jump model are shown in the second and third columns of Table 4.2, respectively.

Since the two models are nested, we can use the likelihood ratio test to examine if the more complex model gives a significantly better fit. The test statistic is $2(l_2 - l_1)$, where l_1 and l_2 are the maximized log-likelihood values for the null model (the no-jump model) and the alternative model (the model with jump effects). We have $l_1 = 147.7368$ and $l_2 = 218.5685$. The test statistic is 141.6633 and the degree of freedom, which equals the difference in the number of parameters, is six. This gives a p-value of zero, which means the model with jump effects gives a significantly better fit than the no jump model.

4.4 Nonconcurrent Transitory Jumps

4.4.1 Mortality Data

In the previous example, one population is a sub-population of the other. As the two populations are closely related to each other, we expect that they are subject to mortality jumps at the same time. A model that permits concurrent transitory jumps is used for this pair of populations.

What if the relation between the two populations being modeled is not that obvious? For instance, sometimes one may need to model two populations that are geographically farther apart. In this situation, the assumption of concurrent jumps might be too stringent.

Parameters	Transitory jumps	No jump
μ_κ	-0.0695	-0.0745
μ_{Δ_κ}	-0.0105	-0.0075
ϕ_{Δ_κ}	0.9807	0.9874
$V_Z(1,1)$	0.0253	0.0946
$V_Z(1,2)$	-0.0005	-0.0088
$V_Z(2,2)$	0.0012	0.0030
$\mu_Y^{(1)}$	-0.0863	N/A
$\mu_Y^{(2)}$	-0.2423	N/A
$V_Y(1,1)$	1.8023	N/A
$V_Y(1,2)$	2.0278	N/A
$V_Y(2,2)$	2.2816	N/A
p	0.0404	N/A

Table 4.2: Estimated parameters in the processes for the period effect indexes, US total and US male populations. (We use $X(i,j)$ to denote the (i,j)th element in a matrix X.)

In what follows, we further generalize the no-jump model to a model that permits nonconcurrent transitory jumps. We illustrate this generalization with historical mortality data from Swedish male and Finnish male populations. For both populations, we consider a sample period of 1900 to 2006 and a sample age range of 25 to 84. The required data (death and exposure counts) are obtained from Human Mortality Database (2011).

4.4.2 An Outlier Analysis

We first estimate the period effect indexes for Swedish males and Finnish males. The estimates are shown graphically in Figure 4.2. Similar to the previous example, the period effect indexes are subject to several jumps. However, as opposed to the previous pair of populations, the jump patterns for Swedish males and Finnish males

Population	Year	Magnitude	Type
Swedish Males	1918	28.837	AO
Finnish Males	1918	49.033	AO
	1939	18.006	TC
	1940	31.655	TC
	1941	17.602	AO
	1944	25.307	AO

Table 4.3: Outliers detected in the mortality indexes for Swedish male and Finnish male populations.

are quite different from each other.

To have a deeper understanding of the jump patterns, we perform an outlier analysis for these two populations. The identified outliers are displayed in Table 4.3. In sharp contrast to the previous example, these two populations have very different timing and frequency of jumps. For Finnish males, five outliers are detected, but for Swedish males, only one outlier is identified. For both populations, an outlier is found in 1918 when the Spanish flu epidemic occurred. On the other hand, there are outliers that affect one population only. For Finnish males, outliers are found in 1939, 1940, 1941 and 1944. These outliers may be attributed to Finland's participation in the Second World War, which lasted from 1939 to 1945. Similar outliers are not found in the period effect index for Swedish males, possibly because Sweden remained neutral in the war.

For this pair of populations, the assumption of concurrent jumps no longer holds. A model that permits the two populations to have different timing and frequency of jumps seems necessary.

4.4.3 Modeling Nonconcurrent Jumps

To introduce nonconcurrent jumps, we modify the way in which the period effect indexes are modeled. As before, we let $\kappa_t^{(i)}$, $i = 1, 2$, be the observed period effect

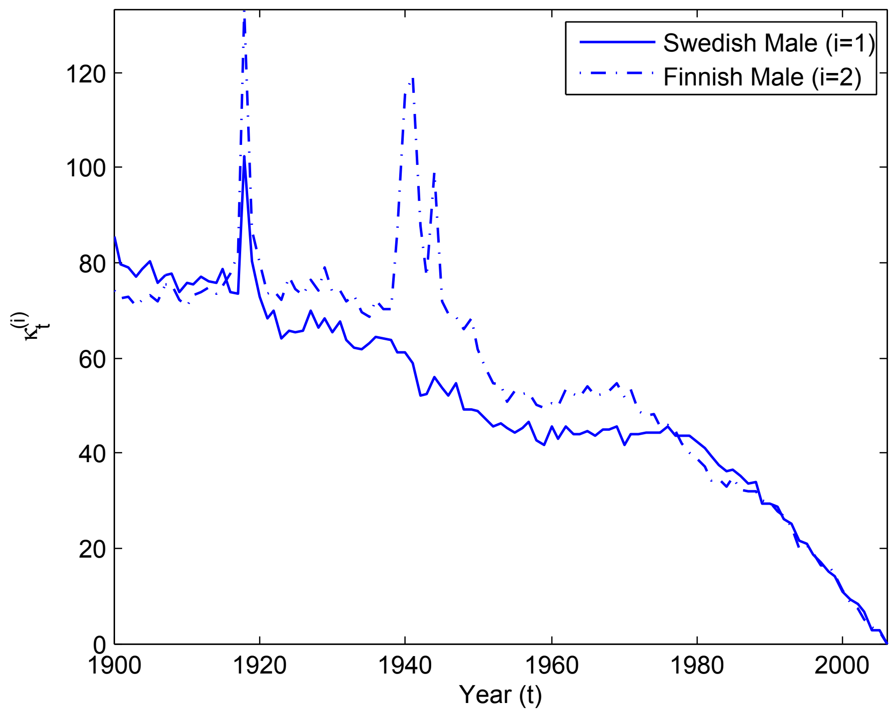


Figure 4.2: Estimates of the period effect indexes for Swedish male and Finnish male populations.

index for population i at time t. The observed index $\kappa_t^{(i)}$ is decomposed into the sum of two components, $\hat{\kappa}_t^{(i)} + N_t^{(i)} Y_t^{(i)}$, where $\hat{\kappa}_t^{(i)}$ is time t value of an unobserved period effect index that is free of jumps, and $N_t^{(i)} Y_t^{(i)}$ is the jump effect at time t. In this generalization, the jump count $N_t^{(i)}$ depends on the population, thereby allowing the two populations to have different timing and frequency of jumps.

As in the concurrent-jump model, we assume that $\hat{\kappa}_t^{(1)}$ follows a random walk with drift and that $\hat{\Delta}_\kappa(t) = \hat{\kappa}_t^{(1)} - \hat{\kappa}_t^{(2)}$ follows a stationary first order autoregressive process. Summing up, the period effect indexes in this generalization are modeled by the following set of equations:

$$\begin{aligned}
\hat{\kappa}_{t+1}^{(1)} &= \hat{\kappa}_t^{(1)} + \mu_\kappa + Z_\kappa(t+1), \\
\kappa_{t+1}^{(1)} &= \hat{\kappa}_{t+1}^{(1)} + N_{t+1}^{(1)} Y_{t+1}^{(1)}, \\
\hat{\Delta}_\kappa(t) &= \hat{\kappa}_t^{(1)} - \hat{\kappa}_t^{(2)}, \\
\hat{\Delta}_\kappa(t+1) &= \mu_{\Delta_\kappa} + \phi_{\Delta_\kappa} \hat{\Delta}_\kappa(t) + Z_{\Delta_\kappa}(t+1), \\
\kappa_{t+1}^{(2)} &= \hat{\kappa}_{t+1}^{(2)} + N_{t+1}^{(2)} Y_{t+1}^{(2)}.
\end{aligned}$$

The error terms $Z_\kappa(t)$ and $Z_{\Delta_\kappa}(t)$ jointly follow a zero-mean bivariate normal distribution with a variance-covariance matrix V_Z. They have no serial dependence and are independent of the jump counts and severities.

For both populations, we assume that there is at most one jump in a given year, since we are using annual data and we consider only transitory mortality jumps . It follows that the joint probability mass function for $N_t^{(1)}$ and $N_t^{(2)}$ can be specified as follows:

$$\begin{aligned}
\Pr(N_t^{(1)} = 1, N_t^{(2)} = 1) &= p_1, \\
\Pr(N_t^{(1)} = 1, N_t^{(2)} = 0) &= p_2, \\
\Pr(N_t^{(1)} = 0, N_t^{(2)} = 1) &= p_3, \\
\Pr(N_t^{(1)} = 0, N_t^{(2)} = 0) &= 1 - p_1 - p_2 - p_3,
\end{aligned}$$

where p_1, p_2 and p_3 are non-negative constants. The jump counts in different years are independent of one another. Notice that the concurrent-jump model is a special case of the nonconcurrent-jump model. Specifically, we recover the concurrent jump model if we set $p_1 = p$ and $p_2 = p_3 = 0$.

The jump severities $Y_t^{(1)}$ and $Y_t^{(2)}$ jointly follow a bivariate normal distribution with mean μ_Y and variance-covariance matrix V_Y. It is assumed that the jump severities are independent of the jump counts and are not serially correlated.

As before, we estimate the processes for the period effect indexes by the method of maximum likelihood. The derivation of the log-likelihood function is presented in Appendix B.

Parameters	Nonconcurrent jumps	Concurrent jumps	No jump
μ_κ	-0.6372	-0.6539	-0.8057
μ_{Δ_κ}	-0.1269	-0.2244	-0.9803
ϕ_{Δ_κ}	0.9184	0.9094	0.8486
$V_\kappa(1,1)$	4.1752	4.2403	17.3162
$V_\kappa(1,2)$	1.0633	1.1018	-11.0413
$V_\kappa(2,2)$	2.7786	3.0961	38.1830
$\mu_Y^{(1)}$	3.5824	2.7746	N/A
$\mu_Y^{(2)}$	12.4430	13.2352	N/A
$V_Y(1,1)$	116.3952	96.3418	N/A
$V_Y(1,2)$	184.9353	161.5611	N/A
$V_Y(2,2)$	293.8356	422.0060	N/A
p	N/A	0.0907	N/A
p_1	0.0622	N/A	N/A
p_2	0	N/A	N/A
p_3	0.0496	N/A	N/A

Table 4.4: Estimated parameters in the processes for the period effect indexes, Swedish male and Finnish male populations. (We use $X(i,j)$ to denote the (i,j)th element in a matrix X.)

We fit the nonconcurrent-jump model to the period effect indexes for Swedish males and Finnish males. The parameter estimates are depicted in the second column of Table 4.4. Recall that in the outlier analysis, we cannot identify any outlier affecting Swedish males but not Finnish males. It is therefore not surprising that the estimate of p_2 is zero. As a matter of fact, the estimates of other parameters remain the same if we impose the constraint that $p_2 = 0$.

We also fit the concurrent-jump model and the no-jump model to the period indexes for Swedish males and Finnish males. The estimates of the parameters in these two more restrictive models are displayed in the third and fourth columns of Table 4.4. We can then perform likelihood ratio tests to compare the three models we fitted.

First, we compare the no-jump model with the concurrent-jump model. The log-likelihood for the null model (the no-jump model) is -634.1980, while that for the alternative model (the concurrent-jump model) is -506.9093. The value of the likelihood ratio test statistic is $2 \times (-506.9093 - (-634.1980)) = 127.2887$ and the degree of freedom is 6. This results in a p-value of 0, indicating that the concurrent-jump model gives a significantly better fit than the no-jump model.

Next, we evaluate the benefit from permitting the two populations to have different timing and frequency of jumps. The log-likelihood for the null model (the concurrent-jump model) is -506.9093, while that for the alternative model (the nonconcurrent-jump model) is -503.0726. The value of the likelihood ratio test statistic is $2 \times (-503.0726 - (-506.9093)) = 7.6734$ and the degree of freedom is one.[3] This gives a p-value of 0.0056, which means that at a significance level of 1%, the nonconcurrent jump model is significantly better than the concurrent jump model.

One may wonder if permitting nonconcurrent jumps would also improve the fit to the period indexes for US total and US male populations. When the nonconcurrent-jump model is fitted to these two populations, the estimates of p_2 and p_3 are both zero, and the estimates of the other parameters are exactly the same as those when only concurrent jumps are permitted. Hence, the allowance of nonconcurrent jumps does not improve the fit at all. The more parsimonious concurrent-jump model is adequate for this particular pair of populations.

4.5 The Impact on Mortality Risk Securitization

4.5.1 An Illustrative Trade

Transitory mortality jumps can result in significant losses to life insurers. At the same time, they affect payouts from securities that are designed for hedging mortality risk.

[3] We imposed the constraint that $p_2 = 0$. Hence, the nonconcurrent jump model has one more parameter than the concurrent jump model.

In this section, we use a hypothetical trade to illustrate how the incorporation of transitory jumps may affect the estimated price of a mortality-linked security.

We consider a trade between two economic agents, Agent A and Agent B. Agent A is a life insurer that holds a portfolio of 10,000 life insurance policies. These policies are issued to Finnish males and are uniformly distributed over the age range of 25 to 44. For each policy, the death benefit is \$0.01 payable at the end of the year of the policyholder's death. For simplicity, we assume that this portfolio is stationary, by which we mean the age composition does not change over time.

It is obvious that Agent A's financial obligation is linked to the mortality of Finnish males at ages 25 to 44. Assuming no small sample risk, the insurance liability due at time t is $L_t = 5\sum_{x=25}^{44} q_{x,t}^{(2)}$, where $q_{x,t}^{(2)}$ is the death probability of Finnish males at age x and in year t. We calculate $q_{x,t}^{(2)}$ by assuming a constant force of mortality between integral ages, which means $q_{x,t}^{(2)} = 1 - e^{-m_{x,t}^{(2)}}$.

To mitigate its exposure to catastrophic mortality risk, Agent A sold in 2006 a mortality bond maturing in three years.[4] For liquidity considerations, the bond is linked to a mortality index that is based on Swedish male population, which is larger than the population associated with the insurance liability. The index is the simple arithmetic average of the central death rates for Sweden males aged between 25 to 44. Mathematically, the time-t value of the index is given by

$$I_t = \frac{1}{20}\sum_{x=25}^{44} m_{x,t}^{(1)}.$$

As of the mortality bond was traded, the values of I_t for $t = 2007, 2008, \ldots$ were not known.

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 calculations. The principal repayment at maturity depends on the values of I_t over the term of the

[4]The last year in the sample period we consider is 2006.

security. Specifically, the principal repayment is specified as follows:

$$\text{Principal Repayment} = \max\left(1 - \sum_{t=2007}^{2009} \text{loss}_t, 0\right),$$

where loss_t is defined by

$$\text{loss}_t = \frac{\max(I_t - aI_{2006}, 0) - \max(I_t - bI_{2006}, 0)}{(a-b)I_{2006}},$$

and a and b are the attachment and exhaustion points. If the mortality index ever exceeds aI_{2006}, then the principal repayment will be reduced, and if the mortality index ever exceeds bI_{2006}, the principal repayment will be exhausted.

We illustrate the trade with two sets of attachment and exhaustion points: (i) $a = 1.3$, $b = 1.4$; (ii) $a = 1.2$, $b = 1.3$. The erosion of capital by an increase in the mortality index is demonstrated in Figure 4.3.

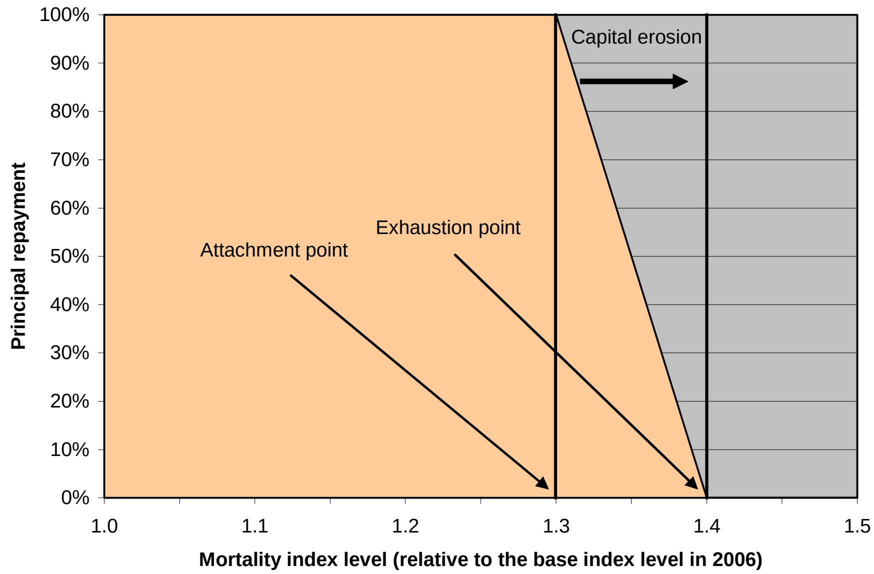
Agent B is an investor who traded the mortality bond with Agent A in 2006 for earning a risk premium. At $t = 2007, 2008, 2009$, Agent B receives payouts from the mortality bonds purchased.

4.5.2 Pricing the Mortality Bond

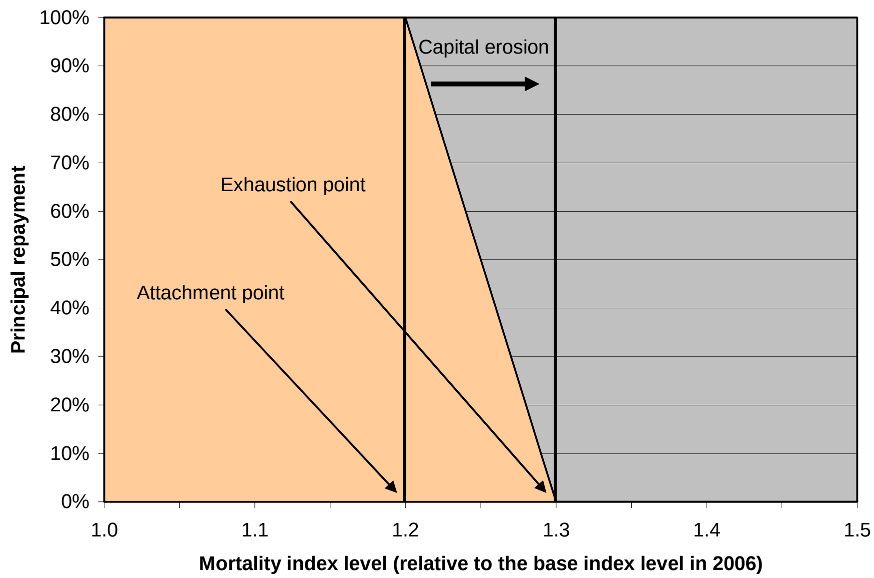
We use the pricing framework proposed in Chapter 3 to price the mortality bond. The framework models the trade between Agents A and B. It is assumed that, given a price, both agents would maximize their expected terminal utility by altering their demand or supply of the security. The price of the security is adjusted iteratively to match the supply and demand, and finally, the estimated price is the price at which the demand and supply are equal, that is, the market clears.

We assume an exponential utility function, $U(x) = 1 - e^{-kx}$, for each agent. Parameter k in the utility function is the absolute risk aversion for all wealth levels. As before, we assume that $k = 1.0$ for Agent A and that $k = 0.5$ for Agent B.

To estimate the price of the mortality bond, we first simulate 10,000 mortality paths from a two-population mortality model. Given the simulated paths, we can use



$a = 1.3,\ b = 1.4$



$a = 1.2,\ b = 1.3$

Figure 4.3: Reduction in principal repayment at different index levels.

Model	$a = 1.3, b = 1.4$		$a = 1.2, b = 1.3$	
	Price	Quantity traded	Price	Quantity traded
No-jump	1.0178	0.2249	0.9748	0.1748
Nonconcurrent-jump	1.0087	0.1699	0.9930	0.1455

Table 4.5: Prices and quantities traded in equilibrium when different mortality models are assumed.

the iterative procedure provided in Chapter 3 to find the price and quantity traded in equilibrium. On the basis of the no-jump model and the nonconcurrent-jump model we fitted in Section 4.4.3, we obtain two sets of bond prices. The estimated prices are displayed in Table 4.5.

As expected, the price of the mortality bond is lower if the attachment and detachment points are smaller. This is because in this case it is more likely that the index I_t will exceed the attachment and detachment points, reducing the expected principal repayment and consequently the price of the bond.

A more interesting observation is that the effect of introducing transitory jumps depends on the attachment and detachment points. The no-jump model produces the highest price when $a = 1.3$ and $b = 1.4$, but the opposite is true when $a = 1.2$ and $b = 1.3$. We will explain the reasons underlying this observation later in this section when we analyze the demand and supply of the mortality bond.

4.5.3 Determinants of Supply and Demand

The economic pricing method we consider is more transparent relative to standard no-arbitrage approaches, in which the price of a security is estimated by extrapolating prices of other similar securities available in the market. On top of the estimated price, the economic pricing method provides us with a pair of demand and supply curves. By examining the response of the demand and supply curves to changes in different factors, we can have a better idea about how and why the price of a mortality-linked security will change under different circumstances.

Let v_L and v_H be the accumulated values (when the bond matures) of the insurance liabilities and the payouts from one unit of the bond, respectively. According to Chapter 3, the supply and demand of the mortality bond depend heavily on the following four factors:

- μ_H, the expected value of v_H
  At a given price, when μ_H increases, Agent B (the buyer) is expected to receive more payouts from the mortality bond, while Agent A (the seller) is expected to make more payouts. Hence, an increase in μ_H would lead to an increase in the demand from Agent B and a reduction in the supply from Agent A.

- σ_H, the volatility of v_H
  When σ_H increases, the security becomes less attractive to Agent B, as he/she needs to take more risk for the same expected payoff from the bond. On the other hand, when σ_H increases, Agent A can achieve the same amount of risk reduction by selling fewer units of the mortality bond. As a result, an increase in σ_H would lead to a reduction in both supply and demand.

- σ_L, the volatility of v_L
  A higher σ_L means that Agent A is subject to more mortality risk. Therefore, it has a stronger need to sell the mortality bond, leading to a greater supply. However, because Agent B's behavior depends only on the mortality bond, a change in σ_L would not affect the demand from Agent B.

- ρ, the correlation between v_L and v_H
  When $|\rho|$ is higher, the mortality bond becomes a more effective hedging instrument. Therefore, at a given price, Agent A is willing to supply more. Same as σ_L, ρ does not affect Agent B's behavior. Hence, a change in ρ would not affect the demand from Agent B.

The effects of these four factors on the supply and demand of the mortality bond are summarized in Table 4.6. In the table, '$\uparrow$' means an increase, '$\downarrow$' means an decrease, and '$-$' means there is no change.

	$\mu_H \uparrow$	$\sigma_H \uparrow$	$\sigma_L \uparrow$	$\lvert\rho\rvert \uparrow$
Supply	$\downarrow$	$\downarrow$	$\uparrow$	$\uparrow$
Demand	$\uparrow$	$\downarrow$	$-$	$-$

Table 4.6: The effects of μ_H, σ_H, σ_L, and ρ on the supply and demand of the hypothetical mortality bond.

These four factors can be estimated from mortality paths simulated from a two-population mortality model. For each simulated mortality path, we calculate the value v_H and v_L. This gives a joint empirical distribution of v_H and v_L, from which we can readily obtain estimates of μ_H, σ_H, σ_L and ρ.

In the rest of this section, we study how the introduction of transitory jumps may affect these four factors and consequently the equilibrium price of the mortality bond.

4.5.4 The Supply and Demand Curves when $a = 1.3$ and $b = 1.4$

The demand and supply curves when $a = 1.3$ and $b = 1.4$ are displayed in Figure 4.4. When transitory mortality jumps are introduced, the supply curve shifts upwards, while the demand curve shifts downwards. This leads to a reduction in the equilibrium price of the mortality bond.

To understand the changes in the demand and supply curves, we examine the four factors we just described. We first estimate them using the no-jump model and the nonconcurrent-jump model. The estimates are presented in Table 4.7.

The quantity μ_H represents the accumulated value of the expected payouts from the mortality bond, which is subject to a principal reduction when the mortality index exceeds the attachment point. The value of μ_H is therefore heavily dependent on the right tail of the distribution of the mortality index. In Figure 4.5 we plot the kernel smoothed density functions for I_{2007}/I_{2006} when different models are assumed.

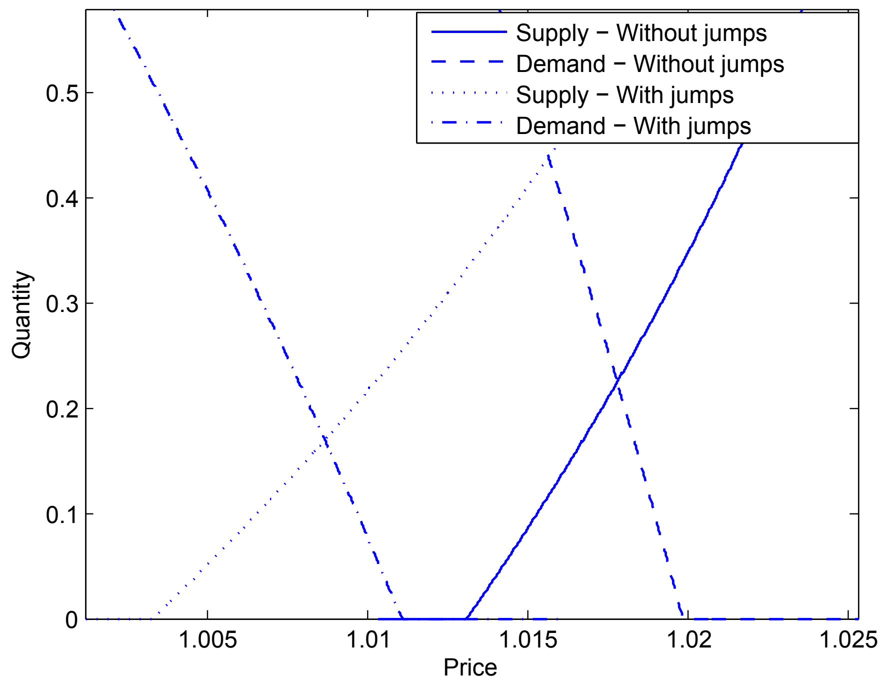


Figure 4.4: Supply and demand curves based on different mortality models, $a = 1.3$ and $b = 1.4$.

The upper panel of Figure 4.5 depicts the entire density functions, while the lower panel gives a closer shot of their right tails.

Recall that the principal repayment of the mortality bond is reduced if the mortality index is higher than $1.3I_{2006}$, and exhausted if it is higher than $1.4I_{2006}$. We observe from Figure 4.5 that beyond the attachment point $a = 1.3$, the value of the density function for I_{2007}/I_{2006} under the no-jump model is always lower than that under the nonconcurrent-jump model. This means that the no-jump model implies a smaller probability (and magnitude) of principal reduction and hence a higher value of μ_H.

The quantity σ_L depends heavily on the volatility of L_t, the insurance liability due in year t, for $t = 2007, 2008, 2009$. In Figure 4.6 we show the kernel smoothed density functions for L_{2007}, based on the no-jump model and the nonconcurrent jump model.

Model	μ_H	σ_H	σ_L	ρ
No-jump	1.0198	0.1258	0.1102	-0.3829
Nonconcurrent-jump	1.0110	0.1573	0.0677	-0.5832

Table 4.7: Estimates of μ_H, σ_H, σ_L and ρ when $a = 1.3$ and $b = 1.4$.

The no-jump model produces a significantly more dispersed (but less heavy-tailed) distribution, because it incorporates the variations caused by the jump process into its volatility term. As a result, the no-jump model gives a higher value σ_L.

The quantity σ_H depends on the volatility of the principal repayment. Because the principal repayment is always 100% unless the index exceeds the attachment point, the value of σ_H depends more heavily on the index's variability in the (right) tail than the overall volatility of the index. The dependence of σ_H on the tail volatility increases with the value of the attachment point. In this example, σ_H is higher when jumps are considered.

The quantity $|\rho|$ is related to the correlation between the mortality rates for the two populations. In the no-jump model, the correlation between the mortality of the two populations is driven entirely by the joint distribution of the innovations, $Z_\kappa(t)$ and $Z_{\Delta_\kappa}(t)$, whereas in the concurrent-jump model, the correlation is driven additionally by the jump count process and the joint distribution of the jump severities, $Y_t^{(1)}$ and $Y_t^{(2)}$. Hence, the introduction of jumps alters the correlation structure, and in this example, it increases $|\rho|$.

Now let us revisit the demand and supply curves. The introduction of transitory jumps increases σ_H and $|\rho|$ but reduces μ_H and σ_L. According to Table 4.6, a lower μ_H and a higher σ_H both exert pressure on the demand, causing the demand curve to shift downwards. On the other hand, the changes in μ_H and $|\rho|$ would lead to an increase in supply, while the changes in σ_H and σ_L would lead to a reduction. Here, the combined effect of μ_H and $|\rho|$ is stronger, causing the supply curve to shift upwards. Overall, the shifts in the demand and supply curves result in a smaller price and a lower quantity traded in equilibrium.

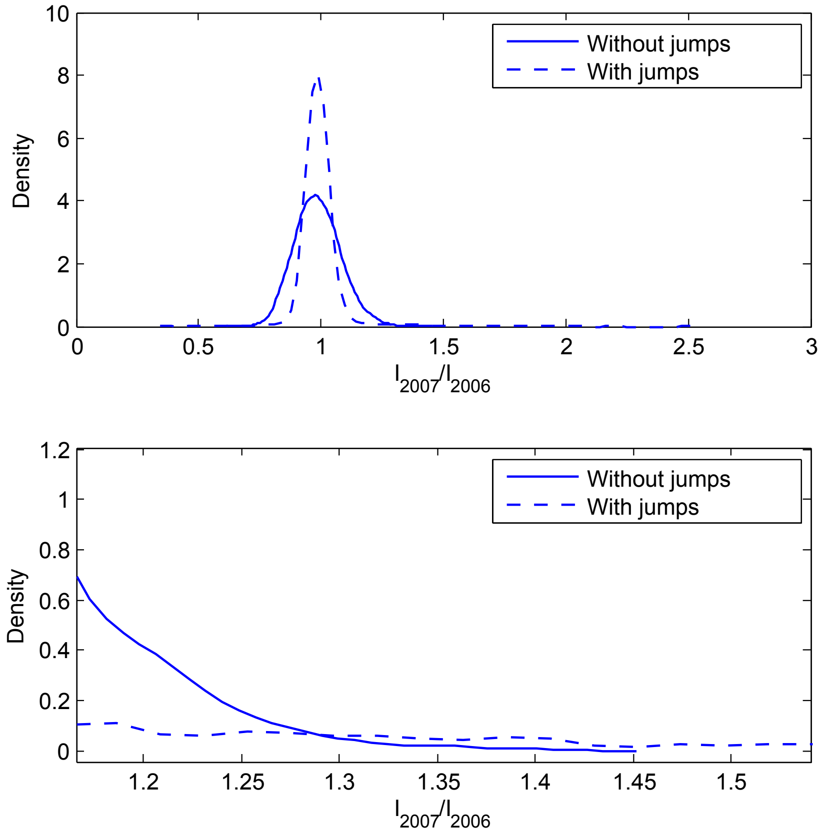


Figure 4.5: Kernel smoothed density functions for I_{2007}/I_{2006} under the no-jump model and the nonconcurrent-jump model.

4.5.5 The Supply and Demand Curves when $a = 1.2$ and $b = 1.3$

Here we repeat the analysis for the case when $a = 1.2$ and $b = 1.3$. The resulting supply and demand curves are shown in Figure 4.7. When mortality jumps are introduced, the supply curve shifts downwards, while the demand curve shifts upwards. This leads to an increase in the equilibrium price of the mortality bond. It is noteworthy that the changes in the demand and supply curves are exactly opposite to those in the previous example, which is based on higher attachment and exhaustion

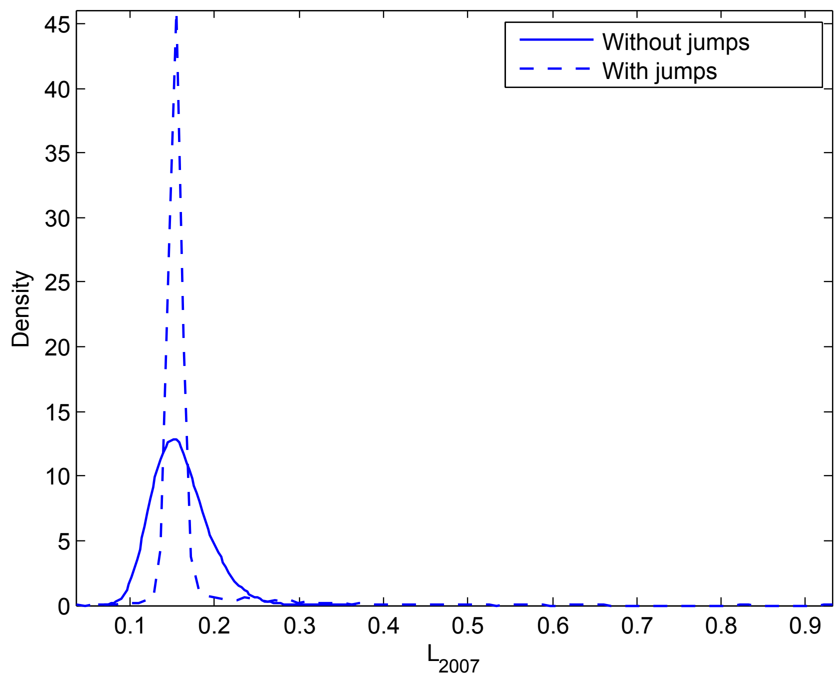


Figure 4.6: Kernel smoothed density functions for L_{2007} under the no-jump model and the nonconcurrent-jump model.

points.

We examine again the four determinants of supply and demand. The estimates of the four factors are shown in Table 4.8. Note that there is no change to the estimates of σ_L, because the insurance liability is not affected by the changes in the attachment and exhaustion points. Also, as in the previous example, the nonconcurrent-jump model implies a higher value of $|\rho|$.

However, in this example, the value μ_H is lower when mortality jumps are not taken into account. The reason behind can be seen from Figure 4.5, which shows the density functions of I_{2007}/I_{2006} under the two mortality models. We observe from the diagram that in most of the interval between the attachment point $a = 1.2$ and the exhaustion point $b = 1.3$, the value of density function based on the no-jump model is higher. This means that the no-jump model implies a higher probability

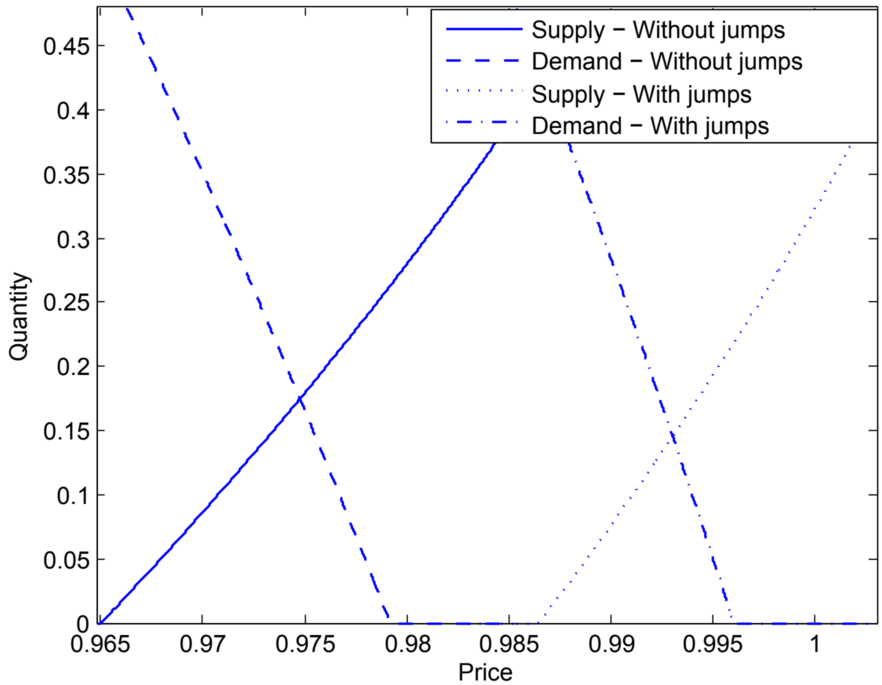


Figure 4.7: Supply and demand curves based on different mortality models, $a = 1.2$ and $b = 1.3$.

(and magnitude) of principal reduction and thus a smaller value of μ_H.

When the attachment point is lower, the value of σ_H is less dependent on the index's variability in the (right) tail but more dependent on the overall volatility of the index. The overall volatility of the index is higher when jumps are not taken into account, because, as we mentioned earlier, the no-jump model incorporates the variations caused by the jump process into its volatility term. As a result, in this example, the value of σ_H is lower when jumps are considered.

Now let us revisit the supply and demand curves in Figure 4.7. The introduction of transitory jumps increases μ_H and $|\rho|$ but reduces σ_H and σ_L. According to Table 4.6, a higher μ_H and a lower σ_H both have a positive effect on the demand, causing the demand curve to shift upwards. On the other hand, the changes in σ_H and $|\rho|$ would lead to an increase in supply, while the changes in μ_H and σ_L would lead to

Model	μ_H	σ_H	σ_L	ρ
No-jump	0.9792	0.2112	0.1102	−0.4980
Nonconcurrent-jump	0.9960	0.1906	0.0677	−0.6034

Table 4.8: Estimates of μ_H, σ_H, σ_L and ρ when $a = 1.2$ and $b = 1.3$.

a reduction. Here, the combined effect of μ_H and σ_L is stronger, causing the supply curve to shift downwards. Overall, the shifts in the demand and supply curves result in a higher price and a lower quantity traded in equilibrium.

4.6 Concluding Remarks

Standardization is a goal of many participants in the market for mortality-related risk. Standardized mortality-linked securities are based on broad-based populations, which are generally different from populations associated with hedgers' portfolios. Therefore, when we evaluate a hedge that is constructed with a standardized instrument, a two-population mortality model is necessary.

Nevertheless, existing two-population models do not incorporate mortality jumps, which could have a significant impact on the securitization of mortality-related risk. In this paper, we generalized a simple two-population model to incorporate transitory jump effects. We proposed two new models, one of which is more parsimonious, applicable to two closely related populations, while the other of which is less restrictive, suitable for modeling populations that are not so closely related to each other.

To study the impact of introducing transitory jumps, we considered a trade of a hypothetical mortality bond between a life insurer and an investor who is willing to take mortality risk for earning a risk premium. The principal repayment of the bond is reduced if the index to which it is linked exceeds a pre-determined attachment point. When jumps are introduced, the distribution of the index becomes less dispersed but heavier-tailed. Consequently, the impact on the estimated bond price depends heavily on the value of the attachment point. If the attachment point is low, then the model without jumps would imply a greater probability of principal reduction and hence a smaller bond price, but if the attachment point is high, the opposite is true.

Chapter 5

Modeling Trades as Nash Bargaining Problems

5.1 Introduction

In Chapter 2 and Chapter 3, we proposed a pricing approach that is applicable to a competitive market. We modeled the pricing problem as a Walrasian tâtonnement process and computed the competitive equilibrium. This approach is based on the most fundamental economic concept: demand and supply. It 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.

However, this approach requires the assumption of a competitive market. A perfectly competitive market exists when every participant is a "price taker", and no participant can influence the price of the product it buys or sells. The competitiveness assumption may not always be satisfied in today's market. For example, the deals in today's longevity market are often customized. The price of the trades are negotiated by the participants. The competitive equilibrium obtained may not be accurate in this case. Nevertheless, we can at least use it as a benchmark in situations when standard no-arbitrage methods are difficult to implement. We may also use it to predict what would happen when the market becomes more mature in future.

In this chapter, we relax the assumption of market competitiveness by modeling the trade of mortality-linked securities as a two-player bargaining game and apply Nash's bargaining solution to obtain a unique pair of trading price and quantity. A bargaining game involves a limited number of players. Each player may influence the pricing results through the bargaining process. The assumption of a competitive market is not required, and this alternative approach may better resemble the current market for mortality-linked securities. Furthermore, the bargaining game preserves the advantages of the tâtonnement approach. It gives a unique solution and does not take the market prices of other products as given. A two-player bargaining game is a starting point for modeling trades of mortality-linked securities with game theoretic methods.

Nash's bargaining solution suppresses many details of the decision making process and explains outcomes by identifying conditions that any outcome arrived at by rational decision makers should satisfy. These conditions are treated as axioms, from which the outcome is deduced using set-theoretical arguments. Nash's bargaining solution has been applied in insurance problems by several researchers. Borch (1974) applied it to reciprocal reinsurance treaties and determined the quota ceded by each player. Kihlstrom and Roth (1982) studied the insurance contracts reached through Nash bargaining and investigated the effect of the insureds' risk aversion on the outcome of bargaining about the terms of an insurance contract.

Boonen et al. (2011) developed a bargaining solution for over-the-counter risk redistribution and applied it to the hedging of longevity risk. They first examined the Pareto optimality of risk redistribution, and then obtained a unique risk redistribution in a Nash bargaining game. They found that proportional risk redistribution is Pareto optimal, when either exponential, power or quadratic utility functions is assumed and players have homogeneous belief about the probability distribution of risk. Proportional risk redistribution may be attained when a fully customized hedge is used. However, it may not be possible to attain proportional risk redistribution by using other hedging instruments. The results of Boonen et al. (2011) are no longer applicable, for example, when the players consider standardized instruments, such as q-forwards.

In this chapter, we model a trade between two counterparties. Counterparty A is a life insurer or pension plan sponsor who wants to hedge its mortality or longevity risk, and Counterparty B is an investor who invests in mortality-linked securities to earn a risk premium. In our set-up, risk is redistributed through trading a pre-specified mortality-linked security, permitting non-proportional risk redistributions.

We illustrate the application of Nash's bargaining solution through pricing a hypothetical mortality bond. We also compare the pricing results with those from a competitive market. Interestingly, we find that the trading quantities obtained in the two approaches are equal. This is because Nash's bargaining solution and the competitive equilibrium are both Pareto optimal under our settings, and all the Pareto optimal outcomes under our settings have the same trading quantity.

The remainder of this chapter is organized as follows. In Section 5.2, we describe the trade under consideration. In Section 5.3, we introduce Nash's bargaining solution, and explain how it can be applied to the pricing of mortality-linked securities. In Section 5.4, we revisit the pricing of mortality-linked securities in a competitive market. In Section 5.5, we structure a hypothetical mortality bond, and price this bond with the methods we introduced. In Section 5.6, we investigate the Pareto optimality of a trading contract. In Section 5.7, we conclude.

5.2 Setting up the Trade

Here, we set up the same trade used in Chapter 4. Recall that there are two counterparties involved in the trade. Counterparty A has life contingent liabilities that are due at times 1, 2, ..., T. The amount due at time t is $f_t(Q_t^L)$. To mitigate its exposure to mortality or longevity risk, it sells (or purchases) a mortality-linked security maturing at time T. At time t, the security makes a payout of $g_t(Q_t^H)$. Counterparty B is an investor who trades the mortality-linked security with Player A, in order to earn a risk premium. At time t, Counterparty B receives (or pays) an amount of $g_t(Q_t^H)$ per unit of the mortality-linked security purchased (or sold).

We let P and θ be the price and quantity that they agree for trading. θ is positive when Counterparty A sells the security, and negative when Counterparty A purchases the security. We call (P, θ) a trading contract.

Let W_t^A and W_t^B be the time-t wealth of Counterparties A and B, respectively. It is easy to see that

$$\begin{aligned} W_T^A(P,\theta) &= W_0^A e^{rT} - \theta e^{rT}(\sum_{t=1}^{T} g_t(Q_t)e^{-rt} - P) - \sum_{t=1}^{T} f_t(Q_t)e^{r(T-t)}, \\ W_T^B(P,\theta) &= W_0^B e^{rT} + \theta e^{rT}(\sum_{t=1}^{T} g_t(Q_t)e^{-rt} - P), \end{aligned}$$

where $W_T^A(P,\theta)$ and $W_T^B(P,\theta)$ indicate that W_T^A and W_T^B are functions of (P,θ).

In the following, we let

$$G = \sum_{t=1}^{T} g_t(Q_t)e^{r(T-t)}, \quad \text{and} \quad F = \sum_{t=1}^{T} f_t(Q_t)e^{r(T-t)},$$

for brevity.

Again, we denote the utility functions for Counterparties A and B by U^A and U^B, respectively. We assume an exponential utility function, $U(x) = 1 - e^{-kx}$, for each player. The two counterparties are assumed to be expected terminal utility maximizers. In our numerical examples, the assumed value of k for Counterparty A and Counterparty B are $k^A = 1$ and $k^B = 0.5$, respectively.

5.3 Modeling the Trade in a Non-Competitive Market

5.3.1 Bargaining Problems

In a non-competitive securities market, participants are not price takers. Each participant has an influence on the price of the security being traded, possibly through bargaining. In this section, we model the trade of a mortality-linked security between two counterparties as a Nash bargaining problem (Nash, 1950). Mathematically, a Nash bargaining problem is a pair $\langle S, d \rangle$, where $S \subset \Re^2$ is a compact and convex set, $d \in S$, and for some $s = (s_1, s_2)$ in S, $s_i > d_i$ for $i = 1, 2$.

The set S is the set of all feasible expected utility payoffs to the counterparties, while the vector $d = (d_1, d_2)$ represents the disagreement payoff; that is, if the counterparties do not come to an agreement, then they will receive utility payoffs of d_1 and d_2, respectively. We require $d \in S$, so that the counterparties can agree to disagree. We also require there exists $s = (s_1, s_2)$ in S such that $s_i > d_i$ for $i = 1, 2$, so that the counterparties have an incentive to reach an agreement (via bargaining).

What we are interested in, of course, is which point in S the bargaining process will lead to. Nash (1950) modeled the bargaining process by a function $\zeta : \mathcal{B} \to \Re^2$, where $\mathcal{B}$ is the set of all bargaining problems. The function ζ, which assigns a unique element in S to each bargaining problem $\langle S, d \rangle \in \mathcal{B}$, is referred to as a bargaining solution.

Rather than explicitly modeling the underlying bargaining procedure, Nash (1951) used a purely axiomatic approach to derive a bargaining solution. In more detail, he specified, as axioms, the following four properties that it would seem natural for a bargaining solution to have.

1. Pareto optimality

   If $\zeta(S, d) = (z_1, z_2)$ and $y_i \geq z_i$ for $i = 1, 2$, then either $y_i = z_i$ for $i = 1, 2$ or $(y_1, y_2) \notin S$.

2. Symmetry

   If (S, d) is a symmetric bargaining problem (i.e., $d_1 = d_2$ and $(x_1, x_2) \in S \Rightarrow (x_2, x_1) \in S$), then $\zeta_1(S, d) = \zeta_2(S, d)$, where $f_i(S, d)$ denotes the ith entry in $\zeta(S, d)$.

3. Independence of irrelevant alternatives

   If (S, d) and (T, d) are bargaining problems such that $S \subset T$ and $\zeta(T, d) \in S$, then $\zeta(S, d) = \zeta(T, d)$.

4. Independence of equivalent utility representatives

   If (S', d') is related to (S, d) in such a way that $d'_i = a_i d_i + b_i$ and $s'_i = a_i s_i + b_i$ for $i = 1, 2$, where a_i and b_i are real numbers and $a_i > 0$, then $\zeta_i(S', d') = a_i \zeta_i(S, d) + b_i$ for $i = 1, 2$.

Properties 1 and 3 are particularly related to bargaining. Specifically, Property 1 implies that if the counterparties agreed on an inferior outcome, then will renegotiate until the Pareto optimal outcome is reached. Property 3, on the other hand, resembles a gradual elimination of unacceptable outcomes. Eliminated outcomes (i.e., those in T but not in S) have no effect on the bargaining solution. We refer readers to Osborne and Rubinstein (1990) for a fuller discussion of these properties.

Nash showed that there is one and only one bargaining solution that satisfies the four axiomatic properties.

Theorem 1. *There is a unique solution which possesses Properties 1-4. The solution, $\zeta^N(S, d) : \mathcal{B} \to \Re^2$, takes the form*

$$\zeta^N(S, d) = \arg\max(s_1 - d_1)(s_2 - d_2),$$

where the maximization is taken over $(s_1, s_2) \in S$, and is subject to the constraint $s_i > d_i$ for $i = 1, 2$.

In other words, Nash's unique bargaining solution selects the utility pair in S that maximizes the product of the counterparties' gain in utility over the disagreement outcome (d_1, d_2). We call $(s_1 - d_1)(s_2 - d_2)$ the Nash product.

5.3.2 The Underlying Bargaining Strategy

Although the derivation of Nash's bargaining solution does not require knowledge on the details of underlying bargaining strategy, one may still wonder how the actual bargaining takes place. A possible underlying strategy is the one proposed by Zeuthen (1930), which we now describe.

At the kth round of the bargaining process, Player 1 proposes an agreement with a payoff vector $y^{1,k} = (y_1^{1,k}, y_2^{1,k})$, while Player 2 proposes another agreement with a payoff vector $y^{2,k} = (y_1^{2,k}, y_2^{2,k})$. If they fail to agree, then they receive disagreement payoff $d = (d_1, d_2)$.

Assume that $d_i < y_i^{j,k} < y_i^{i,k}$, where $i = 1, 2$, $j = 1, 2$ and $i \neq j$. At round $k + 1$, player i can take one of the following actions:

- accept Player j's offer, leading to an agreement;
- make a concession by counter-proposing $y^{i,k+1}$ such that $y_j^{i,k} < y_j^{i,k+1} < y_j^{j,k}$;
- repeat his last offer.

Let

$$p^{i,k} = \frac{y_i^{i,k} - y_i^{j,k}}{y_i^{i,k} - d_i},$$

where $i = 1, 2$, $j = 1, 2$, and $i \neq j$. Zeuthen's bargaining strategy can be summarized as follows:

- if $p^{2,k} < p^{1,k}$, then Player 2 makes a concession;
- if $p^{1,k} < p^{2,k}$, then Player 1 makes a concession;
- if $p^{1,k} = p^{2,k}$, then both players make a concession.

This process above will continue until the two players agree on a solution. Harsanyi (1956) proved that Zeuthen's bargaining strategy leads to the Nash's bargaining solution.

5.3.3 Applying Nash's Bargaining Solution to Mortality-linked Securities

Now, let us turn our focus back to mortality-linked securities. Following the set-up in Section 5.2, the expected utility payoffs to the counterparties are their expected terminal utilities. The utility possible set S is the set of feasible expected utility pairs

$$\left(\mathbb{E}\left[U^A(W_T^A(P,\theta))\right], \mathbb{E}\left[U^B(W_T^B(P,\theta))\right]\right)$$

arising from all possible values of P and θ.[1] Note that the structure of the mortality-linked security is assumed to be fixed. The players are only allowed to bargain over the price and quantity. The disagreement utility payoffs are the expected terminal utilities when there is no trade (i.e., $\theta = 0$). It follows that

$$d = \left(\mathbb{E}\left[U^A(W_T^A(0,0))\right], \mathbb{E}\left[U^B(W_T^B(0,0))\right]\right).$$

It is obvious that $d \in S$. For now, we assume that there exists $s = (s_1, s_2)$ in S such that $s_i > d_i$ for $i = 1, 2$. In Section 5.6, we will discuss under what conditions this assumption will hold.

We can then use Nash's bargaining solution to find the price P and the trading quantity θ upon agreement between the two counterparties:

$$\begin{aligned}
\underset{(P,\theta)}{\text{argmax}} \quad & \left(\mathbb{E}\left[U^A(W_T^A(P,\theta))\right] - \mathbb{E}\left[U^A(W_T^A(0,0))\right]\right) \\
& \times \left(\mathbb{E}\left[U^B(W_T^B(P,\theta))\right] - \mathbb{E}\left[U^B(W_T^B(0,0))\right]\right) \\
\\
\text{subject to} \quad & \mathbb{E}\left[U^A(W_T^A(P,\theta))\right] - \mathbb{E}\left[U^A(W_T^A(0,0))\right] \geq 0 \\
& \mathbb{E}\left[U^B(W_T^B(P,\theta))\right] - \mathbb{E}\left[U^B(W_T^B(0,0))\right] \geq 0 \\
& \theta \geq 0 \text{ (or } \theta \leq 0, \text{ depending on which counterparty is the seller)} \\
& P > 0
\end{aligned}$$

[1] We assume exponential utility functions, which are concave. The use of concave utility functions implies that S is a convex set (see, e.g. Kihlstrom and Roth, 1982; Boonen et al., 2011).

The above may be viewed as a non-linear constrained optimization problem. The solution, which we denote by $(\hat{P}, \hat{\theta})$, can be solved numerically by technical software such as MATLAB. The trading of $\hat{\theta}$ units of the security at a price of $\hat{P}$ makes both counterparties better off relative to the situation when there is no trade, and maximizes the product of the counterparties' expected utility gains.

5.4 Modeling the Trade in a Competitive Market

To see the impact of bargaining on the price and trading quantity, we need to know how the trade would happen if the security market is perfectly competitive. In a competitive market, all participants are price takers. Given a price, each participant decide their supply or demand of the security on the basis of a certain criterion.

Again, we follow the set-up specified in Section 5.2. We assume that the counterparties will choose a supply or demand of the security that will maximize their expected terminal utilities. Let $\hat{\theta}^A$ and $\hat{\theta}^B$ respectively be the quantities that Counterparty A and Counterparty B are willing to trade at a given price P.[2] At equilibrium when the market clears, we have $\hat{\theta}^A + \hat{\theta}^B = 0$. In previous chapters, we postulate this trade as a Walrasian auction, and numerically obtain the equilibrium price and trading quantity by gradually adjusting the price until the excess demand $|\hat{\theta}^A + \hat{\theta}^B|$ becomes zero.

In what follows, we push the results in previous chapters further by deriving a analytic relation between the price and trading quantity at equilibrium. We first formulate the objectives of the counterparties mathematically as follows:

$$\begin{aligned} \text{Counterparty A:} \quad \hat{\theta}^A &= \operatorname*{argsup}_{\theta^A} \mathbb{E}\left[U^A(W_T^A(P, -\theta^A))\right] \\ \text{Counterparty B:} \quad \hat{\theta}^B &= \operatorname*{argsup}_{\theta^B} \mathbb{E}\left[U^B(W_T^B(P, \theta^B))\right]. \end{aligned}$$

Proposition 5. *Suppose that Counterparties A and B have exponential utility functions with absolute risk aversion parameters k^A and k^B, respectively. The competitive*

[2]The quantity for the buyer is positive, while that for the seller is negative.

equilibrium (P^*, θ^*) *should satisfy*

$$P^* = \frac{\mathbb{E}[e^{k^A\theta^* G + k^A F} G]}{e^{rT}\mathbb{E}[e^{k^A\theta^* G + k^A F}]} = \frac{\mathbb{E}[e^{-k^B\theta^* G} G]}{e^{rT}\mathbb{E}[e^{-k^B\theta^* G}]}. \tag{5.1}$$

Proof. Recall that the terminal wealth of Counterparty A is given by $W_T^A(P, -\theta^A) = \theta^A(G - e^{rT}P) - F$. Given a price P, Counterparty A will trade $\hat{\theta}^A$ units of the security so that $\mathbb{E}\left[U^A(W_T^A(P, \hat{\theta}^A)\right]$ is the maximum. This means

$$\frac{\partial}{\partial \theta^A_{|\theta^A = \hat{\theta}^A}} \mathbb{E}\left[U^A(W_T^A(P, -\theta^A))\right] = 0, \quad \frac{\partial^2}{\partial {\theta^A}^2_{|\theta^A = \hat{\theta}^A}} \mathbb{E}\left[U^A(W_T^A(P, -\theta^A))\right] < 0.$$

Assuming that Counterparty A has an exponential utility function with risk aversion parameter k^A, the first order condition can be written as

$$\mathbb{E}\left[k^A(G - e^{rT}P)e^{-k^A\hat{\theta}^A(G - e^{rT}P) + k^A F}\right] = 0,$$

which implies

$$P = \frac{\mathbb{E}[e^{-k^A\hat{\theta}^A G + k^A F} G]}{e^{rT}\mathbb{E}[e^{-k^A\hat{\theta}^A G + k^A F}]}.$$

The second order condition is easy to verify.

On the other hand, the terminal wealth of Counterparty B can be written as $W_T^B(P, \theta^B) = \theta^B(G - e^{rT}P)$. Given a price P, Counterparty B will trade $\hat{\theta}^B$ units of the security such that $\mathbb{E}\left[U^B(W_T^B(P, \hat{\theta}^B))\right]$ is the maximum. This means

$$\frac{\partial}{\partial \theta^B_{|\theta^B = \hat{\theta}^B}} \mathbb{E}\left[U^B(W_T^B(P, \theta^B))\right] = 0, \quad \frac{\partial^2}{\partial {\theta^B}^2_{|\theta^B = \hat{\theta}^B}} \mathbb{E}\left[U^B(W_T^B(P, \theta^B))\right] < 0.$$

Assuming that Counterparty B has an exponential utility function with risk aversion parameter k^B, the first order condition can be written as

$$\mathbb{E}\left[k^B(G - e^{rT}P)e^{-k^B\hat{\theta}^B(G - e^{rT}P)}\right] = 0,$$

which implies

$$P = \frac{\mathbb{E}[e^{-k^B\hat{\theta}^B G} G]}{e^{rT}\mathbb{E}[e^{-k^B\hat{\theta}^B G}]}.$$

The second order condition for counterparty B is also easy to verify.

At equilibrium, $\hat{\theta}^A + \hat{\theta}^B = 0$. Letting $\hat{\theta}^B = \theta^*$ and $\hat{\theta}^A = -\theta^*$, the result follows. $\square$

To find the competitive equilibrium, we first solve the second part of equation (5.1) for θ^* numerically, and then substitute θ^* back to equation (5.1) to obtain P^*. Note that equation (5.1) may not have a solution, which happens where there is no trade between the two counterparties.

5.5 Numerical Illustrations

5.5.1 Mortality Data and Model

We illustrate the pricing methods with historical mortality data from Swedish male and Finnish male populations. We consider the same data used in Chapter 4. The mortality rates of the two populations are modeled by the two-population mortality model with transitory jump effects proposed in Chapter 4. Such a two-population model incorporates mortality jumps is ideal for pricing short-term catastrophe bond in the presence of population basis risk. The parameter estimates and forecasting procedure can also be found in Chapter 4.

5.5.2 Pricing a Hypothetic Mortality Bond

We now price the hypothetic mortality bond considered in Chapter 4. We assume that the attachment and exhaustion points of this mortality bond are 1.3 and 1.4, respectively.

Table 5.1 summarizes the outcomes under the two different models for the trade. The strictly positive utility gains indicate that both counterparties will benefit from the trade of the hypothetical security, no matter if the market is competitive or not. The benefit to Counterparty A (the hedger) is higher if the market is competitive, whereas the opposite is true for the other counterparty.

Method	Competitive Equilibrium	Nash's Bargaining Solution
Trading Price	1.008669	1.007971
Trading Quantity	0.169848	0.169848
Utility Gain for A	7.965399×10^{-4}	5.832860×10^{-4}
Utility Gain for B	1.124217×10^{-4}	1.772784×10^{-4}
Nash Product	8.954836×10^{-8}	1.034040×10^{-7}

Table 5.1: Trading prices, trading quantities and utility values under Nash's bargaining solution and the competitive equilibrium.

According to the first axiomatic property in Section 5.3.1, Nash's bargaining solution is Pareto optimal. It is also well known that, under idealized conditions, any outcome resulting from a competitive equilibrium must be Pareto optimal. An outcome is said to be Pareto optimal if there is no other outcome that makes every counterparty at least as well off and at least one counterparty strictly better off. In other words, a Pareto optimal outcome cannot be improved without hurting at least one counterparty. Therefore, the permission of bargaining can only improve the utility gain of one counterparty, but not both. In this example, Counterparty B benefits.

In this example, the trading price under Nash's bargaining solution is lower than that under the competitive equilibrium. This difference suggest that whether or not market participants are price takers does have an impact on the trading price. Given that the current market for mortality-linked security is not even close to competitive, practitioners should be cautious when they interpret prices estimated from pricing methods that require the assumption of market competitiveness.

By contrast, in this example, the trading quantities under Nash's bargaining solution and the competitive equilibrium are the same. This equality, as we are going to demonstrate in Section 5.6, is not a coincidence, but always true provided that certain conditions are satisfied.

5.6 Conditions for Pareto Optimality

In this section, we discuss the conditions under which the trade between the two counterparties will be Pareto optimal. As both Nash's bargaining solution and the competitive equilibrium are Pareto optimal, knowing the conditions for Pareto optimality may give us some insights about the trading prices and quantities under the two market models.

To obtain the condition for Pareto optimality, we first need to derive the conditions for the equation

$$\mathcal{H}(\theta) = \frac{\mathbb{E}[e^{k^A\theta G + k^A F}G]}{\mathbb{E}[e^{k^A\theta G + k^A F}]} - \frac{\mathbb{E}[e^{-k^B\theta G}G]}{\mathbb{E}[e^{-k^B\theta G}]} = 0. \tag{5.2}$$

to have a unique non-zero solution.

Proposition 6. *Equation (5.2) has a unique non-zero solution if and only if*

- $\mathbb{E}[e^{k^A F}G] - \mathbb{E}[e^{k^A F}]\mathbb{E}[G] < 0$ *when Counterparty A is the seller of the mortality-linked security;*
- $\mathbb{E}[e^{k^A F}G] - \mathbb{E}[e^{k^A F}]\mathbb{E}[G] > 0$ *when Counterparty A is the buyer of the mortality-linked security.*

Proof. We first show that $\mathcal{H}(\theta)$ is a strictly increasing function of θ. Differentiating the first term of $\mathcal{H}(\theta)$ with respect to θ, we have

$$\begin{aligned} & \frac{\partial}{\partial\theta}\frac{\mathbb{E}[e^{k^A\theta G + k^A F}G]}{\mathbb{E}[e^{k^A\theta G + k^A F}]} \\ = \quad & k^A \frac{\mathbb{E}[e^{k^A\theta G + k^A F}G^2]\,\mathbb{E}[e^{k^A\theta G + k^A F}] - \left(\mathbb{E}[e^{k^A\theta G + k^A F}G]\right)^2}{\left(\mathbb{E}[e^{k^A\theta G + k^A F}]\right)^2}. \end{aligned}$$

Using Hölder's inequality, we have

$$\mathbb{E}[e^{k^A\theta G + k^A F}G^2]\,\mathbb{E}[e^{k^A\theta G + k^A F}] - \left(\mathbb{E}[e^{k^A\theta G + k^A F}G]\right)^2 \geq 0.$$

The equality holds if and only if $e^{k^A\theta G+k^AF}G^2 = e^{k^A\theta G+k^AF}$ almost everywhere. This condition, which is equivalent to $G^2 = 1$ almost everywhere, is obviously not satisfied here. Therefore,

$$\frac{\partial}{\partial\theta}\left(\frac{\mathbb{E}[e^{k^A\theta G+k^AF}G]}{\mathbb{E}[e^{k^A\theta G+k^AF}]}\right) > 0.$$

Similarly, we can prove that

$$\frac{\partial}{\partial\theta}\left(-\frac{\mathbb{E}[e^{-k^B\theta G}G]}{\mathbb{E}[e^{-k^B\theta G}]}\right) > 0.$$

As a result, $\frac{\partial}{\partial\theta}\mathcal{H}(\theta) > 0$ for all θ. Since $\mathcal{H}(\theta)$ is a strictly increasing function of θ, the solution to equation (5.2) is unique if it exists.

Now, let us focus on the case that Counterparty A is the seller of the security. Since $\theta \in [0, +\infty)$, equation (5.2) has a unique non-zero solution if and only if

1. $\mathcal{H}(\theta) < 0$, when $\theta = 0$;
2. $\mathcal{H}(\theta) \geq 0$, when $\theta \to +\infty$.

When $\theta = 0$,

$$\mathcal{H}(0) = \frac{\mathbb{E}[e^{k^AF}G] - \mathbb{E}[e^{k^AF}]\mathbb{E}[G]}{\mathbb{E}[e^{k^AF}]},$$

which is negative if and only if $\mathbb{E}[e^{k^AF}G] - \mathbb{E}[e^{k^AF}]E[G] < 0$.

Condition 2 is satisfied if

$$\lim_{\theta\to+\infty}\frac{\mathbb{E}[e^{k^A\theta G+k^AF}G]}{\mathbb{E}[e^{k^A\theta G+k^AF}]} = \sup\{G\}, \tag{5.3}$$

and

$$\lim_{\theta\to+\infty}\frac{\mathbb{E}[e^{-k^B\theta G}G]}{\mathbb{E}[e^{-k^B\theta G}]} = \inf\{G\}. \tag{5.4}$$

For brevity, we let $M = \sup\{G\}$ and $N = \inf\{G\}$. Since $G \geq 0$, we have $0 \leq N < +\infty$.

We now prove equation (5.3). Suppose that $M < +\infty$. For any $\epsilon > 0$, fix $\delta < \frac{\epsilon}{2}$. We have

$$
\begin{aligned}
& \left| \frac{\mathbb{E}\left[e^{k^A\theta G + k^A F} G\right]}{\mathbb{E}\left[e^{k^A\theta G + k^A F}\right]} - M \right| \\
= \quad & \frac{\mathbb{E}\left[e^{k^A\theta G + k^A F} |G - M|\right]}{\mathbb{E}\left[e^{k^A\theta G + k^A F}\right]} \\
= \quad & \frac{\mathbb{E}\left[e^{k^A\theta G + k^A F} |G - M| \mathbb{I}_{|G-M| \leq \delta}\right]}{\mathbb{E}\left[e^{k^A\theta G + k^A F}\right]} + \frac{\mathbb{E}\left[e^{k^A\theta G + k^A F} |G - M| \mathbb{I}_{|G-M| > \delta}\right]}{\mathbb{E}\left[e^{k^A\theta G + k^A F}\right]} \\
\leq \quad & \delta + \frac{\mathbb{E}\left[e^{k^A\theta G + k^A F} |G - M| \mathbb{I}_{|G-M| > \delta}\right]}{\mathbb{E}\left[e^{k^A\theta G + k^A F}\right]} \\
= \quad & \delta + \frac{\mathbb{E}\left[e^{k^A\theta G + k^A F} |G - M| \mathbb{I}_{|G-M| > \delta}\right]}{\mathbb{E}\left[e^{k^A\theta G + k^A F} \mathbb{I}_{|G-M| \leq \frac{\delta}{2}}\right] + \mathbb{E}\left[e^{k^A\theta G + k^A F} \mathbb{I}_{|G-M| > \frac{\delta}{2}}\right]} \\
\leq \quad & \delta + \frac{\mathbb{E}\left[e^{k^A\theta G + k^A F} |G - M| \mathbb{I}_{|G-M| > \delta}\right]}{\mathbb{E}\left[e^{k^A\theta G + k^A F} \mathbb{I}_{|G-M| \leq \frac{\delta}{2}}\right]} \\
< \quad & \delta + \frac{\mathbb{E}\left[e^{k^A\theta (M-\delta) + k^A F} |G - M| \mathbb{I}_{|G-M| > \delta}\right]}{\mathbb{E}\left[e^{k^A\theta (M-\frac{\delta}{2}) + k^A F} \mathbb{I}_{|G-M| \leq \frac{\delta}{2}}\right]} \\
< \quad & \frac{\epsilon}{2} + \frac{e^{-k^A\theta\frac{\delta}{2}} \mathbb{E}\left[e^{k^A F} |G - M| \mathbb{I}_{|G-M| > \delta}\right]}{\mathbb{E}\left[e^{k^A F} \mathbb{I}_{|G-M| \leq \frac{\delta}{2}}\right]},
\end{aligned}
$$

where $\mathbb{I}_A$ is the indicator function for event A. When $\theta \geq -\frac{2}{k^A\delta} \ln \frac{\epsilon \mathbb{E}\left[e^{k^A F} \mathbb{I}_{|G-M| < \frac{\delta}{2}}\right]}{2\mathbb{E}\left[e^{k^A F} |G-M| \mathbb{I}_{|G-M| > \delta}\right]}$, we have

$$
\left| \frac{\mathbb{E}\left[e^{k^A\theta G + k^A F} G\right]}{\mathbb{E}\left[e^{k^A\theta G + k^A F}\right]} - M \right| < \epsilon
$$

for any $\epsilon > 0$. Therefore, equation (5.3) holds when $M < +\infty$.

Suppose that $M = +\infty$. For any $\delta > 0$, we have

$$
\begin{aligned}
& \mathbb{E}\left[e^{k^A\theta G + k^A F}(G-\delta)\right] \\
= \;& \mathbb{E}\left[e^{k^A\theta G + k^A F}(G-\delta)\mathbb{I}_{G\le\delta}\right] + \mathbb{E}\left[e^{k^A\theta G + k^A F}(G-\delta)\mathbb{I}_{\delta<G\le 2\delta}\right] + \mathbb{E}\left[e^{k^A\theta G + k^A F}(G-\delta)\mathbb{I}_{G>2\delta}\right] \\
> \;& \mathbb{E}\left[e^{k^A\theta G + k^A F}(G-\delta)\mathbb{I}_{G\le\delta}\right] + \mathbb{E}\left[e^{k^A\theta G + k^A F}(G-\delta)\mathbb{I}_{G>2\delta}\right] \\
> \;& -\delta\mathbb{E}\left[e^{k^A\theta\delta + k^A F}\mathbb{I}_{G\le\delta}\right] + \delta\mathbb{E}\left[e^{2k^A\theta\delta + k^A F}\mathbb{I}_{G>2\delta}\right] \\
= \;& \delta e^{k^A\theta\delta}\left(-\mathbb{E}\left[e^{k^A F}\mathbb{I}_{G\le\delta}\right] + e^{k^A\theta\delta}\mathbb{E}\left[e^{k^A F}\mathbb{I}_{G>2\delta}\right]\right).
\end{aligned}
$$

When $\theta \ge \frac{\ln\mathbb{E}\left[e^{k^A\theta F}\mathbb{I}_{G\le\delta}\right] - \ln\mathbb{E}\left[e^{k^A\theta F}\mathbb{I}_{G\ge 2\delta}\right]}{k^A\delta}$, we have $\mathbb{E}\left[e^{k^A\theta G + k^A F}(G-\delta)\right] > 0$ and hence

$$
\frac{\mathbb{E}\left[e^{k^A\theta G + k^A F}G\right]}{\mathbb{E}\left[e^{k^A\theta G + k^A F}\right]} - \delta > 0
$$

for any $\delta > 0$. Therefore, equation (5.3) also holds when $M = +\infty$. We conclude that equation (5.3) holds in general.

We then prove equation (5.4). For any $\epsilon > 0$, fix $\delta < \frac{\epsilon}{2}$. We have

$$\begin{aligned}
& \left| \frac{\mathbb{E}\left[e^{-k^B\theta G} G\right]}{\mathbb{E}\left[e^{-k^B\theta G}\right]} - N \right| \\
= \quad & \frac{\mathbb{E}\left[e^{-k^B\theta G}|G-N|\right]}{\mathbb{E}\left[e^{-k^B\theta G}\right]} \\
= \quad & \frac{\mathbb{E}\left[e^{-k^B\theta G}|G-N|\mathbb{I}_{|G-N|\le\delta}\right]}{\mathbb{E}\left[e^{-k^B\theta G}\right]} + \frac{\mathbb{E}\left[e^{-k^B\theta G}|G-N|\mathbb{I}_{|G-N|>\delta}\right]}{\mathbb{E}\left[e^{-k^B\theta G}\right]} \\
\le \quad & \delta + \frac{\mathbb{E}\left[e^{-k^B\theta G}|G-N|\mathbb{I}_{|G-N|>\delta}\right]}{\mathbb{E}\left[e^{-k^B\theta G}\right]} \\
= \quad & \delta + \frac{\mathbb{E}\left[e^{-k^B\theta G}|G-N|\mathbb{I}_{|G-N|>\delta}\right]}{\mathbb{E}\left[e^{-k^B\theta G}\mathbb{I}_{|G-N|\le\frac{\delta}{2}}\right] + \mathbb{E}\left[e^{-k^B\theta G}\mathbb{I}_{|G-N|>\frac{\delta}{2}}\right]} \\
\le \quad & \delta + \frac{\mathbb{E}\left[e^{-k^B\theta G}|G-N|\mathbb{I}_{|G-N|>\delta}\right]}{\mathbb{E}\left[e^{-k^B\theta G}\mathbb{I}_{|G-N|\le\frac{\delta}{2}}\right]} \\
< \quad & \delta + \frac{\mathbb{E}\left[e^{-k^B\theta (N+\delta)}|G-N|\mathbb{I}_{|G-N|>\delta}\right]}{\mathbb{E}\left[e^{-k^B\theta (N+\frac{\delta}{2})}\mathbb{I}_{|G-N|\le\frac{\delta}{2}}\right]} \\
< \quad & \frac{\epsilon}{2} + \frac{e^{-k^B\theta\frac{\delta}{2}}\mathbb{E}\left[|G-N|\mathbb{I}_{|G-N|>\delta}\right]}{\mathbb{E}\left[\mathbb{I}_{|G-N|\le\frac{\delta}{2}}\right]}.
\end{aligned}$$

When $\theta \ge -\frac{2}{k^B\delta}\ln\frac{\epsilon\mathbb{E}\left[\mathbb{I}_{|G-N|<\frac{\delta}{2}}\right]}{2\mathbb{E}\left[|G-N|\mathbb{I}_{|G-N|>\delta}\right]}$, we have

$$\left| \frac{\mathbb{E}\left[e^{-k^B\theta G} G\right]}{\mathbb{E}\left[e^{-k^B\theta G}\right]} - N \right| < \epsilon$$

for all $\epsilon > 0$. Therefore, equation (5.4) holds.

The proof for the case that Counterparty A is the buyer is similar.

□

Having proved proposition 6, we are now ready to derive the conditions for Pareto optimality.

Proposition 7. *Assume that Counterparties A and B have exponential utility functions with risk aversion parameters k^A and k^B, respectively.*

1. *Suppose that Counterparty A is the seller of the mortality-linked security.*

   When $\mathbb{E}[e^{k^A F} G] - \mathbb{E}[e^{k^A F}]\,\mathbb{E}[G] \geq 0$, the outcome $(\tilde{P}, \tilde{\theta})$ is Pareto optimal if and only if $\tilde{\theta} = 0$.

   When $\mathbb{E}[e^{k^A F} G] - \mathbb{E}[e^{k^A F}]\,\mathbb{E}[G] < 0$, the outcome $(\tilde{P}, \tilde{\theta})$ is Pareto optimal if and only if $\mathcal{H}(\tilde{\theta}) = 0$.

2. *Suppose that Counterparty A is the buyer of the mortality-linked security.*

   When $\mathbb{E}[e^{k^A F} G] - \mathbb{E}[e^{k^A F}]\,\mathbb{E}[G] \leq 0$, the outcome $(\tilde{P}, \tilde{\theta})$ is Pareto optimal if and only if $\tilde{\theta} = 0$.

   When $\mathbb{E}[e^{k^A F} G] - \mathbb{E}[e^{k^A F}]\,\mathbb{E}[G] > 0$, the outcome $(\tilde{P}, \tilde{\theta})$ is Pareto optimal if and only if $\mathcal{H}(\tilde{\theta}) = 0$.

Proof. In the trade of the security under consideration, an outcome $(\tilde{P}, \tilde{\theta})$ is Pareto optimal if there does not exist any pair of (P', θ') that satisfy the following conditions:

1. $\mathbb{E}[U^A(W_T^A(P', \theta'))] \geq \mathbb{E}[U^A(W_T^A(\tilde{P}, \tilde{\theta}))]$;
2. $\mathbb{E}[U^B(W_T^B(P', \theta'))] \geq \mathbb{E}[U^B(W_T^B(\tilde{P}, \tilde{\theta}))]$;
3. one of the above two inequalities is strict.

Condition (1) can be rewritten as follows:

$$\begin{aligned}
\mathbb{E}[U^A(W_T^A(P', \theta'))] &\geq \mathbb{E}[U^A(W_T^A(\tilde{P}, \tilde{\theta}))] \\
\mathbb{E}[e^{k^A \theta' G + k^A F}] e^{-k^A P' \theta' e^{rT}} &\leq \mathbb{E}[e^{k^A \tilde{\theta} G + k^A F}] e^{-k^A \tilde{P} \tilde{\theta} e^{rT}} \\
e^{k^A e^{rT} (\tilde{P}\tilde{\theta} - P'\theta')} &\leq \frac{\mathbb{E}[e^{k^A \tilde{\theta} G + k^A F}]}{\mathbb{E}[e^{k^A \theta' G + k^A F}]} \\
P'\theta' - \tilde{P}\tilde{\theta} &\geq \frac{\ln \mathbb{E}[e^{k^A \theta' G + k^A F}] - \ln \mathbb{E}[e^{k^A \tilde{\theta} G + k^A F}]}{k^A e^{rT}}.
\end{aligned}$$

Condition (2) can be rewritten as follows:

$$\begin{aligned}
\mathbb{E}[U^B(W_T^B(P',\theta'))] &\geq \mathbb{E}[U^B(W_T^B(\tilde{P},\tilde{\theta}))] \\
\mathbb{E}[e^{-k^B\theta' G}]e^{k^B P'\theta' e^{rT}} &\leq \mathbb{E}[e^{-k^B\tilde{\theta}G}]e^{k^B\tilde{P}\tilde{\theta}e^{rT}} \\
e^{k^B e^{rT}(P'\theta'-\tilde{P}\tilde{\theta})} &\leq \frac{\mathbb{E}[e^{-k^B\tilde{\theta}G}]}{\mathbb{E}[e^{-k^B\theta' G}]} \\
P'\theta'-\tilde{P}\tilde{\theta} &\leq \frac{\ln\mathbb{E}[e^{-k^B\tilde{\theta}G}]-\ln\mathbb{E}[e^{-k^B\theta' G}]}{k^B e^{rT}}.
\end{aligned}$$

There does not exist any (P',θ') meeting the three conditions above if and only if $(\tilde{P},\tilde{\theta})$ satisfies one of the following conditions:

i. $\frac{\ln\mathbb{E}[e^{k^A\theta' G+k^A F}]-\ln\mathbb{E}[e^{k^A\tilde{\theta}G+k^A F}]}{k^A e^{rT}} \geq \frac{\ln\mathbb{E}[e^{-k^B\tilde{\theta}G}]-\ln\mathbb{E}[e^{-k^B\theta' G}]}{k^B e^{rT}}$ for any (P',θ');

ii. $P'\theta'-\tilde{P}\tilde{\theta} < \frac{\ln\mathbb{E}[e^{k^A\theta' G+k^A F}]-\ln\mathbb{E}[e^{k^A\tilde{\theta}G+k^A F}]}{k^A e^{rT}}$ for any (P',θ');

iii. $P'\theta'-\tilde{P}\tilde{\theta} > \frac{\ln\mathbb{E}[e^{-k^B\tilde{\theta}G}]-\ln\mathbb{E}[e^{-k^B\theta' G}]}{k^B e^{rT}}$ for any (P',θ').

Condition (ii) cannot be satisfied, because we can always choose an arbitrary θ' and then pick a value of P' from the interval

$$\left[\frac{\ln\mathbb{E}[e^{k^A\theta' G+k^A F}]-\ln\mathbb{E}[e^{k^A\tilde{\theta}G+k^A F}]}{k^A e^{rT}\theta'}+\frac{\tilde{P}\tilde{\theta}}{\theta'},+\infty\right).$$

By using a similar argument, we can easily see that Condition (iii) cannot be satisfied, too. All then that remains is Condition (i).

Condition (i) is equivalent to

$$\left(\frac{\ln\mathbb{E}[e^{k^A\theta' G+k^A F}]}{k^A e^{rT}}+\frac{\ln\mathbb{E}[e^{-k^B\theta' G}]}{k^B e^{rT}}\right)-\left(\frac{\ln\mathbb{E}[e^{k^A\tilde{\theta}G+k^A F}]}{k^A e^{rT}}+\frac{\ln\mathbb{E}[e^{-k^B\tilde{\theta}G}]}{k^B e^{rT}}\right)\geq 0,$$

for all θ', which is satisfied if and only if the function

$$\mathcal{V}(\theta)=\frac{\ln\mathbb{E}[e^{k^A\theta G+k^A F}]}{k^A e^{rT}}+\frac{\ln\mathbb{E}[e^{-k^B\theta G}]}{k^B e^{rT}}$$

attains its minimum value at $\tilde{\theta}$.

It is easy to show that $\frac{\partial}{\partial\theta}\mathcal{V}(\theta) = \mathcal{H}(\theta)$. Also, in the proof for proposition 6, we showed that $\frac{\partial^2}{\partial\theta^2}\mathcal{V}(\theta) = \frac{\partial}{\partial\theta}\mathcal{H}(\theta) > 0$ for all θ.

We now focus on the case that Counterparty A is the seller of the mortality-linked security, that is, $\theta \in [0, +\infty)$.

When $\mathbb{E}[e^{k^A F} G] - \mathbb{E}[e^{k^A F}]\,\mathbb{E}[G] < 0$, according to proposition 6, $\frac{\partial}{\partial\theta}\mathcal{V}(\theta) = \mathcal{H}(\theta) = 0$ has a unique non-zero solution. It follows that $\mathcal{V}(\theta)$ is minimized at $\theta = \tilde{\theta}$ such that $\mathcal{H}(\tilde{\theta}) = 0$.

When $\mathbb{E}[e^{k^A F} G] - \mathbb{E}[e^{k^A F}]\,\mathbb{E}[G] \geq 0$, we have

$$\frac{\partial}{\partial\theta}_{|\theta=0}\mathcal{V}(\theta) = \mathcal{H}(0) = \frac{\mathbb{E}[e^{k^A F} G] - \mathbb{E}[e^{k^A F}]\mathbb{E}[G]}{e^{rT}\mathbb{E}[e^{k^A F}]} \geq 0.$$

Also, because $\frac{\partial}{\partial\theta}_{|\theta=0}\mathcal{V}(\theta) \geq 0$ and $\frac{\partial^2}{\partial\theta^2}\mathcal{V}(\theta) > 0$ for all θ, we have $\frac{\partial}{\partial\theta}\mathcal{V}(\theta) > 0$ for all $\theta \in (0, +\infty)$. It follows that $\mathcal{V}(\theta)$ is a strictly increasing function of θ for $\theta \in (0, +\infty)$, and that it attains minimum at $\theta = 0$.

This completes the proof for the case that Counterparty A is the seller. The proof for the remaining case is largely similar.

□

Proposition 7 states explicitly the conditions under which an outcome (P, θ) is Pareto optimal. Depending on the position of Counterparty A in the trade and the sign of the expression $\mathbb{E}[e^{k^A F} G] - \mathbb{E}[e^{k^A F}]\,\mathbb{E}[G]$, Pareto optimality is attained either when there is no trade or when $\tilde{\theta}$ units of the security is traded, where $\tilde{\theta}$ is the solution to the equation $\mathcal{H}(\theta) = 0$, which, according to proposition 6, has a unique non-zero solution.

Suppose that a trade between the two counterparties occurs. Then there is one and only one trading quantity that would lead to Pareto optimality. Because both Nash's bargaining solution and the competitive equilibrium are Pareto optimal, they must yield the same trading quantity. This fact explains why the estimated trading quantities in Table 5.1 are identical. The Pareto optimality conditions, however, do not depend on the trading price P. This means that even though the trading quantities in Nash's bargaining solution and the competitive equilibrium must be the same, the trading prices may not.

Furthermore, proposition 7 provides us with an alternative way to obtain Nash's bargaining solution for the trade under consideration. Specifically, we can first find $\tilde{\theta}$ that solves the equation $\mathcal{H}(\theta) = 0$. Because $\tilde{\theta}$ is the only trading quantity that leads to Pareto optimality, the trading quantity in Nash's bargaining solution must be $\tilde{\theta}$. We can then obtain the trading price in Nash's bargaining solution by maximizing the Nash product evaluated at $\theta = \tilde{\theta}$. This boils down to solving the first order condition,

$$\begin{aligned} k^A \mathbb{E}[e^{k^A \tilde{\theta}(G - e^{rT}P) + k^A F}] - k^B \mathbb{E}[e^{k^A F}] \mathbb{E}[e^{-k^B \tilde{\theta}(G - e^{rT}P)}] & \\ + (k^B - k^A) \mathbb{E}[e^{k^A \tilde{\theta}(G - e^{rT}P) + k^A F}] \mathbb{E}[e^{-k^B \tilde{\theta}(G - e^{rT}P)}] &= 0, \end{aligned}$$

for P.

Given the role of Counterparty A, whether or not a trade will occur depends entirely on the sign of the expression $\mathbb{E}[e^{k^A F} G] - \mathbb{E}[e^{k^A F}]\, \mathbb{E}[G]$, which is precisely $\mathrm{cov}(e^{k^A F}, G)$, the covariance between the random variables $e^{k^A F}$ and G, where F and G are the accumulated values of the cash flows arising from the liability being hedged and the security under consideration, respectively. This condition is highly intuitive. To illustrate, let us consider the situation when Counterparty A hedges its mortality risk exposure by buying a mortality-linked security from Counterparty B. In this situation, if $\mathrm{cov}(e^{k^A F}, G) \leq 0$, which implies $e^{k^A F}$ and G are not positively correlated with each other, then Counterparty A has no reason to purchase the security, as holding the security will increase but not reduce her risk exposure. This is completely in line with Proposition 7, which says that if Counterparty A is the buyer and $\mathrm{cov}(e^{k^A F}, G) \leq 0$, then no trade will occur.

Nash's bargaining solution requires the assumption that there exists $s = (s_1, s_2)$ in the utility possible set S such that $s_i > d_i$ for $i = 1, 2$. In Section 5.3.3, we made this assumption without specifying when it is satisfied and when it is not. Here, we demonstrate that the satisfaction of this assumption is dependent on the sign of $\mathrm{cov}(e^{k^A F}, G)$.

Proposition 8. *Assume that Counterparties A and B have exponential utility functions with risk aversion parameters k^A and k^B, respectively. A necessary and sufficient for satisfying the assumption that there exists $s = (s_1, s_2)$ in S such that $s_i > d_i$ for $i = 1, 2$ is*

1. *$cov(e^{k^A F}, G) < 0$, if Counterparty A is the seller of the mortality-linked security;*
2. *$cov(e^{k^A F}, G) > 0$, if Counterparty A is the buyer of the mortality-linked security.*

Proof. We focus on the case that Counterparty A is the seller of the mortality-linked security.

First, we prove the necessity. The existence of $s = (s_1, s_2)$ in S such that $s_i > d_i$ for $i = 1, 2$ implies that $d = (d_1, d_2)$ is not a Pareto optimal outcome, or equivalently speaking, $\theta = 0$ does not lead to Pareto optimality. According to proposition 7, we must have $\text{cov}(e^{k^A F}, G) < 0$. Therefore, $\text{cov}(e^{k^A F}, G) < 0$ is a necessary condition.

Next we prove the sufficiency. If $\text{cov}(e^{k^A F}, G) < 0$, then according to propositions 6 and 7, any outcome with $\theta = 0$ is not Pareto optimal. It follows that we can always find (P, θ) that leads to expected utility payoffs $(y_1, y_2) \in S$, where $y_1 \geq d_1$, $y_2 \geq d_2$ and at least one of the two equalities does not hold.

If $y_1 > d_1$ and $y_2 > d_2$, then obviously there exists $s \in S$ such that $s_i > d_i$ for $i = 1, 2$.

If $y_1 = d_1$ and $y_2 > d_2$, we can write $y_2 = d_2 + \epsilon$ for some $\epsilon > 0$. Given a fixed θ, y_1 is a continuous increasing function of P, while y_2 is a continuous decreasing function of P.[3] We can always find another price $\hat{P} > P$ that leads to expected utility payoffs $(\hat{y}_1, \hat{y}_2)$, where $d_2 < \hat{y}_2 < d_2 + \epsilon$ and that $\hat{y}_1 > d_1$. Therefore, there exists $s \in S$ such that $s_i > d_i$ for $i = 1, 2$. Using a similar argument, we can verify the existence of such a point in S when $y_1 > d_1$ and $y_2 = d_2$.

[3] It is assumed here that Counterparty A is the seller. Given a fixed θ, the expected utility payoff for Counterparty A is higher if she can sell the security for a higher price. On the other hand, the expected utility payoff for Counterparty B is lower if she has to purchase the security at a higher price.

This completes the proof for the case that Counterparty A is the seller. The proof for the other case is similar.

□

Proposition 8 has significant implications. It says that as long as the security under consideration is an effective hedging instrument (in the sense that $\mathrm{cov}(e^{k^A F}, G)$ takes the desired sign), the bargain between the two counterparties will always lead to a trade of the security. The trade will benefit both counterparties, as it will bring expected utility payoffs that are strictly greater than those when there is no trade.

5.7 Conclusion

In this chapter, we model the mortality-linked security trade as a two-player bargaining game and obtain the security price using Nash's bargaining solution. Nash bargaining assumes that the two players cooperate and maximize the product of their utility gains from a trade agreement. In the bargaining game, each player can exert influence on the trade.

The two-player bargaining game resembles most of the trades in current mortality-linked security market, for example, the buy-in deal in December 2010 between the Dutch food manufacturer Hero and the Dutch insurer Aegon, the longevity swap in February 2011 between J. P. Morgan and the Pall (UK) pension fund, and the first life book reinsurance swap since the Global Financial Crisis between Atlanticlux and institutional investors in June 2011. For such trades, the requirement of competitive market in the tâtonnement approach is not satisfied. Therefore, the Nash bargaining solution is more suitable for the current market than the tâtonnement approach.

Like the tâtonnement approach proposed in previous chapters, the application of Nash bargaining solution has several advantages over no arbitrage approaches. It does not require market price data, which are scarce in today's embryonic mortality-linked security market.

Our illustrations show that the permission of bargaining has an impact on the trading price, but has no effect on the trading quantity. This follows from the facts that both competitive equilibrium and Nash's bargaining solution are Pareto optimal. Pareto optimality is one of the four properties that defines Nash's bargaining solution. The first theorem in the welfare economics says that any competitive equilibrium or Walrasian equilibrium leads to a Pareto efficient allocation of resources. In this paper, we proved that, under the assumption of exponential utility functions, Pareto optimality can be attained only if a unique specific quantity of the security is traded. Therefore, the permission of bargaining does not affect the trading quantity.

Chapter 6

Concluding Remarks and Future Research

Due to the lack of market price data in today's embryonic market, it is difficult to implement any no arbitrage approach for pricing mortality-linked securities. In this thesis, we propose economic pricing approaches for use in both competitive and non-competitive markets. These pricing approaches do not require prices of other similar securities.

We model the trade between two counterparties, 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 both counterparties are expected terminal utility maximizer.

We first develop a pricing approach for use in a competitive market. This approach utilizes a Walrasian tâtonnement process to describe the equilibrium formulation. It is based on the most fundamental economic concept: demand and supply. We present two versions of the approach. The first version, which assumes no trade during the term of security, aims to illustrate the basic principles of the pricing method. The second version is less restrictive, allowing multiple trades by applying sequential decision process. Both versions are illustrated with a hypothetical mortality-linked security.

This approach is then extended to allow for population basis risk. Given the proposed extension, the time-0 price of a standardized mortality-linked security can be readily estimated. Moreover, the trading quantity can be viewed as the quantity of the security required in forming a static longevity hedge, provided that hedger's objective is to maximize his/her expected utility at some future time. Based on the extended pricing approach, we investigate the impact of population basis risk on the behaviors of hedgers and investors in the longevity risk market. We also examine how the hedging strategy would depend on the properties of the populations involved in the trade.

Besides population basis risk, mortality jumps are another important factor when extreme mortality risk is hedged using a standardized mortality-linked security. In order not to underestimate the probability of a catastrophic mortality deterioration, mortality jumps should be taken into account in a multi-population mortality model. We propose a two-population mortality model with transitory jump effects. We then examine the impact of introducing mortality jumps on the trading of mortality/longevity risk, based on the extended pricing approach in a competitive market.

Finally, we model the pricing problem as a bargaining game, which does not require the competitive market assumption, an assumption that is not satisfied in today's market. We apply a two-player Nash bargaining game to pricing mortality-linked securities and compare the results with those derived from the competitive equilibrium. We also examine the Pareto optimality of a trading contract, in order to gain some further insights.

Along the lines of this thesis, the following research ideas can be explored in future.

- *The impact of small sample risk on pricing mortality-linked securities.* In this thesis, we assume that there is no small sample risk (or sampling risk), that is, the risk that the realized mortality experience is different from the true mortality rate. Small sample risk is diversifiable, so it does not matter much for a large population, say one with more than 100,000 lives. However, for smaller populations, the risk may be significant. We may examine the impact of small sample risk on the trading of mortality-linked securities. To achieve this, we may incorporate small sample risk by treating the hedger's population as a random survivorship group, and by modeling the number of deaths in the population with a death process, possibly a Poisson or binomial.

- *A comparison of two-population mortality models.* Recently, Dowd et al. (2011) proposed a gravity model of mortality rates for two populations. We can perform serious validation work, similar to the recent contributions by Cairns et al. (2009) and Dowd et al. (2010a,b), for various two-population mortality models, including Dowd et al. (2011), Cairns et al. (2011a) and the model proposed in Chapter 4.

- *Developing a multi-population mortality model.* The two-population mortality models listed above cannot accommodate three or more populations. It would be interesting to develop an extension that can handle more than two populations simultaneously. Such an extension would have a wide range of applications, including the modeling of a trade involving more than two counterparties, each of which is associated with a different population.

- *Examining trading prices from Nash's bargaining game and the competitive equilibrium.* In Chapter 5, we find that the trading quantities from Nash's bargaining game and the competitive equilibrium are the same. This can be explained by the Pareto optimality of a contract. However, we have not found theoretic evidence to support the order of trading prices. It will be interesting to look into the trading prices and analytically answer the question which counterparty will benefit from the competitiveness.

- *Modeling a trade with multiple counterparties.* A real-world trade may involve more than two counterparties. An investigation of the Pareto optimality of such a trade may help us understand possible results of the trade. We may apply multi-player Nash bargaining game in this case to find a unique trading contract and examine how the risk is redistributed among counterparties.

- *Optimal trading contract.* In Chapter 5, we assume that there is only one type of trading contract available to the players. The structure of the mortality-linked security is fixed. It will be interesting and more realistic to give the players a set of contracts with different structures, and allow players to bargain which structure to be traded, in addition to price and quantity. This extension is different from Boonen et al. (2011) in that it does not allow a fully customized trading contract.

- *Other game theoretic methods.* Nash bargaining solution suppresses the details of the decision process. To understand how bargaining process takes place, we shall consider noncooperative bargaining games. Noncooperative games correspond to particular bargaining processes, and the outcome depends on the process we choose. Some realistic features can be imposed on these bargaining processes, for instance, information asymmetry and delay cost. Through examining these processes, we may get more realistic and accurate pricing results.

APPENDICES

Appendix A

Derivation of the Distributions of $Z_{\Delta_\kappa(t)}|Z_\kappa(t)$ and $Z_{\Delta_\gamma(t)}|Z_\gamma(t)$

Let $(X,Y)'$ be a bivariate normal random vector with mean $(\mu_X,\mu_Y)'$ and variance-covariance matrix S, where S is a symmetric nonnegative-definite 2-by-2 matrix. The probability density function for $(X,Y)'$ is

$$f_{X,Y}(x,y) = \frac{1}{2\pi\sqrt{\det(S)}} e^{-\frac{1}{2}(X-\mu_X,Y-\mu_Y)' S^{-1} (X-\mu_X,Y-\mu_Y)},$$

where

$$S^{-1} = \begin{pmatrix} \frac{S(2,2)}{\det(S)} & -\frac{S(1,2)}{\det(S)} \\ -\frac{S(1,2)}{\det(S)} & \frac{S(1,1)}{\det(S)} \end{pmatrix}.$$

is the inverse of S.

The conditional distribution of Y given $X = x$ is given by

$$
\begin{aligned}
f_{Y|X=x}(y) &= \frac{f_{X,Y}(x,y)}{f_X(x)} \\
&\propto f_{X,Y}(x,y) \\
&\propto e^{-\frac{1}{2}\left[(y-\mu_Y)^2 S^{-1}(2,2)+2(y-\mu_Y)(x-\mu_X)S^{-1}(1,2)\right]} \\
&\propto e^{-\frac{1}{2}\left\{S^{-1}(2,2)y^2-2\left[\mu_Y S^{-1}(2,2)-(x-\mu_X)S^{-1}(1,2)\right]y\right\}} \\
&\propto e^{-\frac{1}{2/(S^{-1}(2,2))}\left[y-\frac{\mu_Y S^{-1}(2,2)-(x-\mu_X)S^{-1}(1,2)}{S^{-1}(2,2)}\right]^2}.
\end{aligned}
$$

Therefore, $Y|X = x \sim \mathrm{N}\left(\frac{\mu_Y S^{-1}(2,2)-(x-\mu_X)S^{-1}(1,2)}{S^{-1}(2,2)}, \frac{1}{S^{-1}(2,2)}\right)$. Substituting S^{-1} into the expression, we obtain

$$
Y|X = x \sim \mathrm{N}\left(\frac{\mu_Y S(1,1) + (x-\mu_X)S(1,2)}{S(1,1)}, \frac{\det(S)}{S(1,1)}\right).
$$

Since $(Z_\kappa(t), Z_{\Delta_\kappa}(t))'$ and $(Z_\gamma(c), Z_{\Delta_\gamma}(c))'$ are zero-mean bivariate normal vectors with variance-covariance matrices V_κ and V_γ, respectively, we have

$$
Z_{\Delta_\kappa}(t)|Z_\kappa(t) \sim \mathrm{N}\left(Z_\kappa(t)\frac{V_\kappa(1,2)}{V_\kappa(1,1)}, \frac{\det(V_\kappa)}{V_\kappa(1,1)}\right),
$$

and

$$
Z_{\Delta_\gamma}(c)|Z_\gamma(c) \sim \mathrm{N}\left(Z_\gamma(c)\frac{V_\gamma(1,2)}{V_\gamma(1,1)}, \frac{\det(V_\gamma)}{V_\gamma(1,1)}\right).
$$

Appendix B

Deriving the Likelihood Function for the Transitory Jump Process

B.1 Likelihood Function for the Concurrent Jump Model

Define $\Delta_\kappa(t) = \kappa_t^{(1)} - \kappa_t^{(2)}$. We can rewrite $\kappa_{t+1}^{(1)}$ as

$$\kappa_{t+1}^{(1)} = \kappa_t^{(1)} + \mu_\kappa + Z_\kappa(t+1) - N_t Y_t^{(1)} + N_{t+1} Y_{t+1}^{(1)},$$

and $\Delta_\kappa(t+1)$ as

$$\Delta_\kappa(t+1) = \phi_{\Delta_\kappa}\Delta_\kappa(t) + \mu_{\Delta_\kappa} + Z_{\Delta_\kappa}(t+1) - \phi_{\Delta_\kappa} N_t (Y_t^{(1)} - Y_t^{(2)}) + N_{t+1}(Y_{t+1}^{(1)} - Y_{t+1}^{(2)}).$$

Let $\xi_t = \kappa_{t+1}^{(1)} - \kappa_t^{(1)}$ and $\varsigma_t = \Delta_\kappa(t+1) - \phi_{\Delta_\kappa}\Delta_\kappa(t)$. We can express ξ_t and ς_t as follows:

$$\begin{aligned}
\xi_t &= \mu_\kappa + Z_\kappa(t+1) - N_t Y_t^{(1)} + N_{t+1} Y_{t+1}^{(1)}; \\
\varsigma_t &= \mu_{\Delta_\kappa} + Z_{\Delta_\kappa}(t+1) - \phi_{\Delta_\kappa} N_t (Y_t^{(1)} - Y_t^{(2)}) + N_{t+1}(Y_{t+1}^{(1)} - Y_{t+1}^{(2)}).
\end{aligned}$$

The likelihood function for the concurrent-jump model is built upon the joint distribution of ξ_t and ς_t.

Suppose that the sample period is $[t_0, t_1]$. We let f_1 be the conditional density function for the random vector $(\xi_{t+1}, \varsigma_{t+1})'$ given $(\xi_t, \varsigma_t)'$, for $t \in \{t_0, \ldots, t_1 - 1\}$, and let f_2 be the density function for the random vector $(\xi_{t_0}, \varsigma_{t_0})'$. Then the log-likelihood can be written as

$$\ln\, f_2\left((\xi_{t_0}, \varsigma_{t_0})'\right) + \sum_{t=t_0}^{t_1-1} \ln\, f_1\left((\xi_{t+1}, \varsigma_{t+1})' | (\xi_t, \varsigma_t)'\right).$$

First, we evaluate f_1. It is easy to see that for $t \in \{t_0, \ldots, t_1 - 1\}$, $(\xi_{t+1}, \varsigma_{t+1})'$ given $(\xi_t, \varsigma_t)'$, N_t, N_{t+1} and N_{t+2} follows a bivariate normal distribution. Let f_3, M_3 and S_3 be the density function, the mean vector and the variance-covariance matrix of this bivariate normal distribution, respectively. The specifications of M_3 and S_3 depend on the value of N_{t+1}. In what follows, we use $W(i)$ to denote the ith element in a vector W and $X(i,j)$ to denote the (i,j)th element in a matrix X.

If $N_{t+1} = 0$, $(\xi_{t+1}, \varsigma_{t+1})'$ does not depend on $(\xi_t, \varsigma_t)'$. In this case, we have

$$\begin{aligned}
\xi_{t+1} &= \mu_\kappa + Z_\kappa(t+2) + N_{t+2} Y_{t+2}^{(1)}, \\
\varsigma_{t+1} &= \mu_{\Delta_\kappa} + Z_{\Delta_\kappa}(t+2) + N_{t+2}(Y_{t+2}^{(1)} - Y_{t+2}^{(2)}),
\end{aligned}$$

and the specifications of M_3 and S_3 are as follows:

$$\begin{aligned}
M_3(1) &= \mu_\kappa + N_{t+2}\mu_Y(1), \\
M_3(2) &= \mu_{\Delta_\kappa} + N_{t+2}(\mu_Y(1) - \mu_Y(2)), \\
S_3(1,1) &= V_Z(1,1) + N_{t+2}V_Y(1,1), \\
S_3(1,2) &= V_Z(1,2) + N_{t+2}(V_Y(1,1) - V_Y(1,2)), \\
S_3(2,2) &= V_Z(2,2) + N_{t+2}(V_Y(1,1) - 2V_Y(1,2) + V_Y(2,2)).
\end{aligned}$$

If $N_{t+1} = 1$, $(\xi_{t+1}, \varsigma_{t+1})'$ depends on $(\xi_t, \varsigma_t)'$ through $Y_{t+1}^{(1)}$ and $Y_{t+1}^{(2)}$. In this case, we have

$$\begin{aligned}
\xi_{t+1} &= -\xi_t + 2\mu_\kappa + Z_\kappa(t+1) + Z_\kappa(t+2) - N_t Y_t^{(1)} + N_{t+2} Y_{t+2}^{(1)}, \\
\varsigma_{t+1} &= -\phi_{\Delta_\kappa}\varsigma_t + (1 + \phi_{\Delta_\kappa})\mu_{\Delta_\kappa} + \phi_{\Delta_\kappa} Z_{\Delta_\kappa}(t+1) + Z_{\Delta_\kappa}(t+2) \\
&\quad -\phi_{\Delta_\kappa}^2 N_t(Y_t^{(1)} - Y_t^{(2)}) + N_{t+2}(Y_{t+2}^{(1)} - Y_{t+2}^{(2)}),
\end{aligned}$$

and the specifications of M_3 and S_3 are as follows:

$$\begin{aligned}
M_3(1) &= -\xi_t + 2\mu_\kappa - (N_t - N_{t+2})\mu_Y(1), \\
M_3(2) &= -\phi_{\Delta_\kappa}\varsigma_t + (1+\phi_{\Delta_\kappa})\mu_{\Delta_\kappa} - (\phi^2_{\Delta_\kappa}N_t - N_{t+2})(\mu_Y(1) - \mu_Y(2)), \\
S_3(1,1) &= 2V_Z(1,1) + (N_t + N_{t+2})V_Y(1,1), \\
S_3(1,2) &= (1+\phi_{\Delta_\kappa})V_Z(1,2) + (\phi^2_{\Delta_\kappa}N_t + N_{t+2})(V_Y(1,1) - V_Y(1,2)), \\
S_3(2,2) &= (1+\phi^2_{\Delta_\kappa})V_Z(2,2) + (\phi^4_{\Delta_\kappa}N_t + N_{t+2})(V_Y(1,1) - 2V_Y(1,2) + V_Y(2,2)).
\end{aligned}$$

Using the results above, we can compute f_1 by using the following formula:

$$\begin{aligned}
& f_1\left((\xi_{t+1}, \varsigma_{t+1})' | (\xi_t, \varsigma_t)'\right) \\
= & \sum_{i_1, i_1, i_3 = 0}^{1} f_3\left((\xi_{t+1}, \varsigma_{t+1})' | (\xi_t, \varsigma_t)', N_t = i_1, N_{t+1} = i_2, N_{t+2} = i_3\right) \\
& \Pr(N_t = i_1)\Pr(N_{t+1} = i_2)\Pr(N_{t+2} = i_3).
\end{aligned}$$

Next, we evaluate f_2. We can express ξ_{t_0} and ς_{t_0} as follows:

$$\begin{aligned}
\xi_{t_0} &= \mu_\kappa + Z_\kappa(t_0+1) - N_{t_0}Y^{(1)}_{t_0} + N_{t_0+1}Y^{(1)}_{t_0+1}, \\
\varsigma_{t_0} &= \mu_{\Delta_\kappa} + Z_{\Delta_\kappa}(t_0+1) - \phi_{\Delta_\kappa}N_{t_0}(Y^{(1)}_{t_0} - Y^{(2)}_{t_0}) + N_{t_0+1}(Y^{(1)}_{t_0+1} - Y^{(2)}_{t_0+1}).
\end{aligned}$$

Given N_{t_0} and N_{t_0+1}, the random vector $(\xi_{t_0}, \varsigma_{t_0})'$ follows a bivariate normal distribution with a density function f_4, a mean vector M_4 and a variance-covariance matrix S_4. The elements in M_4 and S_4 are as follows:

$$\begin{aligned}
M_4(1) &= \mu_\kappa - (N_{t_0} - N_{t_0+1})\mu_Y(1), \\
M_4(2) &= \mu_{\Delta_\kappa} - (\phi_{\Delta_\kappa}N_{t_0} - N_{t_0+1})(\mu_Y(1) - \mu_Y(2)), \\
S_4(1,1) &= V_Z(1,1) + (N_{t_0} + N_{t_0+1})V_Y(1,1), \\
S_4(1,2) &= V_Z(1,2) + (\phi_{\Delta_\kappa}N_{t_0} + N_{t_0+1})(V_Y(1,1) - V_Y(1,2)), \\
S_4(2,2) &= V_Z(2,2) + (\phi^2_{\Delta_\kappa}N_{t_0} + N_{t_0+1})(V_Y(1,1) - 2V_Y(1,2) + V_Y(2,2)).
\end{aligned}$$

We can then calculate f_2 using the following formula:

$$f_2\left((\xi_{t_0}, \varsigma_{t_0})'\right) = \sum_{i_1, i_2 = 0}^{1} f_4\left((\xi_{t_0}, \varsigma_{t_0})' | N_{t_0} = i_1, N_{t_0+1} = i_2\right)\Pr(N_{t_0} = i_1)\Pr(N_{t_0+1} = i_2).$$

Finally, with the expressions for f_1 and f_2, we can calculate the log-likelihood straightforwardly.

B.2 Likelihood Function for the Nonconcurrent Jump Model

As before, we let $\xi_t = \kappa_{t+1}^{(1)} - \kappa_t^{(1)}$ and $\varsigma_t = \Delta_\kappa(t+1) - \phi_{\Delta_\kappa}\Delta_\kappa(t)$. It is easy to show that

$$\begin{aligned}
\xi_t &= \mu_\kappa + Z_\kappa(t+1) - N_t^{(1)}Y_t^{(1)} + N_{t+1}^{(1)}Y_{t+1}^{(1)};\\
\varsigma_t &= \mu_{\Delta_\kappa} + Z_{\Delta_\kappa}(t+1) - \phi_{\Delta_\kappa}(N_t^{(1)}Y_t^{(1)} - N_t^{(2)}Y_t^{(2)}) + N_{t+1}^{(1)}Y_{t+1}^{(1)} - N_{t+1}^{(2)}Y_{t+1}^{(2)}.
\end{aligned}$$

The log-likelihood for the nonconcurrent-jump model is given by

$$\ln\, f_2\left((\xi_{t_0}, \varsigma_{t_0})'\right) + \sum_{t=t_0}^{t_1-1} \ln\, f_1\left((\xi_{t+1}, \varsigma_{t+1})' | (\xi_t, \varsigma_t)'\right),$$

where f_1 is the conditional density function for the random vector $(\xi_{t+1}, \varsigma_{t+1})'$ given $(\xi_t, \varsigma_t)'$, for $t \in \{t_0, \ldots, t_1 - 1\}$, and f_2 is the density function for the random vector $(\xi_{t_0}, \varsigma_{t_0})'$.

First, we evaluate f_1. It is easy to see that for $t \in \{t_0, \ldots, t_1 - 1\}$, $(\xi_{t+1}, \varsigma_{t+1})'$ given $(\xi_t, \varsigma_t)'$, $N_t^{(1)}$, $N_t^{(2)}$, $N_{t+1}^{(1)}$, $N_{t+1}^{(2)}$, $N_{t+2}^{(1)}$ and $N_{t+2}^{(2)}$ follows a bivariate normal distribution. Let f_3, M_3 and S_3 be the density function, the mean vector and the variance-covariance matrix of this bivariate normal distribution, respectively. The specifications of M_3 and S_3 depend on the values of $N_{t+1}^{(1)}$ and $N_{t+1}^{(2)}$.

If $N_{t+1}^{(1)} = 0$ and $N_{t+1}^{(2)} = 0$, then $(\xi_{t+1}, \varsigma_{t+1})'$ does not depend on $(\xi_t, \varsigma_t)'$. In this case, we have

$$\begin{aligned}
\xi_{t+1} &= \mu_\kappa + Z_\kappa(t+2) + N_{t+2}^{(1)}Y_{t+2}^{(1)},\\
\varsigma_{t+1} &= \mu_{\Delta_\kappa} + Z_{\Delta_\kappa}(t+2) + N_{t+2}^{(1)}Y_{t+2}^{(1)} - N_{t+2}^{(2)}Y_{t+2}^{(2)},
\end{aligned}$$

and the specifications of M_3 and S_3 are as follows:

$$
\begin{aligned}
M_3(1) &= \mu_\kappa + N_{t+2}^{(1)}\mu_Y(1), \\
M_3(2) &= \mu_{\Delta_\kappa} + N_{t+2}^{(1)}\mu_Y(1) - N_{t+2}^{(2)}\mu_Y(2), \\
S_3(1,1) &= V_Z(1,1) + N_{t+2}^{(1)}V_Y(1,1), \\
S_3(1,2) &= V_Z(1,2) + N_{t+2}^{(1)}V_Y(1,1) - N_{t+2}^{(1)}N_{t+2}^{(2)}V_Y(1,2), \\
S_3(2,2) &= V_Z(2,2) + N_{t+2}^{(1)}V_Y(1,1) - 2N_{t+2}^{(1)}N_{t+2}^{(2)}V_Y(1,2) + N_{t+2}^{(2)}V_Y(2,2).
\end{aligned}
$$

If $N_{t+1}^{(1)} = 0$ and $N_{t+1}^{(2)} = 1$, then ξ_{t+1} does not depend on ξ_t, while ς_{t+1} depends on ς_t. In this case, we have

$$
\begin{aligned}
\xi_{t+1} &= \mu_\kappa + Z_\kappa(t+2) + N_{t+2}^{(1)}Y_{t+2}^{(1)}, \\
\varsigma_{t+1} &= -\phi_{\Delta_\kappa}\varsigma_t + (1+\phi_{\Delta_\kappa})\mu_{\Delta_\kappa} + \phi_{\Delta_\kappa}Z_{\Delta_\kappa}(t+1) + Z_{\Delta_\kappa}(t+2) \\
&\quad -\phi_{\Delta_\kappa}^2(N_t^{(1)}Y_t^{(1)} - N_t^{(2)}Y_t^{(2)}) + N_{t+2}^{(1)}Y_{t+2}^{(1)} - N_{t+2}^{(2)}Y_{t+2}^{(2)},
\end{aligned}
$$

and the specifications of M_3 and S_3 are as follows:

$$
\begin{aligned}
M_3(1) &= \mu_\kappa + N_{t+2}^{(1)}\mu_Y(1), \\
M_3(2) &= -\phi_{\Delta_\kappa}\varsigma_t + (1+\phi_{\Delta_\kappa})\mu_{\Delta_\kappa} - \phi_{\Delta_\kappa}^2(N_t^{(1)}\mu_Y(1) - N_t^{(2)}\mu_Y(2)) + N_{t+2}^{(1)}\mu_Y(1) \\
&\quad -N_{t+2}^{(2)}\mu_Y(2), \\
S_3(1,1) &= V_Z(1,1) + N_{t+2}^{(1)}V_Y(1,1), \\
S_3(1,2) &= V_Z(1,2) + N_{t+2}^{(1)}V_Y(1,1) - N_{t+2}^{(1)}N_{t+2}^{(2)}V_Y(1,2), \\
S_3(2,2) &= (1+\phi_{\Delta_\kappa}^2)V_Z(2,2) + \phi_{\Delta_\kappa}^4(N_t^{(1)}V_Y(1,1) - 2N_t^{(1)}N_t^{(2)}V_Y(1,2) + N_t^{(2)}V_Y(2,2)), \\
&\quad +N_{t+2}^{(1)}V_Y(1,1) - 2N_{t+2}^{(1)}N_{t+2}^{(2)}V_Y(1,2) + N_{t+2}^{(2)}V_Y(2,2).
\end{aligned}
$$

If $N_{t+1}^{(1)} = 1$ and $N_{t+1}^{(2)} = 1$, then ξ_{t+1} depends on ξ_t and ς_{t+1} depends on ς_t. In this case, we have

$$
\begin{aligned}
\xi_{t+1} &= -\xi_t + 2\mu_\kappa + Z_\kappa(t+1) + Z_\kappa(t+2) + N_{t+2}^{(1)}Y_{t+2}^{(1)} - N_t^{(1)}Y_t^{(1)}, \\
\varsigma_{t+1} &= -\phi_{\Delta_\kappa}\varsigma_t + (1+\phi_{\Delta_\kappa})\mu_{\Delta_\kappa} + \phi_{\Delta_\kappa}Z_{\Delta_\kappa}(t+1) + Z_{\Delta_\kappa}(t+2) \\
&\quad -\phi_{\Delta_\kappa}^2(N_t^{(1)}Y_t^{(1)} - N_t^{(2)}Y_t^{(2)}) + N_{t+2}^{(1)}Y_{t+2}^{(1)} - N_{t+2}^{(2)}Y_{t+2}^{(2)}.
\end{aligned}
$$

and the specifications of M_3 and S_3 are as follows:

$$
\begin{aligned}
M_3(1) &= -\xi_t + 2\mu_\kappa + (N_{t+2}^{(1)} - N_t^{(1)})\mu_Y(1),\\
M_3(2) &= -\phi_{\Delta_\kappa}\varsigma_t + (1+\phi_{\Delta_\kappa})\mu_{\Delta_\kappa} - \phi_{\Delta_\kappa}^2(N_t^{(1)}\mu_Y(1) - N_t^{(2)}\mu_Y(2)) + N_{t+2}^{(1)}\mu_Y(1)\\
&\quad -N_{t+2}^{(2)}\mu_Y(2),\\
S_3(1,1) &= 2V_Z(1,1) + (N_{t+2}^{(1)} - N_t^{(1)})V_Y(1,1),\\
S_3(1,2) &= (1+\phi_{\Delta_\kappa})V_Z(1,2) + N_{t+2}^{(1)}V_Y(1,1) - N_{t+2}^{(1)}N_{t+2}^{(2)}V_Y(1,2)\\
&\quad +\phi_{\Delta_\kappa}^2(N_t^{(1)}V_Y(1,1) - N_t^{(1)}N_t^{(2)}V_Y(1,2)),\\
S_3(2,2) &= (1+\phi_{\Delta_\kappa}^2)V_Z(2,2) + \phi_{\Delta_\kappa}^4(N_t^{(1)}V_Y(1,1) - 2N_t^{(1)}N_t^{(2)}V_Y(1,2) + N_t^{(2)}V_Y(2,2)),\\
&\quad +N_{t+2}^{(1)}V_Y(1,1) - 2N_{t+2}^{(1)}N_{t+2}^{(2)}V_Y(1,2) + N_{t+2}^{(2)}V_Y(2,2).
\end{aligned}
$$

Using the results above, we can compute f_1 by using the following formula:

$$
\begin{aligned}
&f_1\left((\xi_{t+1},\varsigma_{t+1})'|(\xi_t,\varsigma_t)'\right)\\
= &\sum_{i_j=0,j=1,\ldots,6}^{1} f_3\Big((\xi_{t+1},\varsigma_{t+1})'|(\xi_t,\varsigma_t)', N_t^{(1)}=i_1, N_t^{(2)}=i_2, N_{t+1}^{(1)}=i_3, N_{t+1}^{(2)}=i_4,\\
&N_{t+2}^{(1)}=i_5, N_{t+2}^{(2)}=i_6\Big)\Pr(N_t^{(1)}=i_1, N_t^{(2)}=i_2)\Pr(N_{t+1}^{(1)}=i_3, N_{t+1}^{(2)}=i_4)\\
&\Pr(N_{t+2}^{(1)}=i_5, N_{t+2}^{(2)}=i_6).
\end{aligned}
$$

Next, we evaluate f_2. We can express ξ_{t_0} and ς_{t_0} as follows:

$$
\begin{aligned}
\xi_{t_0} &= \mu_\kappa + Z_\kappa(t_0+1) - N_{t_0}^{(1)}Y_{t_0}^{(1)} + N_{t_0+1}^{(1)}Y_{t_0+1}^{(1)},\\
\varsigma_{t_0} &= \mu_{\Delta_\kappa} + Z_{\Delta_\kappa}(t_0+1) + N_{t_0+1}^{(1)}Y_{t_0+1}^{(1)} - N_{t_0+1}^{(2)}Y_{t_0+1}^{(2)} - \phi_{\Delta_\kappa}(N_{t_0}^{(1)}Y_{t_0}^{(1)} - N_{t_0}^{(2)}Y_{t_0}^{(2)}).
\end{aligned}
$$

Given $N_{t_0}^{(1)}$, $N_{t_0}^{(2)}$, $N_{t_0+1}^{(1)}$ and $N_{t_0}^{(2)}$, the random vector $(\xi_{t_0},\varsigma_{t_0})'$ follows a bivariate normal distribution with a density function f_4, a mean vector M_4 and a variance-covariance matrix S_4. The elements in M_4 and S_4 are as follows:

$$\begin{aligned}
M_4(1) &= \mu_\kappa - (N_{t_0}^{(1)} - N_{t_0+1}^{(1)})\mu_Y(1), \\
M_4(2) &= \mu_{\Delta_\kappa} + (N_{t_0+1}^{(1)} - \phi_{\Delta_\kappa} N_{t_0}^{(1)})\mu_Y(1) - (N_{t_0+1}^{(2)} - \phi_{\Delta_\kappa} N_{t_0}^{(2)})\mu_Y(2), \\
S_4(1,1) &= V_Z(1,1) + N_{t_0}^{(1)} V_Y(1,1) + N_{t_0+1}^{(1)} V_Y(1,1), \\
S_4(1,2) &= V_Z(1,2) + \phi_{\Delta_\kappa}(N_{t_0}^{(1)} V_Y(1,1) - N_{t_0}^{(1)} N_{t_0}^{(2)} V_Y(1,2)) + N_{t_0+1}^{(1)} V_Y(1,1) \\
&\quad - N_{t_0+1}^{(1)} N_{t_0+1}^{(2)} V_Y(1,2), \\
S_4(2,2) &= V_Z(2,2) + N_{t_0+1}^{(1)} V_Y(1,1) - 2N_{t_0+1}^{(1)} N_{t_0+1}^{(2)} V_Y(1,2) + N_{t_0+1}^{(2)} V_Y(2,2), \\
&\quad + \phi_{\Delta_\kappa}^2 (N_{t_0}^{(1)} V_Y(1,1) - 2N_{t_0}^{(1)} N_{t_0}^{(2)} V_Y(1,2) + N_{t_0}^{(2)} V_Y(2,2)).
\end{aligned}$$

We can then calculate f_2 using the following formula:

$$\begin{aligned}
& f_2\left((\xi_{t_0}, \varsigma_{t_0})'\right) \\
= & \sum_{i_j=0, j=1,\ldots,4}^{1} f_4\left((\xi_{t_0}, \varsigma_{t_0})' | N_{t_0}^{(1)} = i_1, N_{t_0}^{(2)} = i_2, N_{t_0+1}^{(1)} = i_3, N_{t_0+1}^{(2)} = i_4\right) \\
& \Pr(N_{t_0}^{(1)} = i_1, N_{t_0}^{(2)} = i_2) \Pr(N_{t_0+1}^{(1)} = i_3, N_{t_0+1}^{(2)} = i_4).
\end{aligned}$$

Finally, with the expressions for f_1 and f_2, we can calculate the log-likelihood straightforwardly.

Bibliography

[1] Human mortality database. Technical report, University of California, Berkeley (USA), and Max Planck Institute of Demographic Research (Germany). Available at http://www.mortality.orgorhttp://www.humanmortality.de.

[2] J. Amendinger, P. Imkeller, and M. Schweizer. Additional logarithmic utility of an insider. *Stochastic Processes and their Applications*, 75:263–286, 1998.

[3] K.I. Amin. Jump diffusion option valuation in discrete time. *Journal of Finance*, 48(5):1833–1863, 1993.

[4] K.J. Arrow, H.D. Block, and L. Hurwicz. On the stability of the competitive equilibrium, II. *Econometrica*, 27(1):82–109, 1959.

[5] K.J. Arrow and L. Hurwicz. On the stability of the competitive equilibrium, I. *Econometrica*, 26(4):522–552, 1958.

[6] K.J. Arrow and L. Hurwicz. Some remarks on the equilibria of economic systems. *Econometrica*, 28(3):640–646, 1960.

[7] R.E. Bellman. *Dynamic Programming*. Princeton University Press, Princeton, N.J., 1957.

[8] E. Biffis. Affine processes for dynamic mortality and actuarial valuations. *Insurance: Mathematics and Economics*, 37(3):443–468, 2005.

[9] D. Blake, A.J.G. Cairns, and K. Dowd. Living with mortality: Longevity bonds and other mortality-linked securities. *British Actuarial Journal*, 12(1):153–197, 2006.

[10] D. Blake, A.J.G. Cairns, and K. Dowd. The birth of the life market. *Asia-Pacific Journal of Risk and Insurance*, 3(1):6–36, 2008.

[11] T. Boonen, A.D. Waegenaere, and H. Norde. Bargaining for over-the-counter risk redistributions: The case of longevity risk. 2011. Available at http://www.netspar.nl/files/Evenementen/2011-11-11%20PD/tim%20boonen.pdf.

[12] K. Borch. *The Mathematical Theory of Insurance*. Lexington Books, Lexington, Massachusetts., 1974.

[13] A.J.G. Cairns, D. Blake, and K. Dowd. A two-factor model for stochastic mortality with parameter uncertainty: Theory and calibration. *Journal of Risk and Insurance*, 73(4):687–718, 2006.

[14] A.J.G. Cairns, D. Blake, and K. Dowd. Modelling and management of mortality risk: a review. *Scandinavian Actuarial Journal*, 108(2-3):79–113, 2008.

[15] A.J.G. Cairns, D. Blake, K. Dowd, and G.D. Coughlan. Longevity hedge effectiveness: A decomposition. *Heriot-Watt University*, 2011b. Available at http://www.ma.hw.ac.uk/~andrewc/papers/ajgc60.pdf.

[16] A.J.G. Cairns, D. Blake, K. Dowd, G.D. Coughlan, D. Epstein, A. Ong, and I. Balevich. A quantitative comparison of stochastic mortality models using data from england and wales and the united states. *North American Actuarial Journal*, 13(1):1833–1863, 2009.

[17] A.J.G. Cairns, D. Blake, K. Dowd, G.D. Coughlan, and M. Khalaf-Allah. Bayesian stochastic mortality modelling for two populations. *ASTIN Bulletin*, 41(1):29–59, 2011a.

[18] L.R. Carter and R.D. Lee. Modeling and forecasting US sex differentials in mortality. *International Journal of Forecasting*, 8(3):393–411, 1992.

[19] C. Chen and L.-M. Liu. Joint estimation of model parameters and outlier effects in time series. *Journal of the American Statistical Association*, 88(421):727–751, 1993.

[20] C. Chen and G.C. Tiao. Random level shift time series models, arima approximations, and level shift detection. *Journal of Business and Economic Statistics*, 8(1):83–97, 1990.

[21] H. Chen and S.H. Cox. Modeling mortality with jumps: Application to mortality securitization. *Journal of Risk and Insurance*, 76(3):727–751, 2009.

[22] R. Cole and L. Fleischer. Fast-converging tâtonnement algorithms for one-time and ongoing market problems. *Proceedings of the 40th Annual ACM Symposium on Theory of Computing (STOC)*, pages 315–324, 2008.

[23] G.D. Coughlan. *The Handbook of Insurance-Linked Securities*, chapter Longevity Risk Transfer: Indices and Capital Market Solutions. West Sussex: John Wiley & Sons, 2009.

[24] G.D. Coughlan. Life and longevity markets association: The development of a longevity and mortality trading market. 2010. Presentation at the Sixth International Longevity Risk and Capital Markets Solutions Conference, Sydney, Australia. Available at http://www.longevity-risk.org/Presentations/Coughlan_LLMA.pdf.

[25] G.D. Coughlan, M. Khalaf-Allah, Y. Ye, S. Kumar, A.J.G. Cairns, D. Blake, and K. Dowd. Longevity hedging: A framework for longevity basis risk analysis and hedge effectiveness. *North American Actuarial Journal*, 15(2):150–176, 2011.

[26] S.H. Cox and Y. Lin. Natural hedging of life and annuity mortality risks. *North American Actuarial Journal*, 11(3):1–15, 2007.

[27] S.H. Cox, Y. Lin, and H. Pedersen. Mortality risk modeling: Applications to insurance securitization. *Insurance: Mathematics and Economics*, 46(1):242–253, 2010.

[28] J. Cvitanić and I. Karatzas. Convex duality in constrained portfolio optimization. *Annals of Applied Probability*, 2(4):767–818, 1992.

[29] Y. Deng, P.L. Brockett, and R.D. MacMinn. Longevity/mortality risk modeling and securities pricing. *Journal of Risk and Insurance*, 79(3):697–721, 2012.

[30] M. Denuit, P. Devolder, and A.-C. Goderniaux. Securitization of longevity risk: Pricing survivor bonds with wang transform in the Lee-Carter framework. *Journal of Risk and Insurance*, 74(1):87–113, 2007.

[31] K. Dowd, D. Blake, A.J.G. Cairns, and P. Dawson. Survivor swaps. *Journal of Risk and Insurance*, 73(1):1–17, 2006.

[32] K. Dowd, A.J.G. Cairns, D. Blake, G.D. Coughlan, D. Epsten, and M. Khalaf-Allah. A gravity model of mortality rates for two related populations. *North American Actuarial Journal*, 15(2):334–356, 2011.

[33] H.R. Duffie, D.and Richardson. Mean-variance hedging in continuous time. *Annals of Applied Probability*, 1(1):1–15, 1991.

[34] P. Emms and S. Haberman. Optimal management of an insurer's exposure in a competitive general insurance market. *North American Actuarial Journal*, 13(1):77–105, 2009.

[35] H. Föllmer and D. Sondermann. *Contributions to Mathematical Economics*, chapter Hedging of non-redundant contingent claims., pages 205–223. North-Holland, Amsterdam, 1986.

[36] R.J. Freud. The introduction of risk into a programming model. *Econometrica*, 24(3):253–263, 1956.

[37] R. M. Goodwin. Iteration, automatic computers and economic dynamics. *Metroeconomica*, 3(1):1–7, 1951.

[38] N. Hári, A.D. Waegenaere, B. Melenberg, and T.E. Nijman. Estimating the term structure of mortality. *Insurance: Mathematics and Economics*, 42(2):492–504, 2008.

[39] J. Harsanyi. Approaches to the bargaining problem before and after the theory of games: A critical discussion of zeuthen's, hicks', and nash's theories. *Econometrica*, 24(2):144–157, 1956.

[40] W. Jaffe. Another look at Leon Walras's theory of tâtonnement. *History of Political Economy*, 13(2):313–336, 1981.

[41] D.W. Katzner. Methodological individualism and the walrasian tâtonnement. *Journal of Economic and Social Research*, 1(1):5–33, 1999.

[42] R.E. Kihlstrom and A.E. Roth. Risk aversion and the negotiation of insurance contracts. *Journal of Risk and Insurance*, 49(3):372–387, 1982.

[43] M. Kitti. Convergence of iterative tâtonnement without price normalization. *Journal of Economic Dynamics and Control*, 34(6):1077–1091, 2010.

[44] R. Lee and L. Carter. Modeling and forecasting U.S. mortality. *Journal of the American Statistical Association*, 87(419):659–671, 1992.

[45] J.S.-H. Li. Pricing longevity risk with the parametric bootstrap: A maximum entropy approach. *Insurance: Mathematics and Economics*, 47(2):176–186, 2010.

[46] J.S.-H. Li and M.R. Hardy. Measuring basis risk in longevity hedges. *North American Actuarial Journal*, 15(2):177–200, 2010.

[47] J.S.-H. Li, M.R. Hardy, and K.S. Tan. Uncertainty in mortality forecasting: An extension to the classical Lee-Carter approach. *ASTIN Bulletin*, 39(1):137–164, 2009.

[48] J.S.-H. Li and A.C.Y. Ng. Canonical valuation of mortality-linked securities. *Journal of Risk and Insurance*, 78(4):853–884, 2011.

[49] N. Li and R. Lee. Coherent mortality forecasts for a group of population: An extension of the lee-carter method. *Demography*, 42(3):575–594, 2005.

[50] S.-H. Li and W.S. Chan. Outlier analysis and mortality forecasting: the united kingdom and scandinavian countries. *Scandinavian Actuarial Journal*, 2005(3):187–211, 2005.

[51] S.-H. Li and W.S. Chan. The Lee-Carter model for forecasting mortality, revisited. *North American Actuarial Journal*, 11(1):68–89, 2007.

[52] Y. Lin and S.H. Cox. Securitization of mortality risks in life annuities. *Journal of Risk and Insurance*, 72(2):227–252, 2005.

[53] A. Milidonis, Y. Lin, and S.H. Cox. Mortality regimes and pricing. *North American Actuarial Journal*, 15(2):266–289, 2011.

[54] J.F. Nash. The bargaining problem. *Econometrica*, 18(2):155–162, 1950.

[55] M.J. Osborne and A. Rubinstein. Bargaining and markets. 1990.

[56] PartnerRe. Life reinsurance pricing. Technical report, PartnerRe Ltd. Available at http://www.partnerre.com.

[57] J.W. Pratt. Risk aversion in the small and in the large. *Econometrica*, 32(1/2):122–136, 1964.

[58] M.L. Puterman. *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. John Wiley & Sons, New York; Toronto, 1994.

[59] R.S. Tsay. Outliers, level shifts and variance changes in time series. *Journal of Forecasting*, 7(1):1–20, 1988.

[60] H. Uzawa. Walras' tâtonnement in the theory of exchange. *The Review of Economic Studies*, 27(3):182–194, 1960.

[61] D.A. Walker. Walras's theories of tâtonnement. *Journal of Political Economy*, 95(4):758–774, 1987.

[62] L. Walras. *Élements d'Économie politique pure.* Corbaz, Lausanne, 1874.

[63] L. Walras. *Elements of Pure Economics, or The Theory of Social Wealth.* Irwin, Homewood, IL., 1954. Translated by William Jaffe from the definitive ed.

[64] S. Wang. Premium calculation by transforming the layer premium density. *ASTIN Bulletin*, 26(1):71–92, 1996.

[65] S. Wang. A class of distortion operations for pricing financial and insurance risks. *Journal of Risk and Insurance*, 67(1):15–36, 2000.

[66] S. Wang. A universal framework for pricing financial and insurance risks. *ASTIN Bulletin*, 32(2):213–234, 2002.

[67] K.M. White. Longevity advances in high-income countries, 1955-96. *Population and Development Review*, 28(1):59–76, 2002.

[68] C. Wilson. On the scale of global demographic convergence 1950-2000. *Population and Development Review*, 27(1):155–171, 2001.

[69] F. Zeuthen. *Problems of Monopoly and Economic Warfare.* G. Routledge &. Sons., London, 1930.

[70] Y. Zhu. Analysis of hedge fund investment choice. 2009. Paper presented in 2009 Northern Finance Association Meetings, Niagara-on-the-Lake, Ontario, Canada.