Partial Hedging of Equity-Linked Products in the Presence of Policyholder Surrender Using Risk Measures

Mehran Moghtadai

A Thesis

In The Department

of

Mathematics and Statistics

Presented in Partial Fulfillment of the Requirements

for the Degree of Master of Science (Mathematics) at

Concordia University

Montreal, Quebec, Canada

August 2014

CONCORDIA UNIVERSITY

School of Graduate Studies

This is to certify that the thesis prepared

By: **Mehran Moghtadai**

Entitled: **Partial Hedging of Equity-Linked Products in the Presence of Policyholder Surrender Using Risk Measures**

and submitted in partial fulfillment of the requirements for the degree of

Master of Science (Mathematics)

complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the final Examining Committee:

______________________________ Examiner

Dr. J. Garrido

______________________________ Examiner

Dr. C. Hyndman

______________________________ Thesis Supervisor

Dr. P. Gaillardetz

Approved by ______________________________

Chair of Department or Graduate Program Director

____

Dean of Faculty

Date ______________________________

Abstract

Partial Hedging of Equity-Linked Products in the Presence of Policyholder Surrender Using Risk Measures

Mehran Moghtadai

Throughout the past couple of decades, the surge in the sale of equity linked products has led to many discussions on the valuation of surrender options embedded in these products. However, most studies treat such options as American/Bermudian style options. In this thesis, a different approach is presented where only a portion of the policyholders react optimally, due to the belief that not all policyholders are rational. Through this method, a probability of surrender is found and the product is partially hedged by iteratively reducing the measure of risk to a non-positive value. This partial hedging framework is versatile since few assumptions are made. To demonstrate this, the initial capital requirement for an equity linked product is found under a bivariate equity/interest model with a copula based dependence structure. A numerical example is presented in order to demonstrate some of the dynamics of this valuation method. In addition, a surprising result is found during the adjustment of the surrender parameters which directly implies that under a particular valuation method, an increased number of policy surrenders causes a drop in the initial capital requirement. This counterintuitive result is directly caused by the partial hedging method.

To my parents: Masoud and Shirin.

Acknowledgements

First and foremost, I would like to express my gratitude to my supervisor Dr. Patrice Gaillardetz for his patience, support and input during my master's studies at Concordia University. This thesis would not have been possible without his guidance.

I would like to wholly thank the staff of the Mathematics and Statistics department at Concordia University for creating an environment that is greatly conducive to graduate research. I would also like to thank Dr. José Garrido for his continuous advice during my studies.

My deepest appreciation goes to my committee members Dr. José Garrido, Dr. Cody Hyndman and Dr. Patrice Gaillardetz. Their time and input has greatly improved the quality of this thesis.

In addition, I am forever indebted to my family. My father who encouraged curiosity, my mother who always stayed strong and my brother. Finally I would like to thank my girlfriend who stood by me during the whole of my post-secondary education and encouraged me throughout.

Contents

List of Figures

List of Tables

Introduction

Equity-linked products are a class of products that offer returns based on the stock market. They typically provide a limited participation in the performance of an equity index or a mutual fund, in the case of an equity-indexed annuity (EIA) or variable annuity (VA), respectively. This is done while guaranteeing a minimum rate of return. The monograph by Hardy (2003) has comprehensive discussions on the subject. Introduced by Keyport Life Insurance Co in 1995, EIAs have been the most innovative annuity product over the last 20 years. Equity-linked products have become increasingly popular since their debut and in the case of EIAs the sales have steadily increased and hit a record high of $33.9 billion in 2012. (See LIMRA (2012))

Due to this popularity EIA's have received much attention in the academic literature. Tiong (2000) and Lee (2003) have obtained closed-form formulas for several equity-indexed annuities in the Black-Scholes framework. Lin and Tan (2003) and Kijima and Wong (2007) consider a more general model for EIAs, in which both the external equity index and the interest rate satisfy general stochastic differential equations. Lin et al. (2009) price annuity guarantees under a regime-switching model. In these papers however, mortality is considered to be diversified.

Risk measures (see Artzner et al. (1999)) have also been used to evaluate equity-linked products. A comprehensive introduction to risk measures may be found in Wirch and Hardy (1999). The financial and actuarial approaches have been compared by Boyle and Hardy (1997). Tasche (2000) and Wang (2002) apply risk measures to insurance capital allocation problems.

A much studied approach for capital allocation in incomplete markets is to minimize the expected square of the losses, which is known as quadratic hedging. Incomplete markets arise in finance when the number of securities is less than the number of risk factors. Quadratic hedging was first introduced in the financial context by Föllmer and Sondermann (1986) and was later applied in an actuarial context in equity-linked products by Moller (1998). It has the advantage of being computationally simple, however the disadvantage is that the square of the loss does not distinguish positive losses from negative losses. A more tractable and meaningful approach for capital allocation is the quantile hedging approach introduced by Föllmer and Leukert (1999), applied to equity-linked products by Melnikov and Skornyakova (2005). This is equivalent to the value-at-risk (VaR) risk measure. Melnikov and Smirnov (2012) extend this principle and construct hedging strategies that minimize the Conditional Value-at-Risk (CVaR) with illustrations using equity-linked products. Conditional Value-at-Risk (CVaR) or Conditional Tail Expectation (CTE) has the advantage of being a coherent risk measure, see Artzner et al. (1997) as well economically meaningful, see Laeven and Goovaerts (2004).

In this thesis a partial hedging strategy of equity-linked products with surrender option is evaluated using iterated risk measures. For long-lived contracts such as equity-linked products, it is more reasonable to hedge dynamically. For this reason it is attractive to use an iterated approach such as those developed by Wang (1999). The initial hedging strategy is determined using the iterated value-at-risk (IVaR) and conditional value-at-risk (ICVaR) risk measures. The latter method is used in this thesis in order to determine the initial capital requirement for a partial hedging strategy for a portfolio of equity-linked products. This is done by minimizing the partial hedging costs while constraining the CVaR to non-positive values, made possible through the use of linear programming. Rockafellar and Uryasev (2000) present how linear programming can be used to minimize the CVaR and Gaillardetz and Hachem (2014) apply this idea for use in minimizing hedging costs.

This framework has the advantage of making as few assumptions as possible in order to function with various discrete financial models, risk measures, and products, both path-dependent and path-independent, leaving these decisions to the insurer for their particular market. This thesis takes advantage of this in order to bring forth a new method for evaluating surrenders in a realistic setting, particularly in the case where policyholders react non-optimally.

In Chapter 1 a bivariate economic model is presented. This model uses a discretized regime-switching form of the Cox-Ingersoll-Ross short rate model presented by Cox et al. (1985). The equity model uses a similar discretized regime-switching form of the geometric Brownian motion. These two models are joined using a Markov-switching Copula with Gaussian and student's t regimes. While it may appear that this author abuses regime-switching (Markov-switching) models, it is done with reason. The history of regime switching models goes back to the works of Quandt (1958), however this model was first introduced by Hamilton (1989) in a financial/economical context. The model permits the index to vary between discrete regimes in a Markovian manner. This behaviour consequently induces stochastic volatility while keeping the model simple. In addition, the nature of regime-switching models makes model comparison and reduction easy due to their nested nature.

In Chapter 2, the typically used actuarial notations are presented along with the risk measures used. Chapter 3 presents the dynamic partial hedging portfolio and its valuation through the use of iterated risk measures. Due to the economical model and example product being Markovian and path-independent, respectively, a stochastic mesh is used to evaluate the initial capital requirements of the partial hedging portfolio. In Chapter 4, the stochastic mesh method presented by Broadie and Glasserman (2004) is applied to the problem at hand. In Chapters 5.1 and 5, a numerical example is detailed along with a study of the impacts of different parameters on the initial capital requirement.

Chapter 1

Economic Model

In this chapter a multi-period bivariate discrete model will be presented that describes the dynamic of the stock index and the interest rate. Lattice models like these have been used to model stocks, interest rates, and other financial securities due to their flexibility and tractability; see Panjer et al. (1998) and Lin (2006) for examples. Generally, the premiums obtained from discrete models converge rapidly to the premiums obtained with the corresponding continuous models when considering equity-linked products.

In this chapter, continuous models for equity and interest rates will be presented, discretized and fitted to historical data. The model of choice for the equity and short rate will be the lognormal regime-switching model and the regime-switching Cox-Ingersoll-Ross model, respectively.

1.1 Interest Rate Model

A natural way of modeling interest rates in continuous time is through the use of a short rate model. This rate can be seen as the instantaneous spot rate. That is, the short rate which will be denoted by $r(t)$ from here forth is the annualized continuously compounded interest rate that one can borrow money for an infinitesimal period $\mathrm{d}t$. This notion can be formalized in the following way:

Definition 1.1.1 *Let $r(t)$ be the stochastic state variable for the short rate process and $\mathcal{F}_r(t) = \{r(s) : s \leq t\}$. Then the price of a zero-coupon bond with a face value of 1 at time t maturing at time $t + \Delta$ is given by*

$$P(t, t+\Delta) = \mathbb{E}^Q \left[\exp\left(- \int_t^{t+\Delta} r(s)\ \mathrm{d}s \right) \middle| \mathcal{F}_r(t) \right], \tag{1.1.1}$$

where $\mathbb{E}^Q(\cdot)$ is the expectation under the risk neutral measure.

A desirable characteristic in a short rate model is mean reversion. In theory this implies that the short rate will eventually move back towards the mean since interest rates cannot realistically grow beyond a certain point. An early model that captured this characteristic was introduced by Vasicek (1977). The Vasicek model however, carries two important disadvantages: level independent volatility and the possibility for negative interest rates. An extension to the Vasicek model which resolves these disadvantages is the Cox-Ingersoll-Ross (CIR) model presented by Cox et al. (1985). This model is a particular application of the square-root process. While the CIR model improves on Vasicek's, it still has the disadvantages of having constant volatility and assuming a homogeneous model. This can be improved by using a regime-switching version of the CIR model, which would let the process change to different CIR model states according to a continuous Markov chain.

Let us begin by defining the underlying continuous Markov chain,

Definition 1.1.2 *Let $X(t)$ be the random variable describing the state of the process at time t, and assume that the process is in state i at time t. Then $X(t+\Delta)$ is independent of previous values $\mathcal{F}_X(t) = \{X(s) : s \leq t\}$ and as $\Delta \to 0$ uniformly in t $\forall$ j*

$$\Pr(X(t+\Delta) = j | X(t) = i) = \delta_{ij} + q_{ij}\Delta + O(\Delta), \tag{1.1.2}$$

where

$$\delta_{ij} = \begin{cases} 1, & i = j \\ 0, & otherwise \end{cases}. \tag{1.1.3}$$

Then q_{ij} can be seen as the rate of transition from i to j.

Definition 1.1.3 *For all $t \in [0,T]$, let $X_r(t)$ be the state of a continuous time Markov chain on the finite space $\Omega = \{1,\ldots,K_r\}$ at time t defined as in Definition 1.1.2. A continuous time Cox-Ingersol-Ross regime-switching (RS-CIR) process $r(t)$ is the solution of the stochastic differential equation given by*

$$\mathrm{d}r(t) = \kappa_{X_r(t)}(\mu_{X_r(t)} - r(t))\ \mathrm{d}t + \sigma_{X_r(t)}\sqrt{r(t)}\ \mathrm{d}B(t) \tag{1.1.4}$$

where $B(t)$ is a Brownian motion, and $\kappa_{X_r(t)}, \mu_{X_r(t)}, \sigma_{X_r(t)}$ are the parameters in state $X_r(t)$.

Note that in the single regime CIR process, the drift component $\kappa(\mu - r(t))$ is the same as in the Vasicek model. This ensures that the process reverts to the mean μ with adjustment speed κ. However, while Vasicek's volatility component is a constant σ, CIR's depends on the level of $r(t)$ and as long as $2\kappa\mu > \sigma^2$ it precludes the possibility of negative interest rates.

1.1.1 Discretization

In order to fit the model to historical data a discretized form of the model is needed. Let $r(t)$ be the short rate at time t and $X_r(t)$ be the regime at time t.

Definition 1.1.4 *Let $\mathbf{P}(\Delta)$ be the transition matrix for the Markov chain $X(t)$, then $\mathbf{P}(\Delta) = \mathrm{e}^{\Delta\mathbf{Q}}$, where $\{\mathbf{Q}\}_{ij} = q_{ij}$ as in Definition 1.1.2. Then the transition matrix entries are defined as*

$$p_{ij}(\Delta) = \{\mathbf{P}(\Delta)\}_{ij} = \Pr(X(t+\Delta) = j | X(t) = i) \tag{1.1.5}$$

for $i, j \in \{1,\ldots,K\}$.

Per Definition 1.1.3, the exact discrete model corresponding to the square-root process given the regime is constant in $[t, t+\Delta)$ is given by,

$$\begin{aligned} r(t+\Delta)|r(t), X_r(t) = {} & \mathrm{e}^{-\kappa_{X_r(t)}\Delta} r(t) + \mu_{X_r(t)}\left(1 - \mathrm{e}^{-\kappa_{X_r(t)}\Delta}\right) \\ & + \sigma_{X_r(t)} \int_t^{t+\Delta} \mathrm{e}^{-\kappa_{X_r(t)}(t-s)} \sqrt{r(s)}\ \mathrm{d}B(s) \end{aligned} \tag{1.1.6}$$

Feller (1951) shows that the transition density of this model given regime $k \in \{1, \ldots, K\}$ is given by

$$f_k\left(r(t+\Delta)|r(t), X_r(t)=k\right) = c_k \mathrm{e}^{-u_k - v_k} \left(\frac{v_k}{u_k}\right)^{\frac{q_k}{2}} I_{q_k}(2(u_k v_k)^{\frac{1}{2}}), \tag{1.1.7}$$

where

$$c_k = \frac{2\kappa_k}{\sigma_k^2(1-\mathrm{e}^{-\kappa_k \Delta})}, \tag{1.1.8}$$

$$u_k = c_k r(t) \mathrm{e}^{-\kappa_k \Delta}, \tag{1.1.9}$$

$$v_k = c_k r(t), \tag{1.1.10}$$

$$q_k = \frac{2\kappa_k \mu_k}{\sigma_k^2} - 1, \tag{1.1.11}$$

and $I_{q_k}(\cdot)$ is the modified Bessel Function of the first kind of order q_k. Note that this density can also be seen as that of a non-central χ^2 distribution, where $2c_k r(t+\Delta)$ follows a non-central χ^2 distribution with $2(q_k+1)$ degrees of freedom and non-centrality parameter $2u_k$. It is important to also note that an extra condition exists on the parameters in order for the discrete model to follow this distribution, q_k must be greater or equal to 0. In the CIR model, this condition precludes the possibility of the process ever reaching 0. If $q_k < 0$, then $r(t)$ may occasionally hit 0.

1.1.2 Parameter Estimation

For the estimation of parameters, weekly yields on 4-week US T-Bill data will be used, as they are the shortest term US zero-coupon bonds available and are the closest candidates for an annualized short rate. A better approximation may be had by fitting a yield curve however additional assumptions are necessary to do so. For the purposes presented here, the data will be used directly. This data is obtained from the website of the US Department of the Treasury in weekly steps from the date of August 3rd, 2001 to July 12th, 2013 consisting of 624 data points.

The parameter constraint of $q_k > 0$ presented in the previous section may cause practical problems when fitting a single regime CIR model to this data. From early

August 2001 to early September 2008 the average annualized return on a 4-week US T-Bill was 2.52%, whereas the same average calculated between September 2008 and July 2013 was .0813%. In fact the need for a regime-switching model stemmed from this observed behaviour.

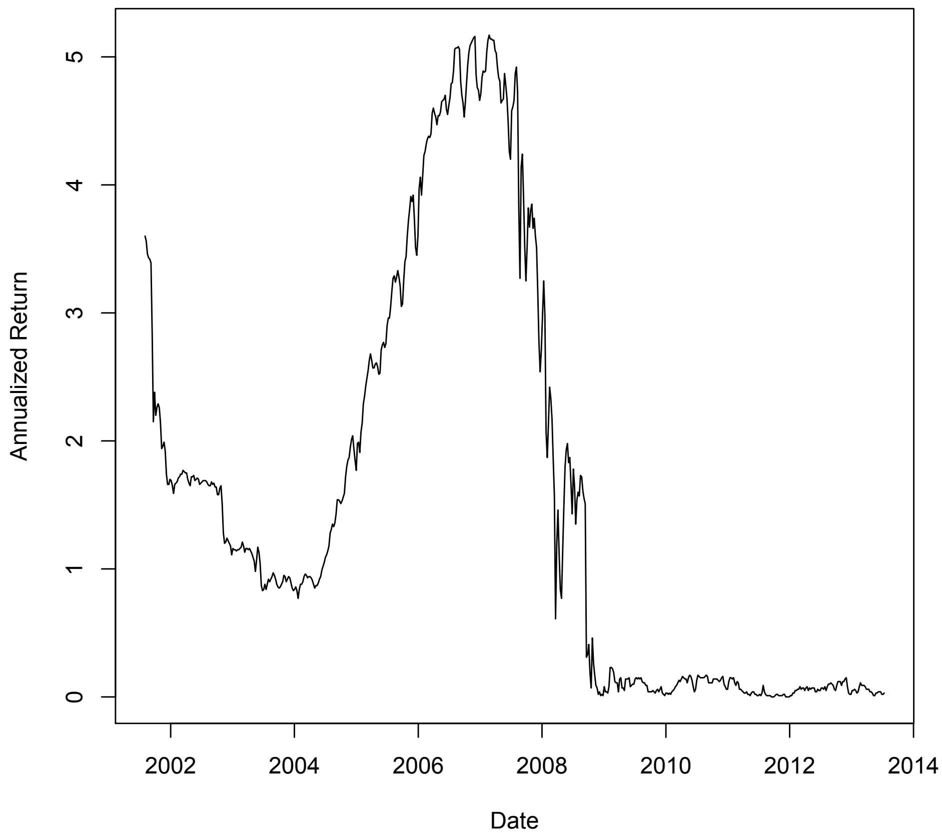

Figure 1.1: 4 Week annualized US T-Bill Returns

This large difference can be seen more effectively when this data is plotted as in Figure 1.1. It then becomes clear that when assuming a CIR model for interest rates, the pre-2008 and post-2008 crash returns do not come from the same process. We use the maximum likelihood method to fit a single regime CIR model and the quasi-likelihood method presented by Hamilton (1989) to fit multi-regime CIR models. The likelihood maximizations are programmed using *The R Project for Statistical Computing*. The code can be found in Appendix A.1. Note that due to the multi-modal nature of the likelihood function, the MLEs are obtained by using randomized starting values in order to find the global maximum.

Table 1.1: RS-CIR Model Comparisons

Regimes	# of Parameters	AIC	BIC
1	3	-6909.05	-6895.74
2	8	-6973.01	-6937.52
3	15	-6917.86	-6851.31

In order to pick the best model with the least number of parameters possible the Akaike Information Criteria (AIC) and Bayesian Information Criteria (BIC) will be used. This is defined as

$$AIC = 2p - 2\log \hat{L}, \tag{1.1.12}$$

$$BIC = p\log N - 2\log \hat{L} \tag{1.1.13}$$

where p is the number of parameters, $\hat{L}$ is the likelihood function evaluated at the estimates and N is the number of data points. In Table 1.1 it can be seen that the results of the AIC & BIC indicate that the best model is the two regime model.

Table 1.2: Estimated Parameters for RS-CIR Model

Regime i	$\hat{p}_{ii}^{(r)}(\frac{1}{52})$	$\hat{\kappa}_i$	$\hat{\mu}_i$	$\hat{\sigma}_i$
1	.99489	.071119	.10744	.12298
2	.58489	5.8046	.036573	.64452

In Table 1.2 the fitted parameters confirm what is observed in Figure 1.1. The second regime has very high mean reversion speed but with a lower mean whereas the first has a very low mean-reversion speed (κ) with a higher mean (μ). It is also worth noting that the persistence ($\hat{p}_{ii}^{(r)}(\frac{1}{52})$) of the first regime is considerably higher implying that the second regime is short lived.

1.2 Equity Model

The geometric Brownian motion (GBM) model famously used by Black and Scholes (1973) is tractable and easily implementable for equities. However the use of an independent lognormal model as the underlying market process carries with itself various assumptions, particularly that of constant volatility. Since then many empirical studies have been done that show how this model fails to capture the long-term extreme movements and variability of the market. Kurpiel and Roncalli (1998) compare the accuracy of option hedging strategies under stochastic volatility assumptions and those under the Black-Scholes model to show these drawbacks. One method to improve on the independent lognormal model would be to let volatility vary over time. A natural extension to the GBM model would be the regime-switching model. Hardy (2001) investigates the effectiveness of this model with the TSX and S&P indices. The regime-switching model allows us to capture the stochastic volatility of an equity index while keeping many of the simplicities of the GBM model. More specifically, the regime-switching lognormal model allows us to randomly vary the price process between K regimes according to a Markov chain.

Then we can define our lognormal regime-switching model in the following way,

Definition 1.2.1 *For all $t \in [0, T]$, let $X_S(t)$ be a continuous time Markov chain on finite space $\Omega = \{1, \ldots, K_S\}$ defined as in Definition 1.1.2. A continuous time lognormal regime-switching model is a stochastic process $S(t)$ which is the solution of the stochastic differential equation given by*

$$\mathrm{d}S(t) = \mu_{X_S(t)} S(t)\ \mathrm{d}t + \sigma_{X_S(t)} S(t)\ \mathrm{d}B(t), \tag{1.2.1}$$

where $B(t)$ is a Brownian motion, $\mu_{X_S(t)}$ and $\sigma_{X_S(t)}$ are the drift and volatility in state $X_S(t)$, respectively.

Hardy (2001) shows that a two regime model gives a good fit while keeping the model simple and meaningful in an economical setting, this will be verified for the given data set.

1.2.1 Discretization and Estimation of the Equity Model

In order to fit the model to historical data a discretized form of the model is needed. Let $S(t)$ be the equity process at time t and $X_S(t)$ be the regime at time t, then we have that,

$$\log \frac{S(t+\Delta)}{S(t)} | X_S(t) \sim \text{Normal}(\mu_{X_S(t)}\Delta, \sigma_{X_S(t)}\sqrt{\Delta}). \tag{1.2.2}$$

Table 1.3: RS-LN Model Comparisons

Regimes	# of Parameters	AIC	BIC
1	2	-2768.63	-2759.76
2	6	-2962.39	-2935.77
3	12	-2926.55	-2873.31

As in Section 1.1.2, a one, two, and three regime model is fitted. Comparing these models using the AIC and BIC in Table 1.3, it can be seen that the same conclusion as Hardy (2001) is found and that the two regime model is the best model.

Then it stands that the parameter vector $\Theta = (\mu_1, \mu_2, \sigma_1, \sigma_2, p_{11}(\Delta), p_{22}(\Delta))$ needs to be estimated. The likelihood functions and the procedure for fitting these parameters can be viewed in detail in Hamilton (1989). The numerical procedure for this fit is similar to the one for the CIR model and is in fact a trivial adjustment in the code presented in Appendix A.1.

Table 1.4: Estimated Parameters for RS-LN Model

Regime i	$\hat{p}_{ii}^{(S)}(\Delta)$	$\hat{\mu}_i\Delta$	$\hat{\sigma}_i\sqrt{\Delta}$
1	.96331	.003262	.01610
2	.9050	-.004639	.04246

The parameter estimates in Table 1.4 are based on weekly data of the S&P500 index between August 2001 and July 2013, in fact, this data is made to be in sync with that of Section 1.1.2. The regimes can be qualified as a low volatility and a high volatility regime. For the sake of consistency State 1 will be associated with low volatility and State 2 with high volatility. Note that the persistence of State 1 is higher than that of State 2, indicating that it is more likely for the process to stay in State 1 having started in State 1 then to stay in State 2 having started in State 2.

1.3 Dependence Model

The dependence between the equity and bond market is something that is thoroughly observed in classical finance. In fact this dependence plays a particularly important role in the capital asset pricing model (CAPM). Financial wisdom implies that as investors sell stock they will use these proceeds to fund bonds, or vice versa. Therefore it is expected that equity and bond prices move in opposite directions. In this framework since rates are used instead, the interpretation is that when higher yields are available on risk free assets, investors expect higher returns in the equity market. In recent years however the government has been keeping bond yields artificially low in order to promote equity participation, as a result looking at recent data, the dependence is rather low. This recent lack of dependence implies a heterogeneity and is essentially the same problem that had to be tackled in the equity and bond model.

Modeling the dependence by assuming a particular multivariate model is a difficult task, particularly when the two marginals can be shown to have widely different dynamics. However with the use of a copula function this task becomes much more manageable as it essentially facilitates a 'divide and conquer' strategy by making different assumptions on the marginals and the dependence structure. The use of copulas in finance was popularized by Embrechts et al. (2002) and Li (2000) who showed that the time-until-default of financial instruments was correlated and could be modeled using copulas. During the recent crisis copulas received bad press due to their use in pricing collateralized debt obligations (CDOs) yet they remain a powerful tool in both financial and actuarial settings when used with caution.

The first introduction of copulas was made by Sklar (1959). The definition of a copula function is given as follows.

Definition 1.3.1 *A d-dimensional copula* $C : [0,1]^d \to [0,1]$ *is a function which is a joint cumulative distribution function (CDF) with uniform marginals.*

This definition relies on the fact that if a random variable X has a continuous CDF F then $F(X)$ is distributed uniformly between 0 and 1. What completes the image is Sklar's Theorem which relates the marginal CDFs and copula to the multi-variate distribution.

Theorem 1.3.1 *(Sklar (1959)) Let* H *be a d-dimensional CDF with marginal CDFs* $F_1, F_2, \ldots, F_d$*. Then there exists a copula function* C *such that*

$$H(x_1, x_2, \ldots, x_d) = C(F_1(x_1), F_2(x_2), \ldots, F_d(x_d))$$

In turn, for any univariate CDF F_i *and any copula* C*, the function* H *is a d-dimensional CDF with marginals* $F_i,\ i \in \{1, \ldots, d\}$*. Furthermore, if* F_i *is continuous* $\forall i \in \{1, \ldots, d\}$*, then* C *is unique.*

The copula framework presented above is very general and opens the door to a vast set of possible copula functions. The question then becomes which copula to use. Embrechts (2009) gives an interesting overview of the historical context of copula functions and their derivations along with cautions. Since the purpose of this thesis is not to analyze this dependence but merely fit a meaningful model, the choice of copula will be limited between Gaussian and t copulas. Gaussian copulas put less emphasis on tail dependence while t-copulas put more. Both of these copulas belong to the elliptical copula family. Jondeau and Rockinger (2006) consider a conditional dependence copula using a Markov switching model that can be used when the marginals follow rather complicated distributions. This copula would then enable periods of high-correlation and low-correlation which can be observed in the data.

The Gaussian and t copulas are defined as follows:

Definition 1.3.2 *Let $R \in \mathbb{R}^{d\times d}$ be a correlation matrix, then the Gaussian copula with parameter R is defined by*

$$C_G(u_1, \ldots, u_d) = \Phi_R(\Phi^{-1}(u_1), \ldots, \Phi^{-1}(u_d)),$$

where Φ is the distribution function of a standard normal random variable and Φ_R is that of multivariate normal distribution with mean vector zero and covariance matrix equal to the correlation matrix R.

Definition 1.3.3 *Let $P \in \mathbb{R}^{d\times d}$ be a correlation matrix, then the Student's-t copula with parameters P, ν is defined by*

$$C_t(u_1, \ldots, u_d) = t_{\nu,P}(t_\nu^{-1}(u_1), \ldots, t_\nu^{-1}(u_d)),$$

where t_ν is the distribution function of a t random variable with ν degrees of freedom and $t_{\nu,P}$ is that of multivariate t distribution with covariance matrix equal to the correlation matrix P and ν degrees of freedom.

Definition 1.3.4 *Let $X_c(t),\ t \in \{0, \Delta, \ldots, \Delta N\}$ be a discrete Markov chain with K states and $\mathbf{P}_c$ be its transition matrix where*

$$p_{ij}^{(c)} = \{\mathbf{P}_c\}_{ij} = \Pr\left(X_c(t+\Delta) = j | X_c(t) = i\right), \tag{1.3.1}$$

for $i, j \in \{1, \ldots, K\}$. Then a d-dimensional Markov Switching Copula (MSC) is defined by

$$C_{MSC}(u_{1,t+\Delta}, \ldots, u_{d,t+\Delta} | X_c(t) = i) = C_i(u_{1,t+\Delta}, \ldots, u_{d,t+\Delta}), \tag{1.3.2}$$

where $i = \{1, \ldots, K\}$ and $u_{k,t}$ is the k^{th} marginal at time t.

Since the marginals together have a total of 14 parameters and a Markov switching copula would be adding at least 4 more, it would be extremely unwieldy to maximize the likelihood of the whole structure in one step. Instead, a two-step estimation will be used: first the marginals (which have already been estimated) then the copula parameters. Patton (2006) shows that this method yields estimators that are asymptotically efficient and normal. He also proves Sklar's Theorem for copulas whose marginals are conditional distributions.

Then the parameter estimate is given by

$$\hat{\mathbf{\Theta}}_c = \operatorname{argmax} \sum_{i=1}^{N} \log c_{MSC} \left(F_S \left(S(\Delta i) | S(\Delta i - 1); \hat{\mathbf{\Theta}}_S \right), F_r \left(r(\Delta i) | r(\Delta i - 1); \hat{\mathbf{\Theta}}_r \right); \mathbf{\Theta}_c \right), \tag{1.3.3}$$

where $c(\cdot)$ is the copula density and $\hat{\mathbf{\Theta}}_c$, $\hat{\mathbf{\Theta}}_S$ and $\hat{\mathbf{\Theta}}_r$ are the estimated parameters for the copula, equity and short rate models respectively. $F_S(\cdot)$ and $F_r(\cdot)$ are the distribution functions of the equity and short rate models respectively. Note that the log likelihood in (1.3.3) disregards the marginals since they have had their parameters already estimated and are in fact constants. The Markov switching copula parameters can be estimated using the same quasi likelihood method detailed by Hamilton (1989).

Table 1.5: MSC Model Comparisons

Model	# of Parameters	AIC	BIC
Gaussian	1	2.10203	6.53818
t	2	-10.1997	-1.327399
Gaussian-Gaussian	4	-21.9025	-4.157899
Gaussian-*t*	5	-31.0640	-8.8832
t-*t*	6	-29.3771	-2.76019

The estimations will be restricted to single and two state models. In Table 1.5 the models are listed in order of complexity along with their AIC. Keeping in mind that the lowest AIC and BIC implies the best fit, the Gaussian-*t* MSC seems to give the best fit for the number of parameters. It is important to keep in mind that the models are

Table 1.6: Estimated Parameters for MSC Model

State i	$\hat{p}_{ii}^{(c)}$	$\hat{\rho}_i$	$\hat{\nu}_i$
1 (Gaussian)	.9965	.05409	-
2 (t)	.9923	.08536	1.0550

simply being compared to each other and no formal goodness-of-fit test is taking place, Genest et al. (2009) gives an overview of goodness-of-fit tests for copulas. It is important however to take into account that due to the conditional nature of the marginals, standard goodness-of-fit tests may be misleading and care should be taken.

While the goodness-of-fit issue will not be tackled here, a likelihood ratio test will be considered in order to justify the use of this dependence model over an independent model. Note that the independent copula, $C_I(u, v) = uv$, can be looked at as a special case of the Markov switching copula presented here, therefore the condition of nested model is satisfied. The test is

$$H_0 = C_I \quad vs \quad H_1 = C_{MSC}. \tag{1.3.4}$$

Then the test statistic is given by

$$t = -2\left(l(\hat{\theta}_I) - l(\hat{\theta}_{MSC})\right). \tag{1.3.5}$$

Note that the likelihood under the null is 1 and hence the log-likelihood, $l(\cdot)$, is equal to 0. Consequently the degrees of freedom under the null is 0 due to the lack of parameters. Then $t = 41.064$ and has a $\chi^2(5-0)$ distribution. The p-value for this test is $9.108\text{e}-08$, hence the null hypothesis is rejected with a very high significance. Therefore it can be stated that the dependence model presented is a worthwhile improvement over the independence assumption.

For the purposes of this thesis the Gaussian-t MSC will be used due to its large improvement on the single state models. The improvement of the t-t MSC does not appear to be worth the increase in complexity. Note that in the two state case the parameter matrix R can be reduced to the correlation coefficient ρ. In Table 1.6 the fitted parameters imply a very high persistence along with low correlation in both states. The fitted degrees of freedom parameter is rather small implying heavy tails in that state along with tail dependence.

Chapter 2

Actuarial Model

In Chapter 1, the model presented assumes the usual frictionless market. In this chapter the actuarial notations for the mortality and surrender process will be presented along with the dynamic risk measures. The simulation framework that will be presented will assume that the insurance provider can only trade in certain intervals, that is, if the simulation makes weekly movements, the provider can only trade monthly or quarterly. It will also be assumed that there is a constant number of trading dates N_1 each year and $\Delta_1 = 1/N_1$ is the length of time between each trading date. Assuming that there are a total of N movements in the simulation, there are $N_2 = N/N_1$ unobserved movements between the trading dates during each Δ_1 period. Also let $\Delta = 1/(N_2 N_1)$ be the length of time between each simulation step.

2.1 Mortality and Surrender Model

The actuarial notation that will be presented is that described by Bowers (1997). Throughout this thesis it will be assumed that the surrenders occur at the very beginning of a period and deaths occur during said period while payments associated with death are made at the end of the period.

Definition 2.1.1 *Let $l_0^{(\tau)}$ be the number of mutually independent policyholders aged $x_1, \ldots, x_{l_0}$ at time of issue, and $T(x_i)$ and $U(x_i)$ be the time until death and time until surrender for policyholder (x_i) at time $t = 0$, respectively. Then the curtate time until exit is defined as:*

$$K(x_i) = \lfloor N_1 \min\left(T(x_i), U(x_i)\right) \rfloor \Delta_1,$$

computed up to multiples of Δ_1 periods. Note that $\lfloor \cdot \rfloor$ is the floor function.

The surrender process is that of a discrete decrement that occurs at the beginning of each period Δ_1, independent from mortality. Let the rate of surrender for the period $[k, k + \Delta_1)$ be denoted by $q_k^{\prime(s)}$. The probability of death for the period $[k, k + \Delta_1)$ is denoted by ${}_{\Delta_1}q_{x+k}^{\prime(d)}$.

Let the probability that (x) remains in the cohort k years be denoted by ${}_kp_x^{(\tau)} = \Pr[K(x) \geq k]$. Then let ${}_{k|\Delta_1}q_x^{(\tau)}$ denote the probability that (x) remains a policyholder for k years and exits within the following Δ_1 period, i.e

$${}_{k|\Delta_1}q_x^{(\tau)} = \Pr[K(x) = k] = {}_kp_x^{(\tau)}\,{}_{\Delta_1}q_{x+k}^{(\tau)} = {}_kp_x^{(\tau)} \left(q_k^{\prime(s)} + p_k^{\prime(s)}\,{}_{\Delta_1}q_{x+k}^{\prime(d)} \right)$$

for $k \in \{0, \Delta_1, 2\Delta_1, \ldots\}$. Similarly the probability of remaining a policyholder by the end of the period is ${}_{\Delta_1}p_{x+k}^{(\tau)} = p_k^{\prime(s)}\,{}_{\Delta_1}p_{x+k}^{\prime(d)}$. Note that the surrender probabilities do not depend on the age of the policyholder. The choice of these probabilities will be presented in CSection 4.2.1.

From here forth, without loss of generality, it will be assumed that l_0 contracts are signed with l_0 policyholders at time 0. It is also assumed that the policyholders are all independent, of the same age x, and their mortality follows the same law.

Definition 2.1.2 *Let $\mathcal{L}_k$ be the cohort of remaining policyholders right before the surrenders at time k:*

$$\mathcal{L}_k = \sum_{i=1}^{l_0} \mathbf{1}_{\{\min\{T(x_i),U(x_i),n_i\}\geq k\}} = l_0\, {}_kp_x^{(\tau)}, \tag{2.1.1}$$

where $T(x_i)$ and $U(x_i)$ are the times until exits as defined in (2.1.1). Then let ${}_{\Delta_1}\mathcal{D}_k^{(s)}$ and ${}_{\Delta_1}\mathcal{D}_k^{(d)}$ be the number of policyholders that surrender at time k and die in the period $[k, k+\Delta_1)$ for $k \in \{0, \Delta_1, \ldots\}$, respectively, i.e.

$$\mathcal{D}_k^{(s)} = \sum_{i=1}^{l_0} \mathbf{1}_{\{U(x_i)=k\}}, \tag{2.1.2}$$

$${}_{\Delta_1}\mathcal{D}_k^{(d)} = \sum_{i=1}^{l_0} \mathbf{1}_{\{T(x_i)\in[k,k+\Delta_1)\}}. \tag{2.1.3}$$

2.2 Risk Measures

Risk measures have been widely used by financial institutions such as insurance and investments companies to evaluate the risk level of business lines. This widespread use is mainly due to its meaningfulness in a business setting. Mathematically speaking a risk measure is defined as a mapping from a set of random variables to the real line. This definition is rather broad and encompasses many mappings that do not necessarily give any idea as to the level of risk in a corporate sense. The most common and well known risk measure is the Value-at-Risk (VaR). VaR is widely used due to its ease of implementation and interpretability in risk management, and regulatory requirements. More recently however the Conditional Value-at-Risk (CVaR) has been lauded as a more meaningful and appropriate risk measure because of the recognition that coherence, as defined by Artzner et al. (1999), is a desirable property of risk measures. Essentially, a coherent risk measure is said to possess the properties of monotonicity, sub-additivity, positive homogeneity, and translation invariance. Note that VaR does not possess the sub-additivity property and is therefore not a coherent risk measure. In this thesis the focus will be mostly on CVaR, due to the fact that CVaR retains its coherence for discrete distributions.

Definition 2.2.1 *The Value-at-Risk at time k is defined as the c-quantile of the discounted loss random variable L. That is*

$$R_1\left(L|\mathcal{F}(k)\right) = \inf\{y \in \mathbb{R} : \Pr[L > y|\mathcal{F}(k)] \leq 1 - c\} \tag{2.2.1}$$

for $c \in [0,1]$, $k \in \{0, \Delta_1, 2\Delta_1, \ldots\}$ and $\mathcal{F}(k)$ is the natural filtration associated with the information on L up to time k.

Definition 2.2.2 *The CVaR risk measure, which represents the expected value of the worst $(1-c)$ losses, is given at time k by*

$$R_2\left(L|\mathcal{F}(k)\right) = \mathbb{E}\left[L|\mathcal{F}(k), L > R_1(L)\right]. \tag{2.2.2}$$

In the case of a discrete loss random variable this can be expressed explicitly as

$$R_2\left(L|\mathcal{F}(k)\right) = \pi_c + \frac{\mathbb{E}\left[(L-\pi_c)\mathbf{1}_{\{L>\pi_c\}}|\mathcal{F}(k)\right]}{1-c}, \tag{2.2.3}$$

where $\pi_c = R_1\left(L|\mathcal{F}(k)\right)$.

Wang (1999) proposes the use of risk measures in an iterative manner for discounted random variables, which implies that the risk measure at time k is a function of the loss random variable and the future capital requirements. This class of risk measures is particularly useful in cases where there is an evolution in the future liabilities. This concept has been applied to equity-linked insurance products with risk measures (see Hardy and Wirch (2004) and Gaillardetz and Moghtadai (2014)).

Definition 2.2.3 *The iterative risk measure at time k for a discounted random loss is given by*

$$DR_j\left(L|\mathcal{F}(k)\right) = R_j\left(DR_j(L|\mathcal{F}(k+\Delta_1))|\mathcal{F}(k)\right), \tag{2.2.4}$$

for $j \in \{1,2\}$, and $k \in \{0, \Delta_1, \ldots, n-2\Delta_1\}$, where $DR_j\left(L|\mathcal{F}(n-\Delta_1)\right) = R_j\left(L|\mathcal{F}(n-\Delta_1)\right)$ and n represents the latest possible liability payment.

In other words the iteratively calculated measure of risk at time k is the measure of risk as a function of the iterative measure of risk at time $k + \Delta_1$.

Chapter 3

Valuation

In this chapter the general iterative evaluation method will be presented for a product that is contingent on mortality and surrender behaviors. Independently from this contingency is the involvement in the financial market, which is the unique and defining feature of equity-linked products. Typically, this involvement is through the level of survival and death benefit that is linked to the performance of the financial market. The random payoff denoted $B(x, K(x))$, at time 0 of this contract for the policyholder (x) is given by

$$\begin{cases} D(x, K(x) + \Delta_1)\mathbf{1}_{\{T(x) \leq U(x)\}} \\ \quad + (1 - C_{K(x)})D(x, K(x))\mathbf{1}_{\{T(x) > U(x)\}} & , \text{ if } K(x) \in \{0, \Delta_1, \ldots, n - \Delta_1\} \\ D(x, n) \ , & \text{if } K(x) \in \{n, \ldots\} \end{cases} \tag{3.0.1}$$

where $C_{K(x)} \in [0, 1]$ is the surrender fee at time $K(x)$ and $D(x, K(x))$ is the deatch benefit for policyholder (x) at time $K(x)$. Note that here we assume that the death and surrender benefits are paid at the end and beginning of the period respectively.

The framework put forth will use the discounted loss random variables that represent the aggregate losses incurred by the issue of these contracts. These include the gains and expenses that occur in the portfolio. It will be assumed that the initial capital will be invested in a hedge portfolio denoted W which consists of index shares and a risk free asset. The purpose of this framework is to calculate the initial capital requirement as well as the hedge ratio.

Definition 3.0.4 *Let $\mathcal{A} = \{a(k), b(k)\}$ denote the hedging strategy at time $k \in \{0, \Delta_1, \ldots\}$, where $a(k)$ is the dollar portion invested in index shares and $b(k)$ is the dollar portion invested in the risk free asset at time k.*

Definition 3.0.5 *Let $W(t,k),\ k > t$ denote the value of the accumulated aggregate hedge portfolio at time k given $\mathcal{F}(t) = \{S(u), r(u), \mathcal{L}_u : u \leq t \cup \mathcal{D}_t^{(s)}\}$. Then $W(k,k) = a(k)+b(k)$ for $k \in \{0, \Delta_1, \ldots\}$, and $W(0,0) = a(0)+b(0)$ is the initial capital requirement of this hedging strategy. Similarly $W(k, k+\Delta_1^-)$ denotes the value of the hedge portfolio prior to any benefit payment at time $k+\Delta_1$ given $\mathcal{F}(k)$, then we have that the accumulated hedge portfolio at time $k+\Delta_1^-$ is the hedge portfolio from time k accumulated for 1 period,*

$$W(k, k+\Delta_1^-) = a(k)\frac{S(k+\Delta_1)}{S(k)} + b(k)\mathrm{e}^{r(k)\Delta_1} \tag{3.0.2}$$

for $k \in \{0, \Delta_1, \ldots, n-\Delta_1\}$.

3.1 Loss Random Variable

By leveraging the use of dynamic risk measures as in Definition 2.2.3 it is possible to establish a discounted loss random variable L by using the conditional losses between time k and $k+\Delta_1$. Starting with the last period, the conditional loss variable at time $n-\Delta_1$ is the discounted difference between the guarantees paid and the accumulated hedge portfolio,

$$\begin{aligned} L|\mathcal{F}(n-\Delta_1) = {} & (1 - C_{n-\Delta_1}) \sum_{i=1}^{\mathcal{D}_{n-\Delta_1}^{(s)}} D(x, n-\Delta_1) \\ & + \mathrm{e}^{-r(n-\Delta_1)\Delta_1} \left(\sum_{i=1}^{\mathcal{L}_{n-\Delta_1} - \mathcal{D}_{n-\Delta_1}^{(s)}} D(x_i, n) - W(n-\Delta_1, n^-) \right). \end{aligned} \tag{3.1.1}$$

For periods $k \in \{0, \Delta_1, \ldots, n-2\Delta_1\}$ the discounted loss random variable needs to take into consideration the benefits for the current period as well as the risk associated with the remaining cohort. The conditional discounted loss random variable is the discounted difference between the expenses and the accumulated investment portfolio,

$$
\begin{aligned}
L|\mathcal{F}(k) &= (1 - C_k) \sum_{i=1}^{\mathcal{D}_k^{(s)}} D(x_i, k) \\
&\quad + \mathrm{e}^{-r(k)\Delta_1} \left(\sum_{i=1}^{\Delta_1 \mathcal{D}_k^{(d)}} D(x_i, k + \Delta_1) + \sum_{i=1}^{\mathcal{L}_{k+\Delta_1}} DR_j(L|\mathcal{F}(k + \Delta_1)) - W(k, k + \Delta_1^-) \right),
\end{aligned}
\tag{3.1.2}
$$

for $j = 1, 2$.

The probability distribution associated with this random variable will be presented in Section 4.3. This loss random variable can then be used to define the time 0 loss associated with issuing such a contract. Typically it is desirable to set the equity-linked product parameters such that the time 0 measure of loss is set to 0. These parameters are said to set the fair value.

3.2 Dynamic Hedging Portfolio

The iterative risk measures method that will be used to evaluate the initial capital requirement is similar to the framework first presented by Gaillardetz and Moghtadai (2014), where a portfolio of equity shares and risk-free assets are iteratively rebalanced while minimizing cost and setting the measure of risk to zero in a two step optimization process. This method is an extension of the iterative CTE method presented by Hardy and Wirch (2004), where they iteratively discount the payoff. A consequence of the latter method is that it assumes the initial capital is wholly invested in risk-free assets leading to potentially higher initial capital costs. The former hedging strategy however uses the fact that one can always lower their risk by increasing the available capital, in other words one can find a particular hedging portfolio which sets the measure of risk at a certain level to zero. In fact, due to the unboundedness of risk measures there is an uncountably infinite set of hedging portfolios for which the measure of risk at a certain level is zero. The method in this chapter will similarly minimize the cost, however instead of setting the measure of risk to zero, it constrains it to be non-positive.

The level of risk measure and the type of risk measure need to be set by the insurance company in a balancing act. The minimum cost at a high level of risk may still be very costly and make the contract unattractive, while minimizing the cost under a low level of risk may cause insolvency due to market shock in the latter years of the contract.

Rockafellar and Uryasev (2000) present a method for minimizing the CVaR through linear programming using approximations. The main approximation step relies on the fact that the CVaR for a continuous loss random variable can be found by generating a finite number of random variables through Monte Carlo simulation and finding the discrete CVaR instead, in essence turning Equation (2.2.2) into Equation (2.2.3). Then through the use of auxiliary variables they transform the CVaR into a linear expression that can be used in linear programming. This is convenient since it will permit the hedging portfolio to be set by using a randomly generated tree. While their purpose is to minimize the CVaR, here it is the cost that is optimized while the CVaR is a constraint, this is presented by Gaillardetz and Hachem (2014).

The procedure will involve finding the optimum portfolio denoted $\mathcal{A}^*(k) = \{a^*(k), b^*(k)\}$ iteratively for $k \in \{0, \Delta_1, \ldots, n - \Delta_1\}$ starting from time $n - \Delta_1$. The linear expression that is to be minimized at each step is simply the sum of the portions in equity shares and risk free assets, i.e.

$$\{a^*(k), b^*(k)\} = \underset{a(k), b(k) \geq 0}{\text{argmin}}\ a(k) + b(k). \tag{3.2.1}$$

This objective is then minimized with respect to the constraint of $DR_2(L|\mathcal{F}(k)) \leq 0$ at level c. To do this let $L^{(j)}|\mathcal{F}(k)$ for $j \in \{1, \ldots, m\}$ represent the j-th discounted loss random variable that is generated from the random variable $L|\mathcal{F}(k)$ and p_j be the probability associated with that outcome. Using the formulation presented by Rockafellar and Uryasev (2000), the CVaR constraint can be written in a linear way as follows,

$$-L^j|\mathcal{F}(k) + \alpha^+ - \alpha^- + u_j > 0, \quad \text{for } j \in \{1, \ldots, q\}, \tag{3.2.2}$$

$$u_j \geq 0, \quad \text{for } j \in \{1, \ldots, m\}, \tag{3.2.3}$$

$$\alpha^-, \alpha^+ \geq 0, \tag{3.2.4}$$

$$\alpha^+ - \alpha^- + \sum_{j=1}^{m} \frac{u_j p_j}{1-c} \leq 0, \tag{3.2.5}$$

where u_j for $j \in \{1, \ldots, m\}$ are auxiliary variables and $(\alpha^+ - \alpha^-)$ is the VaR at level c.

Note that the variables that are being optimized, namely $a(k)$ and $b(k)$, are part of the $W(k, k + \Delta_1^-)$ portion of loss random variable. Since this optimization is done iteratively starting with the last period, it is assumed that when finding $\mathcal{A}^*(k)$, $DR_2(L|\mathcal{F}(k + \Delta_1))$ was calculated using $\mathcal{A}^*(k + \Delta_1)$. This process will then eventually give us the initial capital requirement $\mathcal{A}^*(0)$.

As previously stated, equity-linked contracts are generally issued with a fair value parameter. Some parameters are set and one of the parameters is determined using numerical methods such that the measure of loss at time 0 is equal to zero. In the above framework it is also possible to find such a parameter for a homogeneous group of policyholders.

Chapter 4

Stochastic Mesh

Due to the nature of the problem presented in Section 3.2, everything must take place in discrete time. One of the most popular and tractable models for pricing options and option style liabilities is the binomial (or multinomial) option pricing model first presented by Cox et al. (1979). This model has the advantage of converging to the continuous model (for certain types of models) and being extremely tractable by having the recombining property depending on the type of liability being priced. Bollen (1998) extends this model into a pentanomial recombining lattice that can be used to price options when the underlying equity movements follow a regime-switching model. While this latter extension can be used for our equity and interest rate models separately, it is not appropriate under the dependence conditions presented here.

It is important to note that in the scenario presented in this thesis there are four dimensions: equity, interest rate, mortality and surrender. In this situation one might want to use simulation. One method for doing so is the random tree method presented by Broadie and Glasserman (1997) where m paths are simulated from the starting node and m paths from each of those and so on (see Figure 4.1). It is then easy to see that this method is not very tractable as the size of the tree explodes in an exponential manner, in fact the computation requirements are $O(m^N)$, where N is the total number of steps. While they show that only using 3 steps for a 1 year option on a single asset they are able to get relatively accurate results, it is hard to extend this for long term multi-dimensional liabilities. Another method that can be used to tackle this problem is the stochastic mesh presented by Broadie and Glasserman (2004), this will be the method used in this thesis.

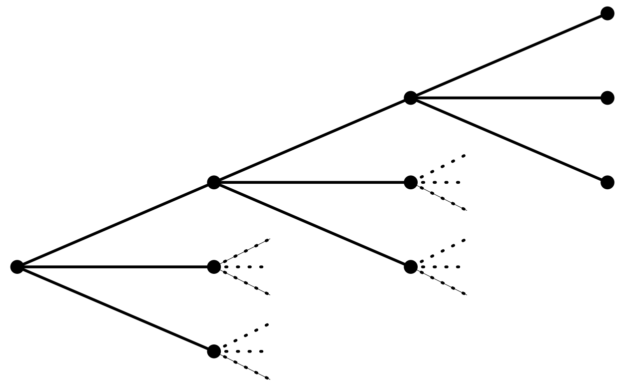

Figure 4.1: Random Tree ($m = 3$)

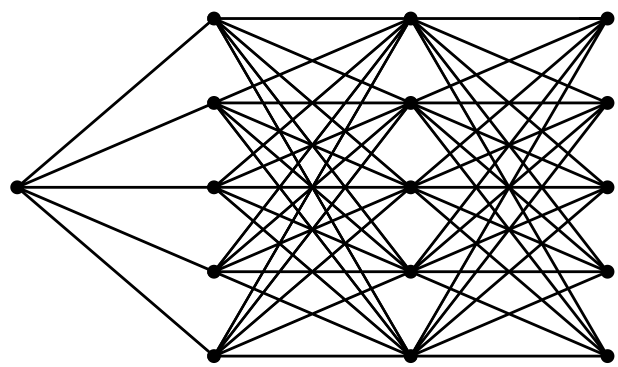

Figure 4.2: Stochastic Mesh ($m = 5$)

Broadie and Glasserman (2004) propose that under certain conditions, one can simulate m independent paths with N steps each and at each node use all m states from the next time step by assigning an appropriate weight W_{jk}^{t}, where $t \in \{0, \Delta_1, \ldots, n-\Delta\}$ and $j, k \in \{1, \ldots, m\}$, in other words the weight for going from node j at time step t to k at time step $t + \Delta_1$ (see Figure 4.2). The advantage over the random tree becomes evident as the number of nodes at each time step is fixed, avoiding the exponential growth, in fact this method is $O(Nm^2)$.

Let $\mathbf{Y}_i = (Y_{t1}, \ldots, Y_{tm})$ denote all nodes at step $t \in \{0, \Delta_1, \ldots, n - \Delta\}$, then the conditions under which a stochastic mesh holds are

1. $\{\mathbf{Y}_0, \ldots, \mathbf{Y}_{t-\Delta_1}\}$ and $\{\mathbf{Y}_{t+\Delta_1}, \ldots, \mathbf{Y}_n\}$ are independent given $\mathbf{Y}_t$, for all $t \in \{0, \ldots, n-\Delta_1\}$

2. Each weight W_{jk}^t is a deterministic function of $\mathbf{Y}_t$ and $\mathbf{Y}_{t+\Delta_1}$.

3. For all $t \in \{0, \Delta_1, \ldots, n-\Delta_1\}$ and all $j \in \{1, \ldots, m\}$,

$$\frac{1}{m}\sum_{k=1}^{m} \mathbb{E}\left[W_{jk}^t V_{t+\Delta_1}(Y_{t+\Delta_1,k})|\mathbf{Y}_t\right] = G_t(Y_{tj}), \tag{4.0.1}$$

   where $G_t(Y_{tj})$ is the continuation value in state Y_{tj} and $V_{t+\Delta_1}(Y_{t+\Delta_1,k})$ is the value of the option in state $Y_{t+\Delta_1,k}$.

Condition 1 is satisfied by our economic model since every process involved is Markovian by nature, and the same is true for the surrender and mortality processes. Condition 2 and 3 depend on the choice of weights. Broadie and Glasserman (2004) suggest the use of likelihood weights. These weights can be used to satisfy Conditions 2 and 3 as long as the paths are generated using a model that satisfies Condition 1. They can then be defined in the following manner,

$$W_{jk}^t = \begin{cases} \dfrac{f(Y_{t+\Delta_1,k}|Y_{tj})}{f(Y_{t+\Delta_1,k}|Y_{00})}, & k \neq j \\ 1, & k = j, \end{cases} \tag{4.0.2}$$

where $f(\cdot|Y_{tj})$ and $f(\cdot|Y_{00})$ are transition densities from each respective state. The general idea behind these weights is that since for $j \neq k$, $Y_{t+\Delta_1,k}$ is not generated from Y_{tk}, the weights need to be adjusted with respect to the state from which they were generated which is Y_{00}. Using these weights the value of the American option at time t is given by,

$$\hat{V}_t(Y_{tj}) = \max\left\{\frac{1}{m}\sum_{k=1}^{m} W_{jk}^t e^{\Delta r(Y_{tj})}\hat{V}_{t+\Delta_1}(Y_{t+\Delta_1,k}), D_t(Y_{tj})\right\}, \tag{4.0.3}$$

where $D_t(Y_{tj})$ is the immediate exercise value at time t in path j.

Broadie and Glasserman (2004) mention that this framework produces high bias estimates that converge to the real value as $m \to \infty$ and $\Delta \to 0$. They suggest the use of an improved estimator that interleaves the high bias estimator with a low bias estimator. However the application of this estimator is not computationally straightforward with this model and the purpose of this thesis is not to get an accurate price but to study the impact of surrender on the valuation method. Hence, a modified form of the original high bias estimator will be used.

While Equation (4.0.3) is convenient for calculating expectations, it is not appropriate for the valuation method used in this thesis. Since the valuation method essentially uses a tail expectation, it is more appropriate and convenient to use probabilities instead of weights. Equation (4.0.3) would suggest that $\frac{W_{jk}^t}{m}$ is the probability of going from state Y_{tj} to Y_{tk}. However this would not necessarily sum to 1 and is therefore not appropriate. To correct for this, the sum of the weights are used instead of m. These probabilities are then defined in the following way,

$$\Pr\left(Y_{t+\Delta,k}|Y_{tj}\right) = p_{jk}^t = \frac{W_{jk}^t}{W_{j\bullet}^t}, \tag{4.0.4}$$

where $W_{j\bullet}^t = \sum_k W_{jk}^t$

In order to show that these two methods are essentially equivalent, an American option on a risk asset modeeled as a geometric Brownian motion is used for a numerical example. The price of an American Call is calculated with initial stock price of 1 and strike of 1. The constant interest rate is set to 3% and the volatility is 40%. The price calculated is for a 1 year option and the length of the time steps Δ is set to 0.1. Fifty simulations are used for each number of paths tested. The exact price of the option is calculated using the Black-Scholes formula.

In Figure 4.3 it can be seen that for a low number of paths the presented probabilities actually give a more accurate result. However as the number of paths increases, both methods converge to the true price. It should be noted that at each simulation the same simulated paths are used for both methods and it can be seen that the presented method is always lower than the alternate method. The code for this test can be seen in Appendix A.2.1

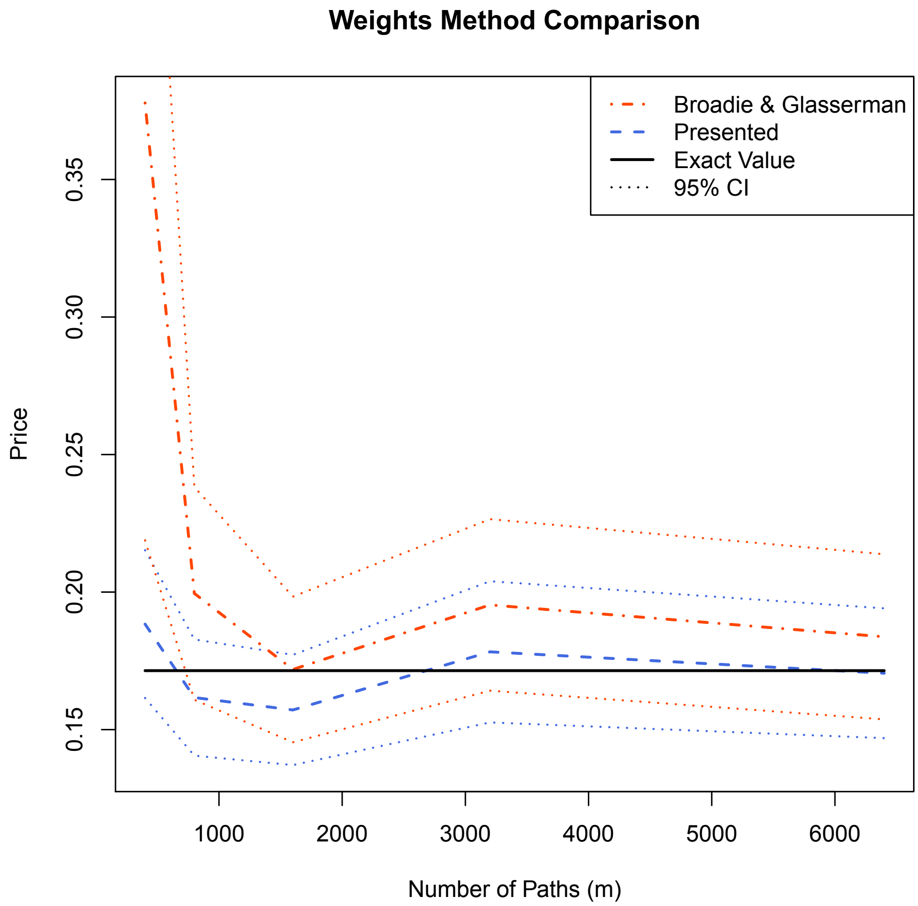

Figure 4.3: Weighing Methods

4.1 Economic Model Simulation

Following the general framework presented in the previous section, the simulation of paths simplifies due to the independent path generation scheme. For the economic model presented in Chapter 1 this can be done in two steps for each path, first by simulating N copulas following the appropriate dynamic and second by using the inversion method with the copula generated probabilities in order to generate a path for each process (equity and interest rate).

The copula generation can be done by first simulating $N-1$ states from the $X_c(t)$ Markov chain with a known $X_c(0)$ and then simulating N pairs of uniform random variables $\mathbf{v}_t = (v_t^{(S)}, v_t^{(r)}),\ t \in \{\Delta, 2\Delta, \ldots, n\}$ from either a t-copula or a Gaussian copula depending on the outcome of $X_c(t), t \in \{0, \Delta, 2\Delta, \ldots, n-\Delta\}$. More explicitly,

$$\mathbf{v}_t = \{(v_t^{(S)}, v_t^{(r)})\} = \mathbf{1}_{\{\hat{X}_c(t)=1\}}\Phi\left(\Phi^{-1}_{(\rho_1)}(u)\right) + \mathbf{1}_{\{\hat{X}_c(t)=2\}}t_{(\nu_2)}\left(t^{-1}_{(\rho_2,\nu_2)}(u)\right), \tag{4.1.1}$$

where $\Phi^{-1}_{(\rho_1)}(\cdot)$ is the inverse bivariate normal CDF and $\Phi(\cdot)$ is the standard normal CDF applied element wise. Similarly $t^{-1}_{(\rho_2,\nu_2)}(\cdot)$ is the inverse bivariate t CDF and $t_{(\nu_2)}(\cdot)$ is the univariate t CDF applied element wise. Also u is a generated uniform random variable. Note that the inverse bivariate normal CDF evaluated at a point represents an uncountable set. Instead a pair of correlated values can be found using the Cholesky decomposition, that is, $\Phi^{-1}_{(\rho_1)}(u)$ is fixed to a particular set by generating 2 independent uniform random variables (u_1, u_2),

$$u_1' = u_1, \tag{4.1.2}$$

$$u_2' = \rho_1 u_1' + u_2\sqrt{1-\rho_1^2}. \tag{4.1.3}$$

Finally the generated copula given that $\hat{X}_c(t) = 1$ is $\mathbf{v}_t = (\Phi(u_1'), \Phi(u_2'))$.

For the t copula the multivariate normal variance mixture representation $t \sim \sqrt{W}\mathbf{Z}$ can be used where $\mathbf{Z}$ is a multivariate normal with mean 0 and correlation matrix $\mathbf{R}$ and $\nu/W \sim \chi^2_\nu$. Then similar to the Cholesky decomposition above,

$$u_1' = \rho_2 u_2 + \frac{u_1\sqrt{1-\rho_2^2}}{\sqrt{w\nu_2}}, \tag{4.1.4}$$

$$u_2' = \frac{u_2}{\sqrt{w\nu_2}}, \tag{4.1.5}$$

where w is a generated $\chi^2_{\nu_2}$ random variable and u_1, u_2 are independently generated uniform random variables. Finally the generated copula given that $\hat{X}_c(t) = 2$ is $\mathbf{v}_t = \left(t_{(\nu_2)}(u_1'), t_{(\nu_2)}(u_2')\right)$.

This is generally where a variance reduction method would be applied. However, one must be careful in how this is done since the dependence structure can be broken by either using unadjusted antithetic variables or Latin hypercube sampling. Packham and Schmidt (2008) demonstrate a modified Latin hypercube sampling with dependence by using rank statistics. In this thesis however, through experimentation it became apparent that variance minimization biases the results since the valuation method is based on the tail expectation of the simulated paths. Therefore variance reduction was omitted.

4.1.1 Interest Rate Path Generation

Note that in order to generate a path for the regime switching Cox-Ingersoll-Ross (RS-CIR) model it is necessary to know the conditional distribution for each step. Given that the regime is k at time t, $2c_k r(t+\Delta)$ follows a χ^2 distribution with degrees of freedom $2(q_k+1)$ and non-centrality parameter $2u_k$ as described in Section 1.1.1. For generating the value at $t=\Delta$, $X_r(0)=\omega$ is known and therefore the value of $r(\Delta)$ can be simulated by simply inverting the appropriate CDF, i.e. $H_\omega^{(r)-1}(v_\Delta^{(r)})/2/c_\omega$, where $H_\omega^{(r)}(\cdot)$ is the distribution function of a non-central χ^2 with state ω parameters, this must be found numerically. However for $t=\Delta, 2\Delta, \ldots, N\Delta$ the regime is unknown, instead the probability of being in either state 1 or 2 is known. These probabilities are defined as $p_{\omega 1}(t)$ and $p_{\omega 2}(t)$. Therefore the distribution is in fact a mixture. This mixture distribution with parameters adjusted for Δ length steps is defined by,

$$F_\Delta^{(r)}(r(t+\Delta)|r(t), X_r(0)=\omega) = H_1^{(r)}(2c_1 r(t+\Delta))p_{\omega 1}(t) + H_2^{(r)}(2c_2 r(t+\Delta))p_{\omega 2}(t), \tag{4.1.6}$$

for $t \in \{0, \Delta, 2\Delta, \ldots, n-\Delta\}$. This function can then be used to find $\hat{r}(t+\Delta) = F_\Delta^{(r)-1}(v_{t+\Delta}^{(r)}|\hat{r}(t), X_r(0)=\omega)$, which can be done numerically. The numerical method can be accelerated by using an Euler type approximation in order to obtain a good starting value. To do so, the implicit Milstein scheme can be used. Alfonsi (2005) gives an overview of various discretization methods for CIR processes. This method has the advantage of being computationally simple while still converging to the true process. For a single regime this is given by,

$$\hat{r}'(t+\Delta) \approx \frac{\hat{r}(t) + \kappa\mu\Delta + \sigma\sqrt{\Delta\hat{r}(t)}Z}{1+\kappa\Delta}, \tag{4.1.7}$$

this can then be extended to two regimes:

$$\hat{r}'(t+\Delta) \approx \mathbf{1}_{\{X_r(t)=1\}} \frac{\hat{r}(t) + \kappa_1\mu_1\Delta + \sigma_1\sqrt{\Delta\hat{r}(t)}Z}{1+\kappa_1\Delta} + \mathbf{1}_{\{X_r(t)=2\}} \frac{\hat{r}(t) + \kappa_2\mu_2\Delta + \sigma_2\sqrt{\Delta\hat{r}(t)}Z}{1+\kappa_2\Delta}, \tag{4.1.8}$$

however as previously stated, $\mathbf{1}_{\{X_r(t)=1\}}$ is a random variable at time t, therefore its expected value will be used instead to simulate the value at $t+\Delta$, then

$$\hat{r}'(t+\Delta) \approx p_{\omega 1}(t) \frac{\hat{r}(t) + \kappa_1\mu_1\Delta + \sigma_1\sqrt{\Delta\hat{r}(t)}Z}{1+\kappa_1\Delta} + p_{\omega 2}(t) \frac{\hat{r}(t) + \kappa_2\mu_2\Delta + \sigma_2\sqrt{\Delta\hat{r}(t)}Z}{1+\kappa_2\Delta}, \tag{4.1.9}$$

for $t \in \{\Delta, \ldots, (N-1)\Delta\}$, where $Z \sim N(0,1)$. Therefore the starting value $\hat{r}'(t+\Delta)$ can be generated by generating Z through inversion using $v_{t+\Delta}^{(r)}$. Note that the Milstein scheme produces positive interest rates as long as the parameters satisfy the following condition: $4\kappa\mu > \sigma^2$ (this is the case in both regimes).

4.1.2 Equity Path Generation

Using $v_t^{(S)}$ it is possible to generate the equity path since $\frac{S(t+\Delta)}{S(t)}|X_S(t)$ is normally distributed. However since $X_S(t)$ is not exactly known, it is a mixture of normals, that is,

$$\log \frac{S(t+\Delta)}{\hat{S}(t)} | \hat{S}(t), X_S(0) = \omega \sim N([\mu_1 p_{\omega 1}(t) + \mu_2 p_{\omega 2}(t)]\Delta, [\sigma_1 p_{\omega 1}(t) + \sigma_2 p_{\omega 2}(t)]\sqrt{\Delta}). \tag{4.1.10}$$

This can then be generated by inverting a standard normal CDF using $v_t^{(S)}$ and then applying the mean and variance transformation. In other words,

$$\hat{S}(t+\Delta) = \hat{S}(t) \exp\left\{[\mu_1 p_{\omega 1}(t) + \mu_2 p_{\omega 2}(t)]\Delta + [\sigma_1 p_{\omega 1}(t) + \sigma_2 p_{\omega 2}(t)]\sqrt{\Delta}Z\right\}, \tag{4.1.11}$$

for $t \in \{0, \Delta, \ldots, n-\Delta\}$, given starting values $S(0)$ and $X_S(0)$, and where Z is a generated $N(0,1)$ random variable.

4.1.3 Economic Model Mesh

Finally using the above schemes m pairs of paths can be constructed, the elements of these paths will be denoted by $(\hat{S}^{(i)}(t), \hat{r}^{(i)}(t))$, for $i \in \{1, \dots, m\}$ and $t \in \{\Delta_1, 2\Delta_1 \dots, n\}$. Note that while simulations are done in N times Δ length steps, only N_1 states are observable at steps of length Δ_1. Using this information the weights for the economic model mesh denoted by We can then be constructed,

$$We^t_{jk} = \begin{cases} \dfrac{f_{\Delta_1}\left(S(t+\Delta_1) = \hat{S}^{(k)}(t+\Delta_1), r(t+\Delta_1) = \hat{r}^{(k)}(t+\Delta_1) | S(t) = \hat{S}^{(j)}(t), r(t) = \hat{r}^{(j)}(t)\right)}{f_{t+\Delta_1}\left(S(t+\Delta_1) = \hat{S}^{(k)}(t+\Delta_1), r(t+\Delta_1) = \hat{r}^{(k)}(t+\Delta_1) | S(0), r(0)\right)}, & k \neq j \\ 1, & k = j \end{cases} \tag{4.1.12}$$

where $f_h(\cdot)$ is the density of the bivariate model with parameters adjusted from for time steps of length h, i.e.

$$f_h(X = x, Y = y | X_0 = x_0, Y_0 = y_0) = c_{h,MSC}\left(F^{(S)}_h(x|x_0), F^{(r)}_h(y|y_0)\right) f^{(S)}_h(x|x_0) f^{(r)}_h(y|y_0). \tag{4.1.13}$$

Note that the function $c_{MSC}(\cdot, \cdot)$ depends on the Markov chain state and again since the states are not known for the equity and interest rate, both $F^{(S)}_h, f^{(S)}_h$ and $F^{(r)}_h, f^{(r)}_h$ are mixtures that depend on their respective Markov chains. It is important to note that the parameters need to be adjusted for time steps of length Δ_1 and $t + \Delta_1$ for the numerator and denominator, respectively. Finally the probabilities are given by $pe^t_{jk} = \frac{We^t_{jk}}{We^t_{j\bullet}}$ as defined in (4.0.4).

4.2 Mortality and Surrender Simulation

Due to the assumed independence, mortality can be simulated independently from the economic model. However, in order to simulate the cohort alive at each step it is also necessary to simulate the number of policyholders that surrender their contract. This surrender scheme will depend on the outcomes of the economic model. The surrender probabilities will depend on a moneyness ratio which is denoted by MR. This ratio is found by treating the equity-linked product as a Bermudian option with an exercise fee. More precisely, it is the ratio of the surrender value to the continuation value. It is found independently from the hedging portfolio presented in Chapter 3 and is instead found by looking at this option from the perspective of a rational option holder. From this perspective at each possible exercise time the policyholder will evaluate his/her position by comparing the exercise value to the prospective value measured by staying in the contract and will always pick the choice with the highest value. Let $\hat{V}^{(k)}(t)$ be the continuation value for the period $[k, k+\Delta)$, $\hat{P}^{(k)}(t)$ be the option value and $(1-C_t)\hat{D}^{(k)}(t)$ be the value and the surrender value at time $t \in \{0, \Delta_1, 2\Delta_1, \ldots, n\}$ for path $k \in \{1, \ldots, m\}$, respectively. Then,

$$MR^{(k)}(t) = \frac{(1-C_t)\hat{D}^{(k)}(t)}{\hat{V}^{(k)}(t)}, \tag{4.2.1}$$

$$\hat{P}^{(k)}(t) = \max\left\{(1-C_t)\hat{D}^{(k)}(t), \hat{V}^{(k)}(t)\right\}, \tag{4.2.2}$$

where $V^{(k)}(t)$ is the continuation value. This continuation value is then found under a risk-neutral measure. This is done in this fashion since the product is treated as a tradable option and therefore the price needs to be adjusted with respect to the investors/policyholders risk profile. However under a complete market with no arbitrage opportunities this price can be found by using a risk-neutral measure which incorporates the market risk premium. Such conditions lead to a unique price.

In the market conditions presented in this thesis, no assumption is made about the completeness of the market. In fact the presented valuation partial hedging method is most useful under the assumption that the market is incomplete. In this situation the physical probabilities can be adjusted in order to find a risk-neutral measure. However in an incomplete market setting the price is no longer unique. By proceeding this way the replication cost can be found directly by calculating the expectation in the following manner:

$$\hat{V}^{(k)}(t) = \mathrm{e}^{\Delta_1 \hat{r}^{(k)}(t)} \sum_{i=1}^{m} \hat{P}^{(i)}(t+\Delta_1) q e^{t}_{ki}, \tag{4.2.3}$$

for $t \in \{0, \Delta_1, \ldots, n - \Delta_1\}$ where $\hat{V}^{(i)}(n) = D^{(i)}(n)$ and qe^t_{ki} is the probability of going from state X_{tk} to $X_{t+\Delta_1,i}$ under a risk neutral measure. For the purposes of finding such a measure, the Esscher transform presented by Gerber and Shiu (1994) will be used:

$$qe^t_{ki}(h) = \frac{\exp\left(h e^{\left(\Delta_1 \hat{r}^{(k)}(t)\right)} \dfrac{\hat{S}^{(i)}(t+\Delta_1)}{\hat{S}^{(k)}(t)}\right) pe^t_{ki}}{\sum_{i=1}^{m} \exp\left(h e^{\left(\Delta_1 \hat{r}^{(k)}(t)\right)} \dfrac{\hat{S}^{(i)}(t+\Delta_1)}{\hat{S}^{(k)}(t)}\right) pe^t_{ki}}, \tag{4.2.4}$$

$h \in \mathbb{R}$. It should be noted that this implies that there are an uncountably infinite number of possible risk-neutral measures. The impact and choice of the Esscher parameter h will be discussed with numerical examples in Section 5.4.

4.2.1 Surrender Scheme

There are various ways to consider the surrender behaviour. One reasonable way is to assume that the policyholder is rational and will surrender optimally. Grosen and Jørgensen (1997) do exactly this when they show that a product with a revenue guarantee is essentially equivalent to an American option and can hence be calculated through that principle. Grosen and Jørgensen (2000) take this idea a step further by separating an equity linked contract into 3 components. The risk-free bonds, bonus option and surrender option components are then priced separately using Monte Carlo simulation. Bacinello (2004) calculates the price of a contract where the surrender value is calculated endogenously. In all these scenarios the valuation is done while assuming that the policyholder is rational. While this method is theoretically sound it leads to high capital requirements due to the necessary rationality assumption. Note that as shown in Section 4.2 using the stochastic mesh, the contract can simply be treated as an option. By adding mortality into the valuation process, $P(0)$ can then be considered as the total initial capital required for this contract.

While rational behaviour is a reasonable assumption, there are many arguments against it. The behavioural finance field studies these particular issues. The implications of the existence of irrational investors are rather impactful since many fundamental financial theories such as the efficient market hypothesis rely heavily on this notion. Another important factor to keep in mind is that in the case of equity-linked products, the policyholders are all categorized as individual investors. Barber and Odean (2011) show that this category of investors often under-performs benchmarks through various irrational behaviours. Therefore it is reasonable to assume that a portion of the cohort will not act in an optimal way. Forsyth and Vetzal (2014) examines the effects of irrationality by comparing the hedging costs of a variable annuity under optimal surrender behaviour and sub-optimal surrender behaviour. The sub-optimal behaviour is defined as the surrender occurring when the moneyness ratio is larger than a certain threshold. In the scheme presented in this thesis, the sub-optimal behaviour presented by Forsyth and Vetzal (2014) is extended using the moneyness ratio defined in Equation (4.2.1).

The surrender rate $q_t'^{(s)}$ is set to be a function of this ratio. This idea rests on the hypothesis that policyholders do not all act the same. For example, a portion of the cohort will surrender independently of the market for various reasons such as personal money issues. Another portion of the population will only surrender if the moneyness ratio is high enough. For this reason, the surrender rate $q_t'^{(s)(k)}$, $k \in \{1, \ldots, m\}$ will be a piecewise function (Figure 4.4) that will be a constant ϕ for $MR^{(k)}(t) < 1$ and increasing for $MR^{(k)}(t) \geq 1$ until a particular threshold ψ where the probability of surrender becomes 1. More explicitly,

$$q_t'^{(s)(k)} = \begin{cases} \phi, & MR^{(k)}(t) < 1, \\ \phi + \dfrac{1-\phi}{\psi}\left(MR^{(k)}(t) - 1\right), & 1 \leq MR^{(k)}(t) < \psi + 1, \\ 1, & MR^{(k)}(t) \geq \psi + 1, \end{cases} \tag{4.2.5}$$

for $\phi, \psi > 0$. Note that as $\psi \to 0$ then,

$$q_t'^{(s)(k)} = \begin{cases} \phi, & MR^{(k)}(t) < 1, \\ 1, & MR^{(k)}(t) \geq 1, \end{cases} \tag{4.2.6}$$

which implies that all policyholders will surrender as soon as the moneyness ratio is higher than 1. When ϕ is also set to 0, it implies optimal behaviour from the policyholder.

Similarly as $\psi \to \infty$, the probability of survival is $q_t^{\prime(s)(k)} = \phi$ and does not depend on the moneyness ratio. Also note that the method presented by Forsyth and Vetzal (2014) can be written as,

$$q_t^{\prime(s)(k)} = \begin{cases} 0, & MR^{(k)}(t) < \psi + 1 \\ 1, & MR^{(k)}(t) \geq \psi + 1, \end{cases} \tag{4.2.7}$$

in this framework.

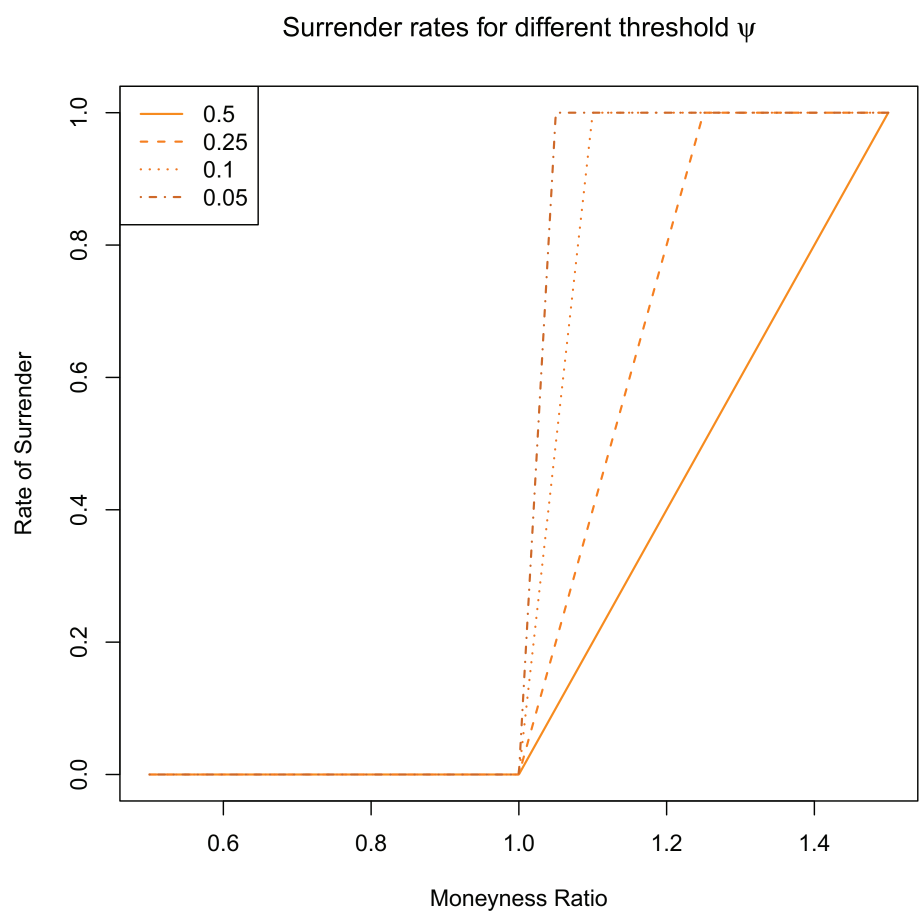

Figure 4.4: Threshold ψ

4.2.2 Surrender and Mortality Path Generation

Once the surrender probabilities for the mesh are found, the surrender and mortality paths can then be generated. Keeping in mind that surrenders occur at the beginning of the year, this can be done by generating $\mathcal{D}_t^{(s)(k)}$ and ${}_{\Delta_1}\mathcal{D}_t^{(d)(k)}$ in two ordered steps, for $k \in \{1, \ldots, m\}$ and $t \in \{0, \Delta_1, \ldots, n - \Delta_1\}$. The generation of these random variables is simplified due to the assumption of a homogeneous cohort in Section 2.1. $\mathcal{D}_t^{(s)(k)}$ follows a binomial distribution with size $\widehat{\mathcal{L}}_t^{(k)}$ and probability of success $q_t'^{(s)(k)}$. Next, ${}_{\Delta_1}\mathcal{D}_t^{(d)(k)}$ follows a binomial distribution with size $\widehat{\mathcal{L}}_t^{(k)} - \widehat{\mathcal{D}}_t^{(s)(k)}$ and probability of success ${}_{\Delta_1}q_{x+t}^{(d)}$. Once $\widehat{\mathcal{D}}_t^{(s)(k)}$ and ${}_{\Delta_1}\widehat{\mathcal{D}}_t^{(d)(k)}$ are generated, then $\widehat{\mathcal{L}}_{t+\Delta_1}^{(k)} = \widehat{\mathcal{L}}_t^{(k)} - \widehat{\mathcal{D}}_t^{(s)(k)} - {}_{\Delta_1}\widehat{\mathcal{D}}_t^{(d)(k)}$. Given that $\mathcal{L}_0^{(k)} = l_0 \; \forall \; k \in \{1, \ldots, m\}$, this process can be repeated sequentially for $t = \{\Delta_1, \ldots, n - \Delta_1\}$ in order to generate mortality for each path.

4.2.3 Surrender and Mortality Mesh

Once the paths are generated, it is then necessary to define the weights for each state. This can be done independently from the economical model weights due to the assumed independence. The weights for this can then be defined by Wm,

$$Wm_{jk}^t = \mathbf{1}_{\left\{\widehat{\mathcal{L}}_t^{(j)} - \widehat{\mathcal{D}}_t^{(s)(j)} \geq {}_{\Delta_1}\widehat{\mathcal{D}}_t^{(d)(k)}\right\}} \binom{\widehat{\mathcal{L}}_t^{(j)} - \widehat{\mathcal{D}}_t^{(s)(j)}}{{}_{\Delta_1}\widehat{\mathcal{D}}_t^{(d)(k)}} \tag{4.2.8}$$

$$\left({}_{\Delta_1}q_{x+t}^{(d)}\right)^{{}_{\Delta_1}\widehat{\mathcal{D}}_t^{(d)(k)}} \left(1 - {}_{\Delta_1}q_{x+t}^{(d)}\right)^{\mathcal{L}_t^{(j)} - \widehat{\mathcal{D}}_t^{(s)(j)} - {}_{\Delta_1}\widehat{\mathcal{D}}_t^{(d)(k)}}, \tag{4.2.9}$$

for $j, k \in \{1, \ldots, m\}$ and $t \in \{0, \Delta_1, \ldots, n - \Delta_1\}$. Unlike the equity weights, these weights are not divided by the probability of reaching cohort state $Y_{t+\Delta_1,k}$ since the value of $\mathcal{D}_t^{(s)(j)}$ is always known at the time of valuation and the lives are considered independently and identically distributed.

4.3 Complete Mesh Weights

In order to use the valuation method presented in Chapter 3, it is then necessary to have a complete model which describes the probabilities of going from state X_{tj} to $X_{t+\Delta_1,k}$. This can be done by putting the weights together,

$$W^t_{jk} = We^t_{jk} \times Wm^t_{jk} \tag{4.3.1}$$

and finally the probabilities for the transition are given by

$$p^t_{jk} = \frac{W^t_{jk}}{W^t_{j\bullet}}, \tag{4.3.2}$$

for $j, k \in \{1, \ldots, m\}$ and $t \in \{0, \Delta_1, \ldots, n - \Delta_1\}$.

In summary, in order to achieve these weights several steps were needed. First the economic model needs to be simulated and from there the economic mesh weights can be found. When this is done the moneyness ratio is found for every node and the mortality/surrender mesh can be simulated. Once this last step is done the mortality weigths needed can be found which completes the mesh weights given above.

Chapter 5

Numerical Example

5.1 Equity-Linked Products

Equity-linked products appeal to investors because they offer the same protection as conventional annuities by limiting the financial risks, but are also linked to the performance of an equity market. A variety of these products currently exist on the market such as variable annuities, equity-indexed annuities, universal life insurance and variable universal life insurance.

When concluding the contract, the insured may usually opt for optional guarantees, such as *Guaranteed Minimum Death Benefits* (GMDB) as well as *Guaranteed Minimum Living Benefits* (GMLB). The risk profile of the investor is set by the specific selection of mutual funds.

The guaranteed minimum death benefits (GMDB) consists of a death benefit that is payable if the insured were to die during the deferment period. The simplest form is the *Return of Premium Death Benefit* where the maximum of the current account value at the time of death and the single premium is paid.

The guaranteed minimum living benefits (GMLB) are separated into three types, *Guaranteed Minimum Accumulation Benefits* (GMAB), *Guaranteed Minimum Income Benefits* (GMIB) and more recently *Guaranteed Minimum Withdrawal Benefit* (GMWB).

GMAB is the simplest form the these benefits, where the insured is entitled to the single premium or a roll-up benefit base at maturity. The roll-up benefit base is defined by Bauer et al. (2008) as the theoretical value of the compounded single premium with a constant interest rate, namely the *roll-up rate*. The GMIB offers the choice to obtain the account value, annuitize the account value or annuitize some guaranteed amount at specified rates. The GMWB offers the possibility to withdraw a certain amount in small portions annually. The focus will be on GMDB and GMAB guarantees, where the annual capital requirements may be obtained using the framework presented in this paper.

To illustrate an equity-linked product valuation, the simplest design of EIAs is used, known as the point-to-point with term-end design where the index growth is based on the growth between two time points over the entire term of the annuity. This design has embedded GMDB and GMAB guarantees with the payoff at time t represented by

$$D(k) = \max\left[\min\left[1 + \gamma R(t), (1+\zeta)^t\right], \beta(1+g)^t\right], \tag{5.1.1}$$

for $t \in \{\Delta_1, \ldots, n\}$ with an embedded surrender option with value $(1 - C_t)D(t)$, where γ represents the participation in the index and C_t is the surrender fee at time t. The "gain" $R(t)$ is defined by

$$R(t) = \frac{S(t)}{S(0)} - 1. \tag{5.1.2}$$

EIAs provide a protection against the loss from a down market $\beta(1+g)^t$. The cap rate $(1+\zeta)^t$ reduces the cost of such a contract since it imposes an upper bound on the maximum return.

Financial options embedded in equity-indexed annuity contracts are usually dependent on a set of parameters. These include the participation rate, the minimum guaranteed rate, the guaranteed fraction of premium, etc. Therefore, the premium of the contract is a function of this set of parameters.

To study the effects and side effects of surrender on the initial capital requirement, a numerical example for the product presented above will be used. That is, a point-to-point EIA where the index and interest rate is governed by the models presented in Chapter

1. It will be assumed that the index and interest rates make weekly movements and the partial hedging portfolio is rebalanced monthly according to the valuation method in Chapter 3. The initial capital requirements will be calculated for a homogeneous cohort with 5 year contracts with investment of 1. The cohort is assumed to all be of age 50 at time of issue and their mortality follows that of the illustrative table in Bowers (1997). The initial capital requirement is normalized with respect to the size of the cohort (i.e. divided by the size).

Since the valuation using the model presented relies on simulation, the confidence interval for each point estimate will also be presented in the plot. However it is important to note that the confidence interval presented is a point-wise confidence interval and not a simultaneous one.

5.2 Economical Model Effects

Initially the effects of the model parameters on a typical EIA product are inspected. This is done under the assumption of no surrenders. The EIA contract is fixed with a guarantee of 100% of the premium ($\beta = 1$) with a guaranteed annual return of 1% ($g = .01$), the participation rate is fixed to 60% ($\gamma = .6$) and there is no cap ($\zeta = \infty$). The valuation is done using CVaR at the 95% level for a cohort of 1 policyholder.

The first parameter that is altered is the initial interest rate $r(0)$. This is done for the fitted model and for an alternate model which assumes an independent and constant interest rate equal to the initial interest rate. This is done for $r(0) \in \{.5\%, 1\%, 2\%, 4\%\}$. The initial capital requirement is calculated using the average of 10 initial capital requirements each using an independently simulated mesh with 1,000 paths ($m = 1000$).

As expected, in Figure 5.1, it can be seen that a higher initial interest rate does in fact lower the hedging costs. This is expected since with higher interest rates the insurance company can invest larger proportions in the risk free market and therefore can reduce the CVaR. The second observation is that the effect of initial interest rate on the initial

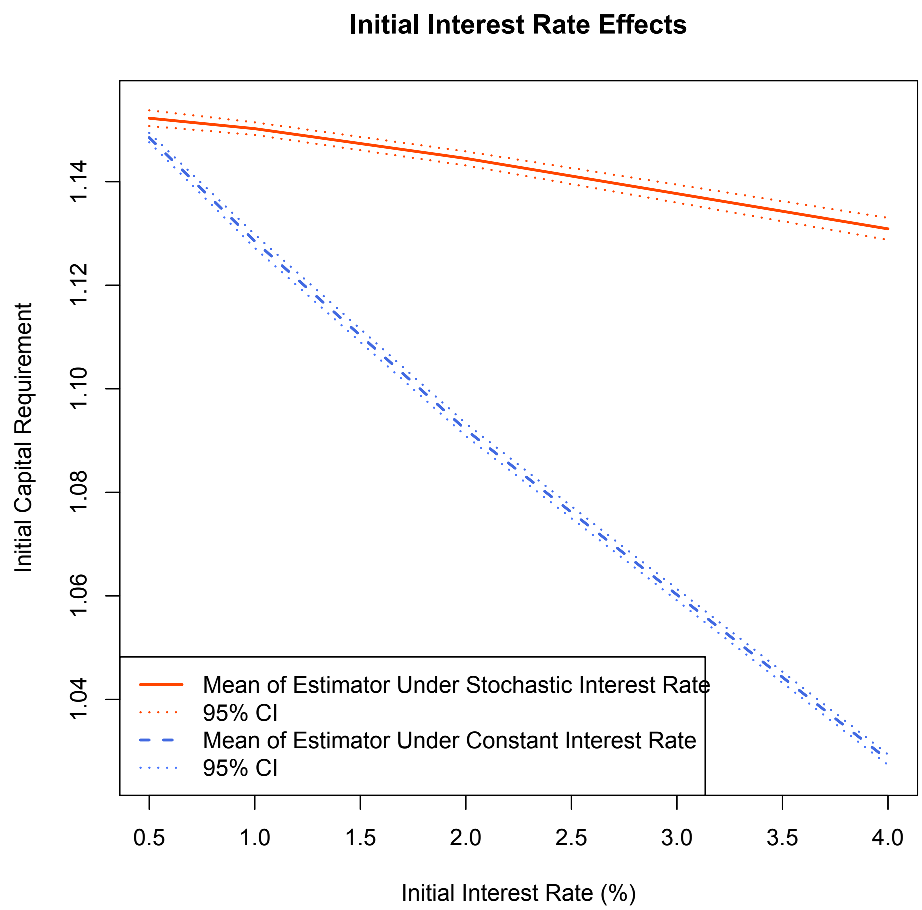

Figure 5.1: Initial Interest Rate Impact

capital requirement is mostly linear, that is, for every percent increase in the initial interest rate the initial capital requirement is reduced by approximately 0.0071 under stochastic interest rate and by 0.038 under independent constant interest rate. These values are found through the use of regression on the simulated initial capital requirement.

One of the reasons why the initial interest rate appears to have less of an impact on the initial capital requirement under the stochastic interest rate can be attributed directly to the model. The model that was presented in Section 1.1 has the property of mean reversion. Due to this property even when starting with different initial interest rates, the process will revert to the model mean. This means that over time the impact of the initial interest rate will be reduced which is not the case in the constant interest rate scenario. Finally it can be observed that the initial capital requirement under stochastic and dependent interest rate is higher than that under independent and constant interest rate. This increase can again be attributed to the mean-reverting nature of the interest rate model. As it can be seen in Figure 5.1 for small initial interest rates the initial capital requirement is close for both schemes. Then noticing the fitted parameters for the second regime in Section 1.1.2, the high mean-reverting regime has a rather low mean which implies that the process spends most of its time at lower interest rates. These lower interest rates increase the hedging costs. The numerical data can be seen in Table B.1 in Appendix B.

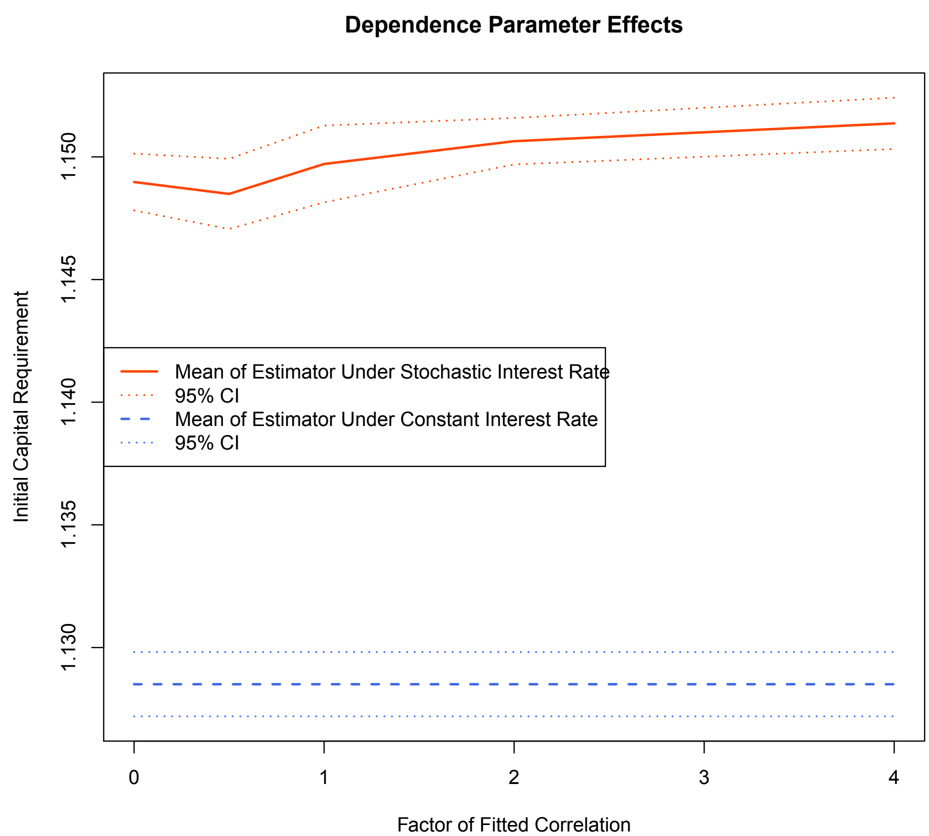

Figure 5.2: Dependence Parameter Impact

Next the dependence is altered using the ρ parameters which are the dependence correlations. The fitted parameters $\hat{\rho}_i$ $i \in \{1, 2\}$ are considered as a baseline. Multiples of the baseline: $0\hat{\rho}_i$, $0.5\hat{\rho}_i$, $1\hat{\rho}_i$, $2\hat{\rho}_i$, and $4\hat{\rho}_i$ are considered while the initial interest rate is fixed to 1%. In Figure 5.2, it can be seen that the stochastic interest rate still produces more expensive portfolios than under a constant independent interest rate whether there is dependence or not. It can be also be noted that increasing the dependence from a factor of 0.5 and beyond increases the initial capital requirement. Keeping in mind that the fitted copula parameters implied a positive correlation between the two process, the increase may be attributed to the fact that the two models will move more strongly together which means higher interest rates will happen alongside higher equity returns. While higher interest rates imply lower hedging costs as seen in Figure 5.1, higher equity returns imply higher payoffs and hence a higher hedging cost. From the plot, stronger dependence implies higher initial capital requirements and it can then be concluded that the higher returns have a larger impact. Note that the type of product that is being evaluated is bounded from below by the guarantee. Therefore the interest rate and equity returns moving downward together would both imply a higher hedging cost. The numerical data can be seen in Table B.2 in Appendix B.

5.3 EIA Parameter Effects

Typically EIA's are priced through their parameters, usually either the spread, participation rate (γ), or cap (ζ). The contracts will be separated into two categories, one which returns a minimum of 90% ($\beta = 0.9$) of initial investment plus guaranteed return (g) and another with a minimum of 100% ($\beta = 1$). These will be calculated for a cohort of 1 with no surrender option and with an initial interest rate of 1% ($r(0) = 0.01$). The simulation parameters are same as in Section 5.2.

Looking at Figures 5.3 and 5.4 it can be seen that the effect of these parameters are linear. As expected it can also be seen that given β, an increase in the guaranteed return causes an increase in the initial capital requirement. However it can also be seen that the increase is more pronounced for contracts with $\beta = 100\%$ than for contracts with

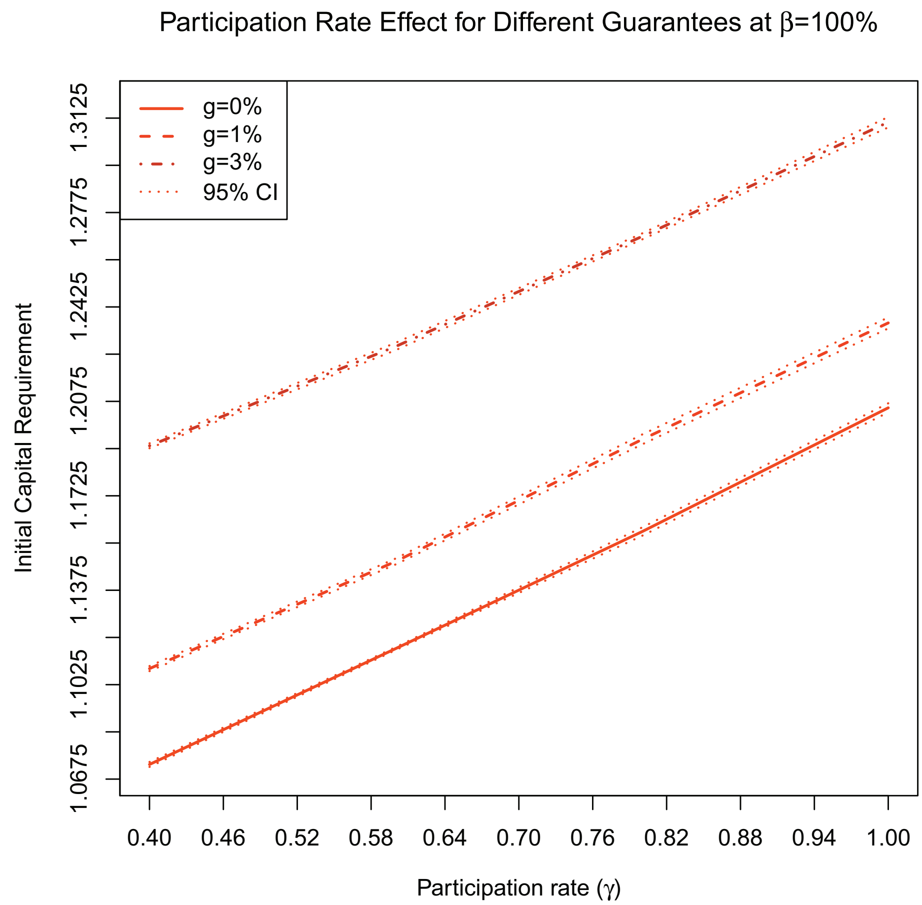

Figure 5.3: Participation Rate Effects @ $\beta = 1$, 95% CVaR

$\beta = 90\%$.

Using least-squares regression the slopes can be found and compared. In Figure 5.5 an interesting behaviour can be seen. In the case where 100% of the investment is guaranteed, the effect of the participation rate appears to be reduced as the guaranteed rates are increased. However in the case where 90% of the investment is guaranteed, the effect of the participation rate appears to be increased as the guarantees are increased. The change of behaviour may be explained by a change in leverage.

Note that due to the linearity of the effects, the fair value participation rate can be found by simply interpolating from the linear regression given that we have data for contracts which the initial capital requirements are below 1.

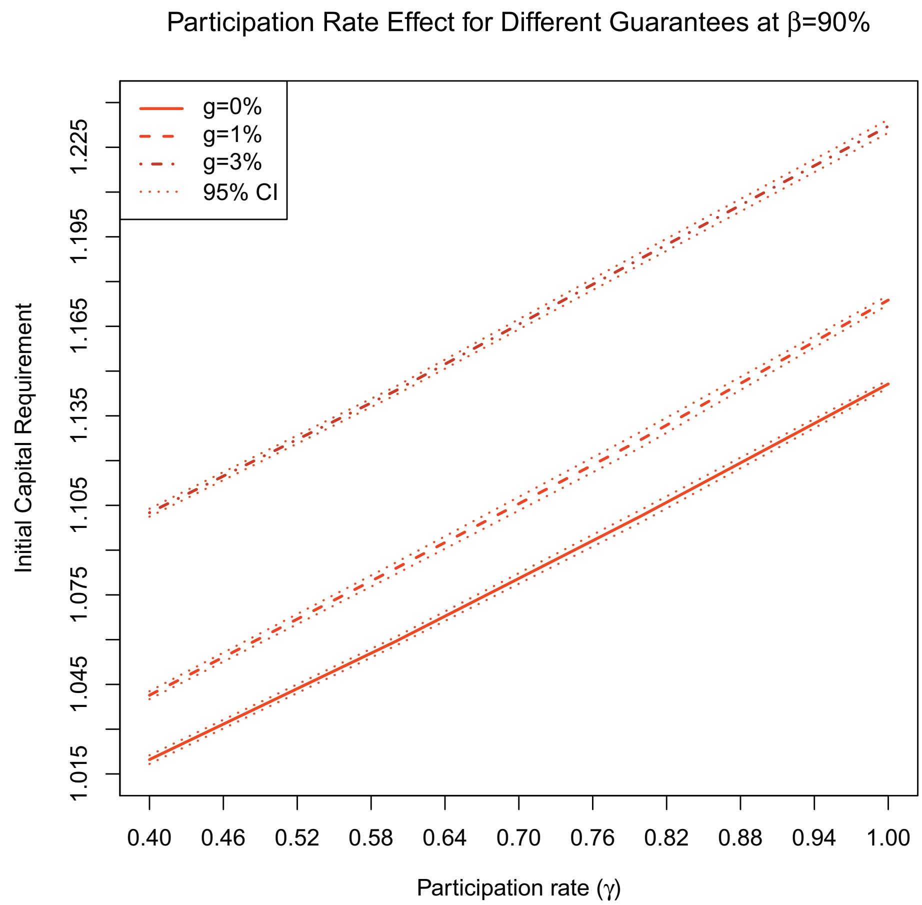

Figure 5.4: Participation Rate Effects @ $\beta = 0.9$, 95% CVaR

The impact of the cap parameter can be seen in Figures 5.6 and 5.7. In contrast to the participation rate, the cap does not have a linear effect. In fact it can be seen that the initial capital requirement is sensitive to low caps. However as the cap increases past 15% this sensitivity is gone, which is due to the fact that the index rarely makes gains of more than 15% per year. As with the participation rate it can be seen that for a given β, an increase in the guaranteed rate of return (g) causes an increase in the initial capital requirements. Similarly, for the same guaranteed rate of return, an increase in β causes an increase in the initial capital requirements as well. The numerical data can be seen in Tables B.4 and B.5 in Appendix B.

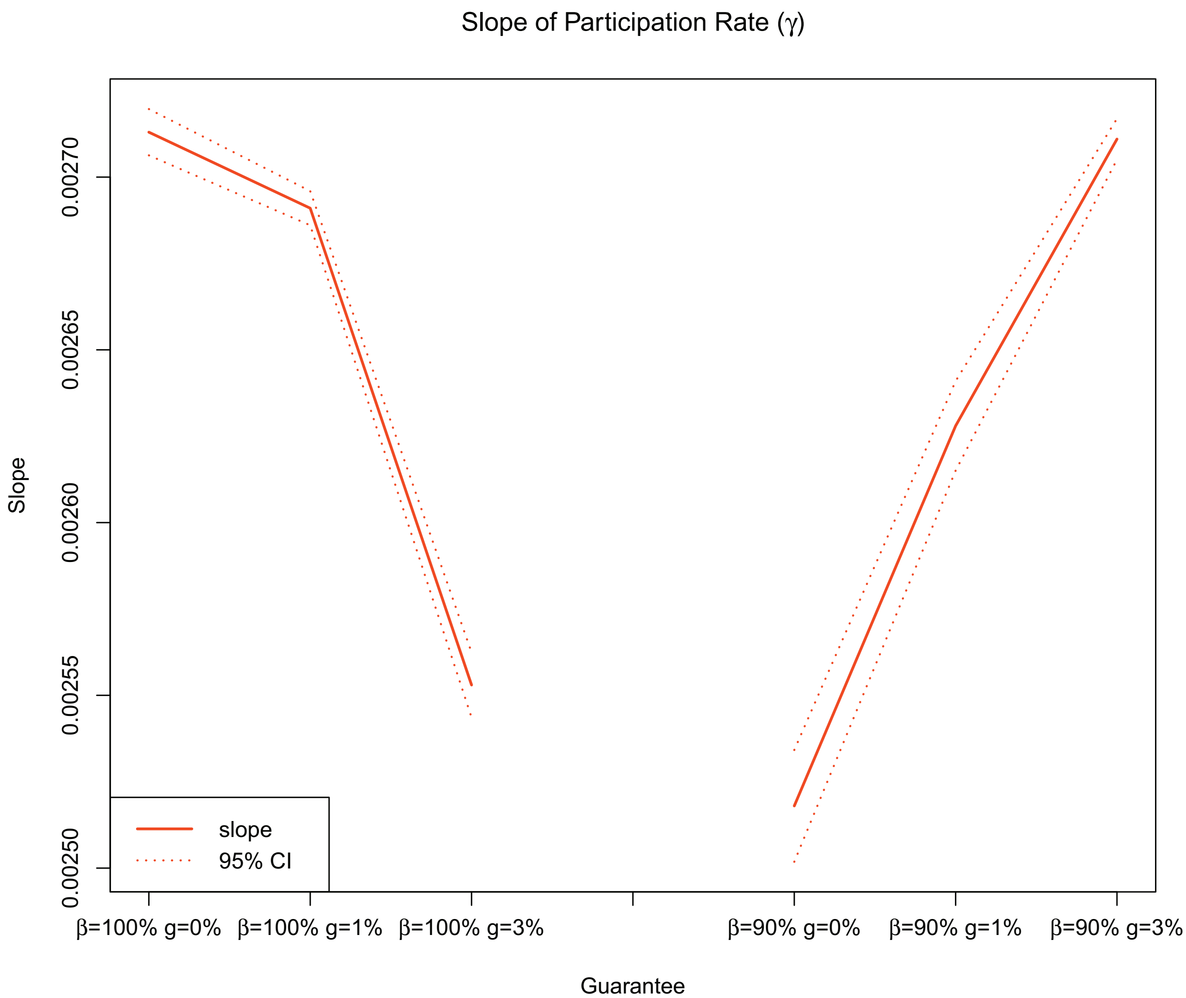

Figure 5.5: Participation Rate Effects

5.4 Esscher Parameter Effects

Before considering a surrender option it is necessary to pick a meaningful parameter for the Esscher transform that is used to find the moneyness ratio presented in Section 4.2. To evaluate the impact of the Esscher parameter on the initial hedging costs, a contract similar to 5.2 will be used. In addition, the initial interest rate will be set to 1%. Also it will be assumed that the policyholder is rational (i.e. $\psi = 0$ and $\phi = 0$) and the contract has a constant surrender fee of 1%.

It is important to first note that the Esscher transform essentially skews the distri-

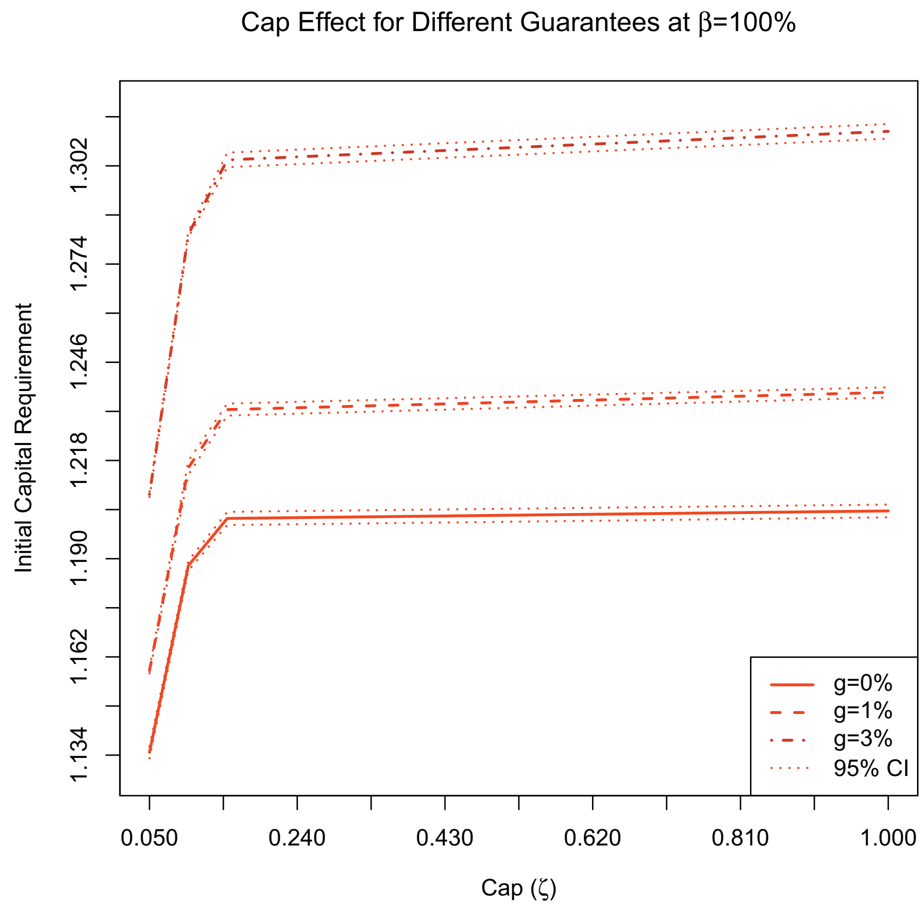

Figure 5.6: Cap Effects @ $\beta = 1$, 95% CVaR

bution. In particular, when $h < 0$ the distribution is skewed in such a way that lower valued outcomes have a higher density and when $h > 0$ the distribution is skewed in such a way that higher valued outcomes have a higher density. In Figure 5.8 this impact can be seen particularly for the initial value of the Bermudian option. Since a higher h implies that higher returns are more probable, it is then expected that the initial replicating cost increases with respect to h. Due to this introduced skewness it can also be deduced that the continuation values under a higher h will be higher than those calculated with a smaller h.

This impact is not as great in the initial capital requirements using the iterative risk measures (IRM) method and is in fact very subtle for $h > -5$. Recall that h is used to

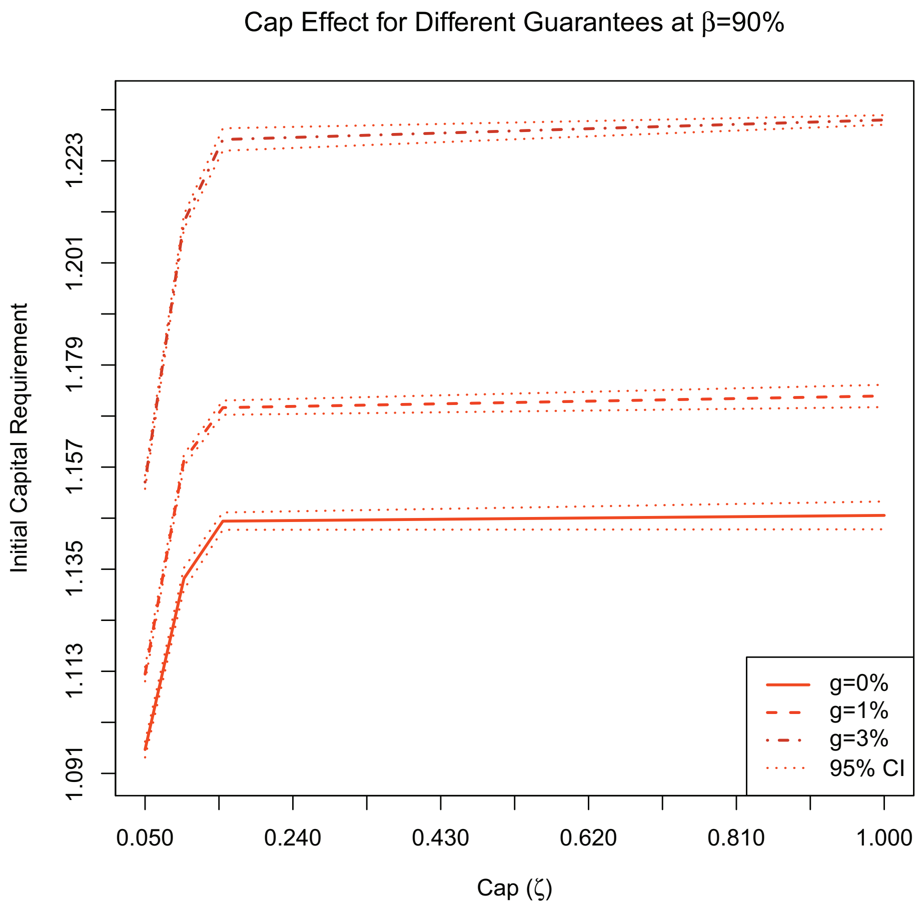

Figure 5.7: Cap Effects @ $\beta = .9$, 95% CVaR

calculate the continuation value which in return is used to determine the moneyness ratio at a particular node. This ratio is then used to determine the probability of surrender which in this case is either 0 or 1 depending on whether the moneyness ratio is larger or smaller than 1. As previously stated due to the impact of h on the distribution, a higher h would imply lower surrender opportunities since staying in contract is beneficial due to the increased expected return. Conventional wisdom would imply that more frequent rational surrenders would drive up the hedging costs, therefore it is not expected that the initial capital requirements increase with respect to h. There are two factors to take into account in this scenario.

The first is that at each node the costs calculated for the Bermudian option and the

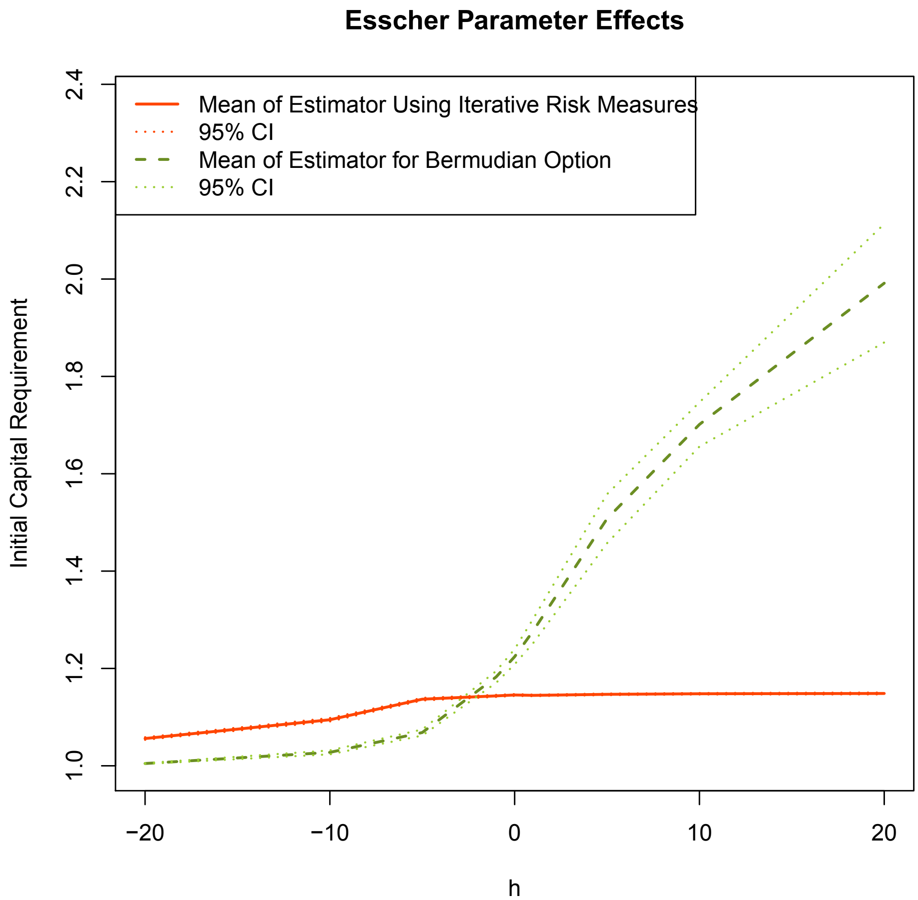

Figure 5.8: Esscher Parameter Impact

EIA using IRM represent fundamentally different things. In the case of the Bermudian option it is the replication cost whereas under the IRM method it is the replication cost of the worst 5%. Therefore it is completely possible that even though the policyholder rationally surrenders, the surrender value is lower than the cost of hedging the worst 5% future outcomes, hence the act of surrendering eliminates the risk. This behaviour will be more thoroughly addressed in Section 5.5.

The second factor to take into account is that since a higher h induces higher possible returns, the fact that a policyholder rationally surrenders under a high h may not necessarily translate to it being worthwhile under the physical measure which is used to calculate the initial capital requirements under the IRM method.

The final conclusion here is that the choice of h has a minimal impact on the initial capital requirements under the IRM method compared to the Bermudian option costs. The numerical data can be seen in Table B.3 in Appendix B.

5.5 Surrender Effects

To investigate the effect of surrenders on the framework, the contract parameters will be fixed similarly to Section 5.2 with the addition of the initial interest rate being 1%. As shown in Section 5.4, the initial capital requirement for the partial hedging strategy calculated using iterative risk measures is rather robust in relation to the Esscher parameter. For this chapter, the Esscher parameter will be fixed at $h = -2.5$ as at this value the Bermudian option value is equal to the initial capital requirement under this particular contract.

Initially, the effect of the surrender fee will be looked at under different surrender scenarios. In Figure 5.9 the moneyness ratio threshold is set to 0 ($\psi = 0$), that is as soon as the moneyness ratio (MR) is above 1, the probability of surrender is 1. This is calculated for different values of non-rational surrender rates (ϕ). Note that here the rates are stated as annual rates. As one would expect, it can be observed that as the fee increases from 0% to 5%, the initial capital requirements (ICR) decrease for any given non-rational surrender rate. It can also be observed that with a fee of 5% as the non-rational surrender rate increases, the ICR decreases. Since the threshold is set to 0 every other surrender that will occur due to the non-rational surrender rate will be at a time where $MR < 1$. This act would in fact be beneficial to the insurance company due to the non-zero surrender fee. In the case where the surrender fee is 0%, while one would still expect a decrease in the ICR, this decrease is not as large.

In order to better understand this result, the effect of non-rational surrender is assumed to be constant and through linear regression, the slope and the associated significance is found. This is done for various thresholds in Table 5.1. It can be seen that as the surrender

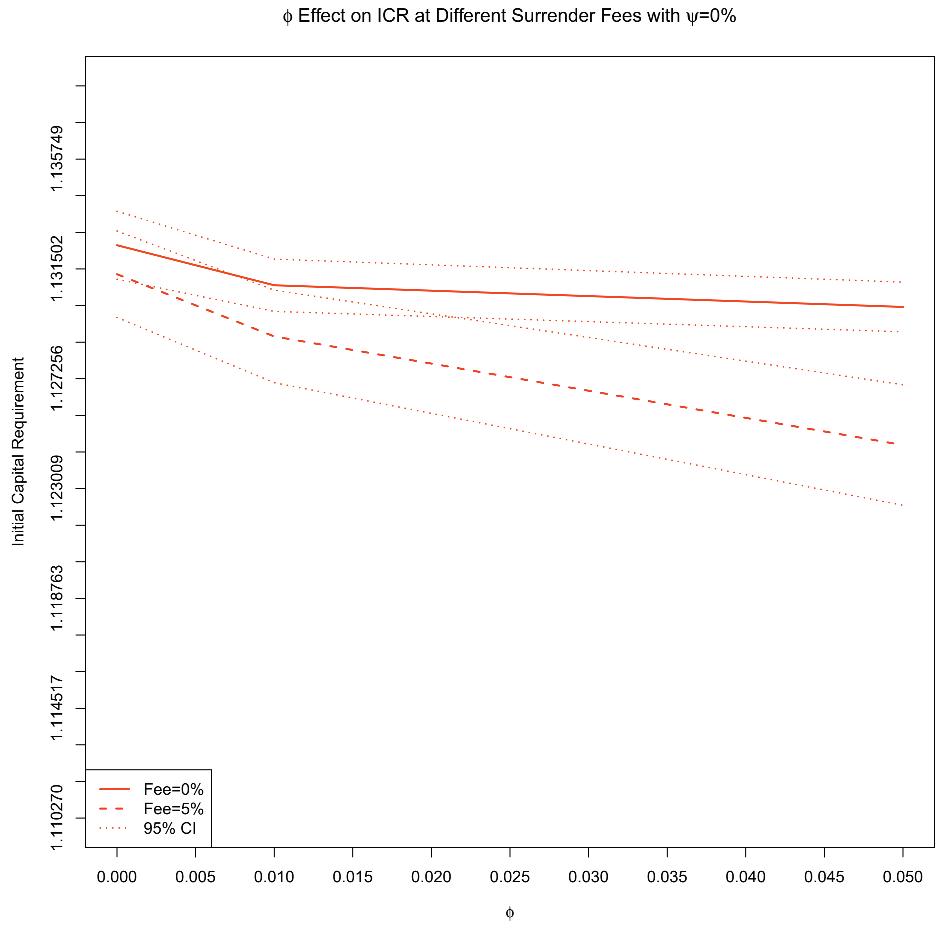

Figure 5.9: Effect of Surrender Fees

fee increases, the slope consistently decreases for all thresholds. This implies that non-rational surrenders have a greater impact when the fees are higher. As previously stated this is intuitive since a non-rational surrender occurring when $MR < 1$ would be beneficial to the insurance company. This effect is then augmented by higher fees since this would make the cases where $MR < 1$ more common.

It can also be noted that as the threshold (ψ) increases, the effect of non-rational surrenders (ϕ) is reduced in the presence of a 5% surrender fee. This may be explained by comparing the two extremes. When $\psi = 0$, policyholders react optimally and hence every other surrender occurring due to ϕ will be non-optimal which would reduce the price. When $\psi = \infty$, no optimal surrender is occuring which means that a portion of the non-optimal surrenders may in fact happen to be optimal simply by chance. This would then cause the effect of non-rational surrenders ϕ to be dampened.

Table 5.1: Non-Rational Surrender Effect

ψ (%)	Fee (%)	Slope	P-Value
0	0	-0.040	1.02E-01
	5	-0.124	1.27E-05
.1	0	-0.064	4.63E-05
	5	-0.119	7.99E-06
1	0	-0.065	2.09E-04
	5	-0.090	2.08E-05
∞	0	-.038	3.43E-02
	5	-0.090	1.24E-05

Another interesting dynamic to observe is the effect of the surrender threshold ψ on the initial capital requirement at different non-rational surrender rate levels ϕ. In Figures 5.10 and 5.11 these are calculated for a contract with a 5% surrender fee. The first observable feature is that non-rational surrenders decrease the ICR at any level of ψ when comparing $\phi = 0\%$ and $\phi = 5\%$. This doesn't appear to be uniformly true for $\phi = 0\%$ and $\phi = 1\%$, however this may simply be due to the fact that the difference are small and the confidence intervals wide. This is due to the fact that non-rational surrenders may occur during non-optimal times, in fact the odds of this occuring at a non-optimal time is higher partly due to the surrender fee (see Figure 5.13).

The most noteworthy feature however, is that as the threshold increases the initial capital requirement does so as well. This behaviour was partially observed in Section 5.4. This appears counter-intuitive since a higher threshold ψ implies that most people will wait until a higher moneyness ratio is reached before surrendering and hence at lower moneyness ratio there are less people surrendering their contracts while it may still be

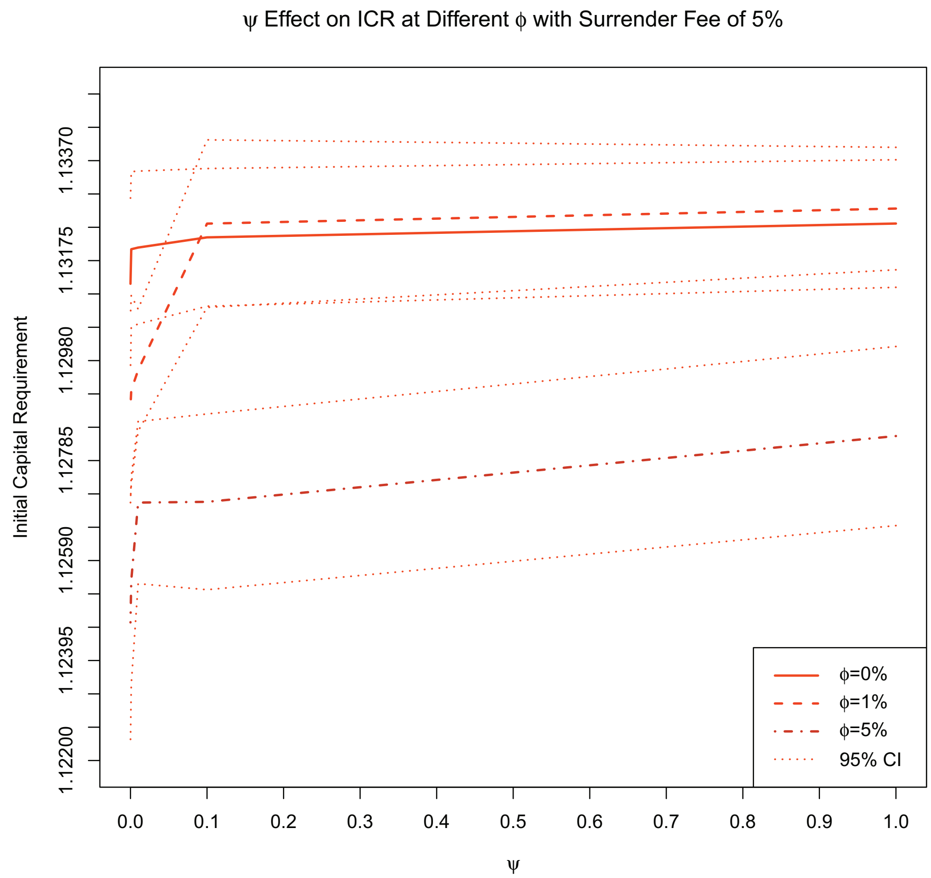

Figure 5.10: Threshold Effects on Initial Capital Requirements

optimal. A fully rational policyholder would in fact surrender as soon as MR is larger than 1. In a pricing scenario where the whole contract is being hedged this should cause an increase in the cost. However because in the presented valuation method only the worst portion of the contract is hedged, a surrender, even when optimal, will decrease the risk. In Figure 5.12 a histogram of the future payoffs can be seen. In this particular case the moneyness ratio is at 1.0505, therefore a surrender is highly optimal. It can be observed that indeed the surrender value is higher than the expected discounted payoff if the policyholder were to stay. However the 95% CVaR is still considerably higher than the expectation. This difference makes it so that a surrender becomes beneficial even under optimal conditions.

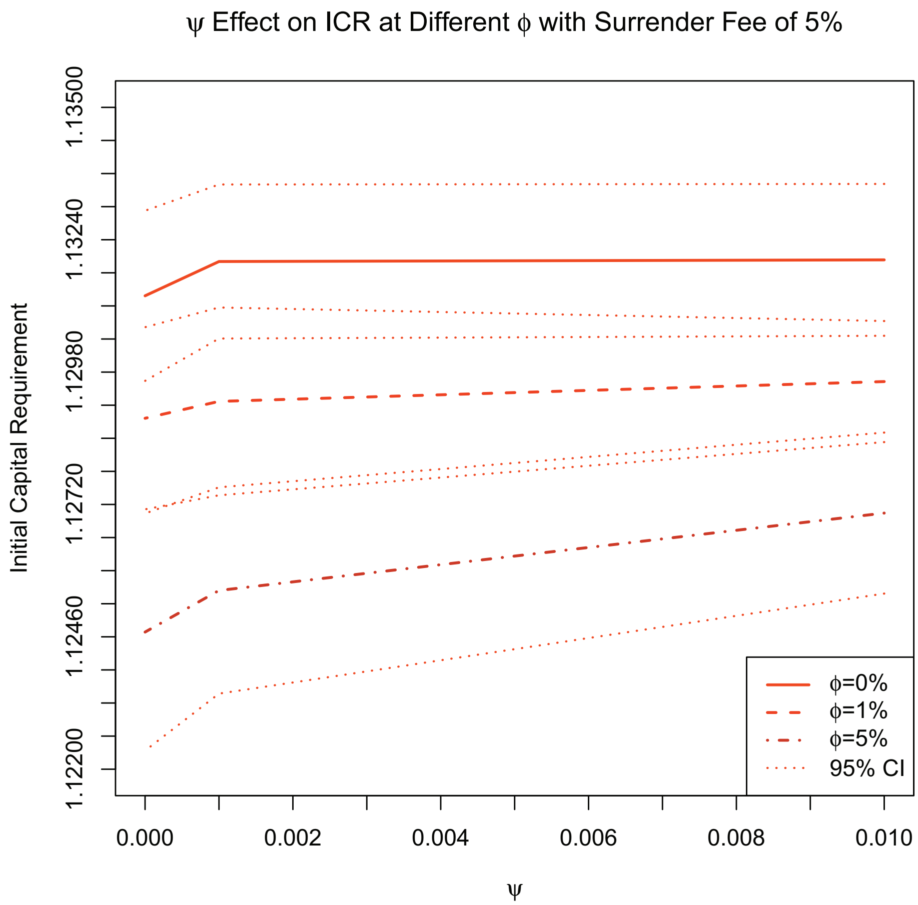

Figure 5.11: Threshold Effects on Initial Capital Requirements

The other noteworthy feature is the shape of the curve with respect to the threshold. One can observe that the rate of increase is very high between thresholds of 0 and 10%. This can be explained by the fact that given a moneyness ratio higher than 1, the density of these ratios in the mesh are concentrated very close to 1. This can be seen in Figures 5.13 and 5.14. In fact, about 80% of these ratios are less than 1.005. This implies that the initial capital requirement for the partial hedge is very sensitive to the threshold parameter for low thresholds. This then explains the shape of Figure 5.10 since at low thresholds the differences between the odds of surrender opportunities changes most.

As explained in Section 4.2.1, the presented method of surrender is an extension on the method presented by Forsyth and Vetzal (2014). They present a binary scenario

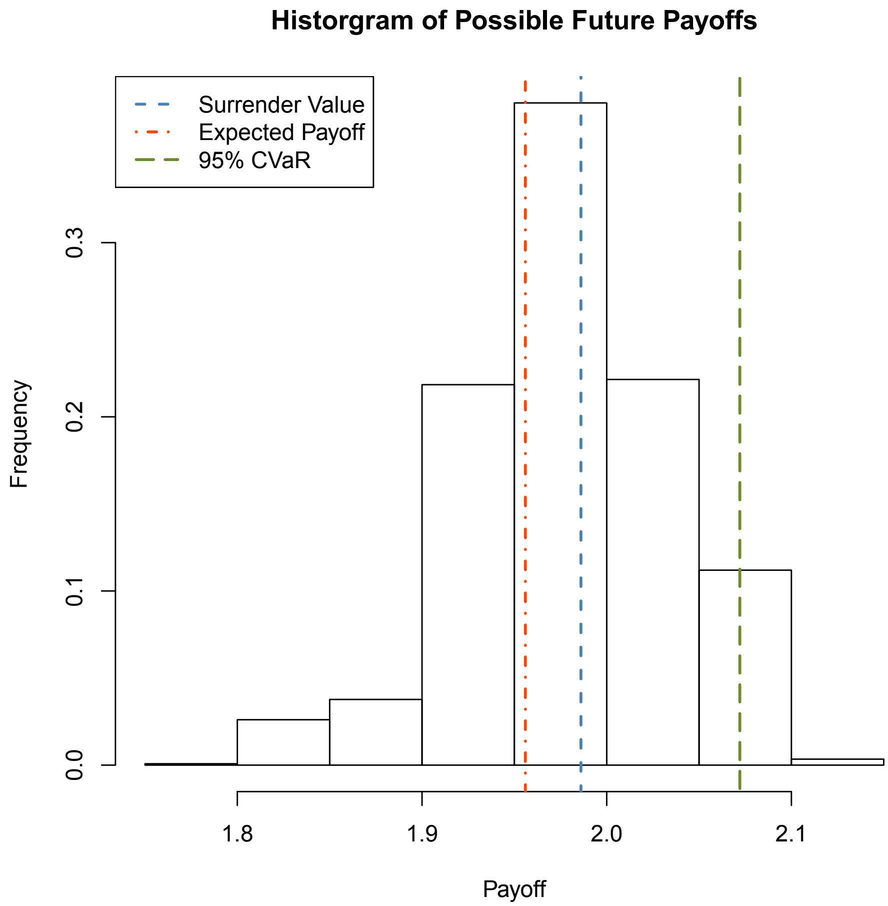

Figure 5.12: Histogram of Future Payoffs

where 100% of the policyholders surrender at a moneyness ratio of larger than $1+\psi$ and no one surrenders at lower ratios. In the method presented in this thesis, between the moneyness ratio of 1 and $1+\psi$ there is a probability of surrender. In the latter method one would expect a higher number of surrenders for a given ψ. This behaviour can be observed in Figure 5.15 and 5.16. In this particular case the surrender fee is set to 5% and the non-rational surrender rate is set to $\phi = 0\%$. Note that at a threshold of 0 ($\psi = 0$), both methods are equivalent. However as the threshold increases, the ICR under the Forsyth method quickly increases and converges whereas under the presented method this convergence is slower. Due to the non-binary nature of the presented surrender curve, between a moneyness ratio of 1 and $1+\psi$ there is an increasing portion of surrenders.

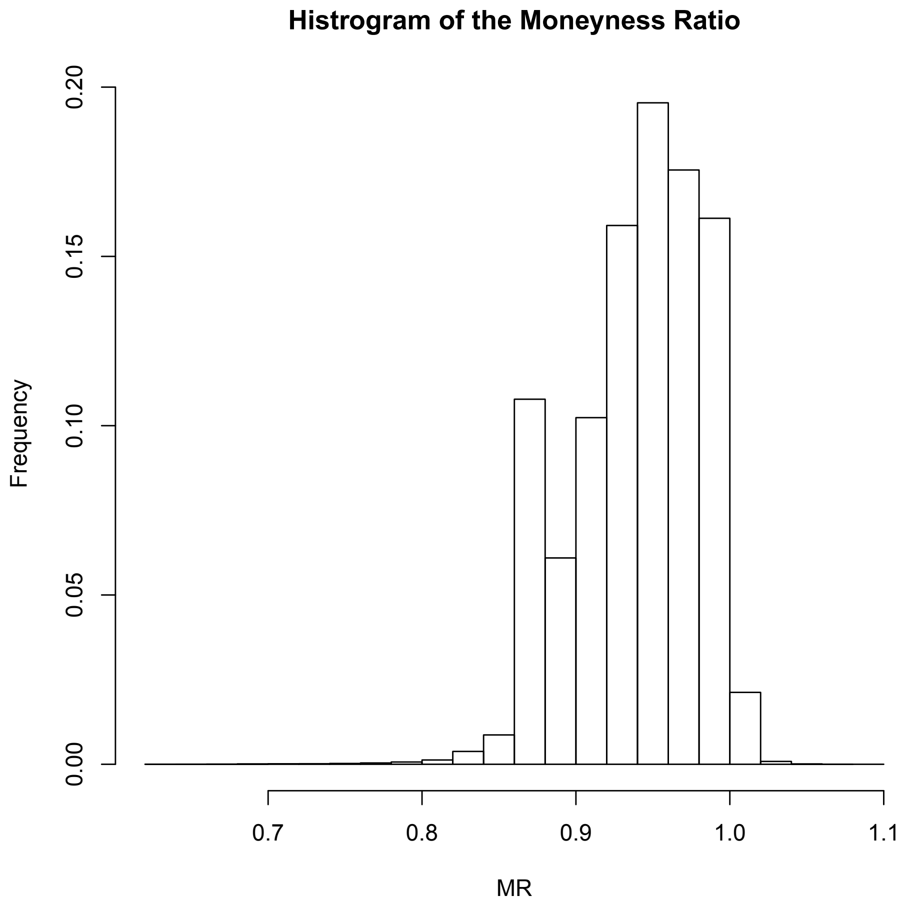

Figure 5.13: Histogram of Moneyness Ratios a Mesh

This then causes more surrenders to occur on average than under the binary Forsyth curve and as previously stated, under the presented valuation method, surrenders decrease the initial capital requirements. The difference in initial capital requirement is not visibly significant in this scenario but in Figure 5.16 it can be seen that at a threshold of $\psi = .01$ this difference is in fact statistically significant. The numerical data can be seen in Table B.6 and B.7 in Appendix B.

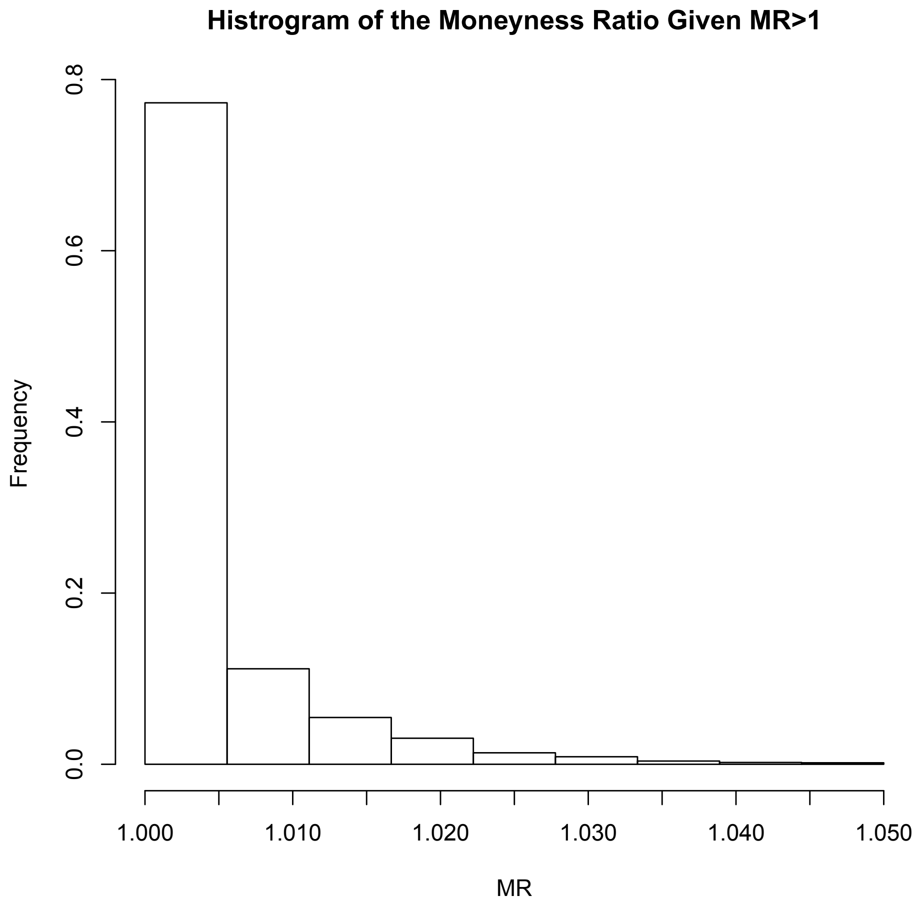

Figure 5.14: Histogram of Moneyness Ratios a Mesh Given $MR > 1$

5.6 Cohort Size Effects

The valuation method presented in Chapter 3 makes use of the CVaR risk measure. As previously stated one of the nicer features of this risk measure is sub-additivity. That is, the CVaR of the sum of random variables is less than the sum of the CVaR of the same random variables. Due to this feature one would expect that adding individuals to the cohort would in fact result in a reduction of tail risk and hence a reduction in initial capital requirements. Figure 5.17 confirms this expectation. Note that it can be seen that increasing the cohort size uniformly decreases the initial capital requirement for the partial hedging strategy across different thresholds ψ. In this plot, the initial

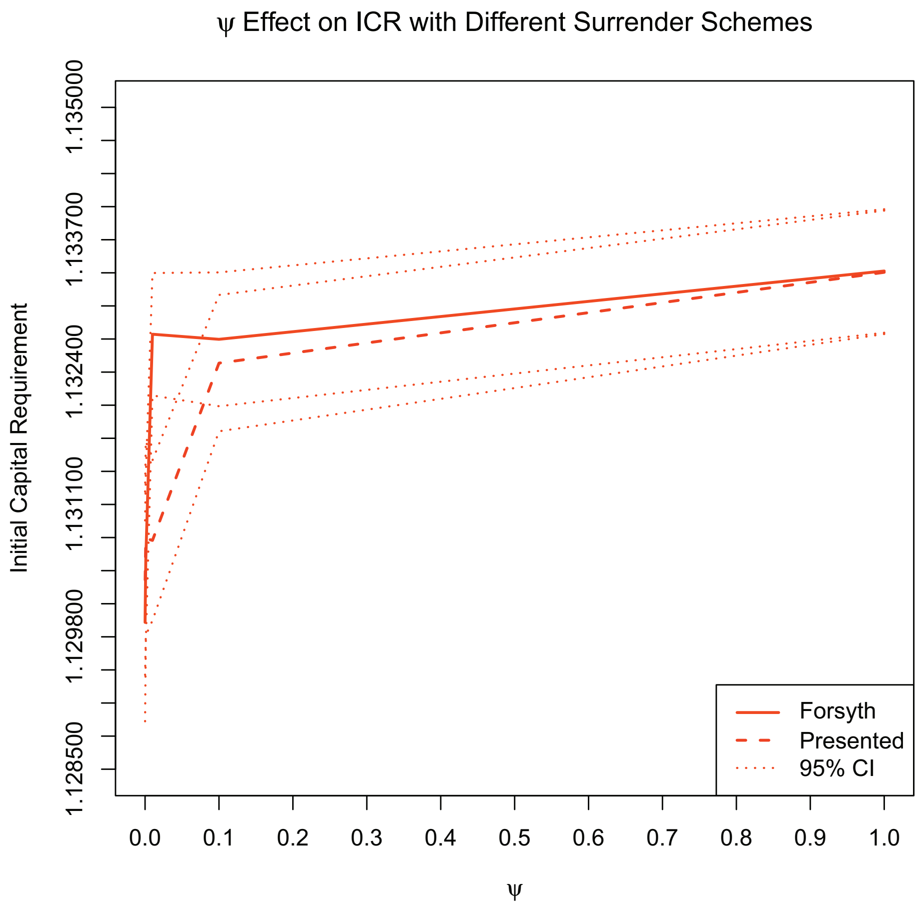

Figure 5.15: Surrender Methods

capital requirements are calculated for a contract similar 5.5 while keeping ϕ at 1% and the surrender fee at 5%. The numerical data can be seen in Table B.8 in Appendix B.

5.7 Programming Note

The stochastic mesh presented in Chapter 4 makes it possible to have a meaningful and rather realistic numerical example. However some notes should be made about the use of stochastic mesh with the pricing framework presented in Chapter 3 and the model presented in Chapter 1. The numerical example was programmed in *C/C++*. This decision was made since neither *R* nor *MATLAB* are particularly efficient when it comes

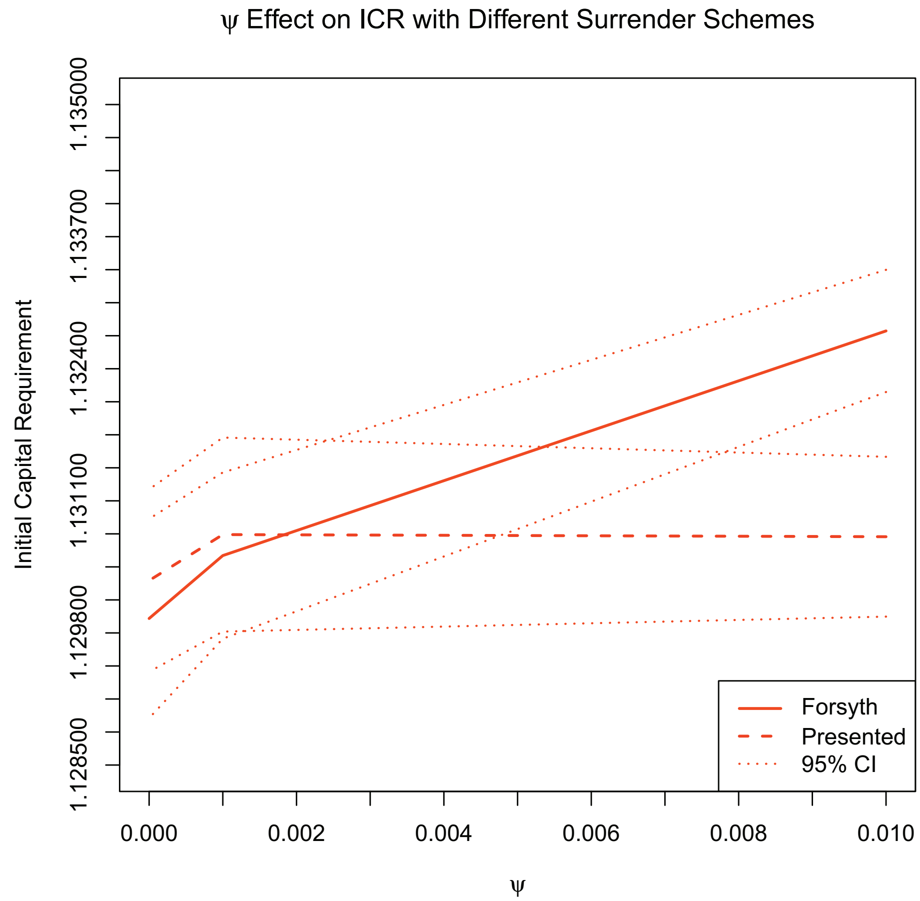

Figure 5.16: Surrender Methods

to concurrency and loops. In terms of runtime, a single run for a 5 year contract rebalanced monthly with $m = 1,000$ takes anywhere between 12 to 14 hours using *MATLAB*. Using *C/C++* this same run takes between 3-4 minutes. Both these times are for a machine running an Intel Core i5-2500k running at 4.28GHz. While it can be argued that using faster matrix based algorithms in *MATLAB* would reduce the runtime, the same can be said of *C/C++*.

As previously stated the stochastic mesh relies on the transition density of the model. While this is typically not an issue, the transition density of the CIR model, namely the non-central χ^2 distribution, is not available in closed form. This distribution relies on the modified Bessel function of the first kind, the implementation that is used is the

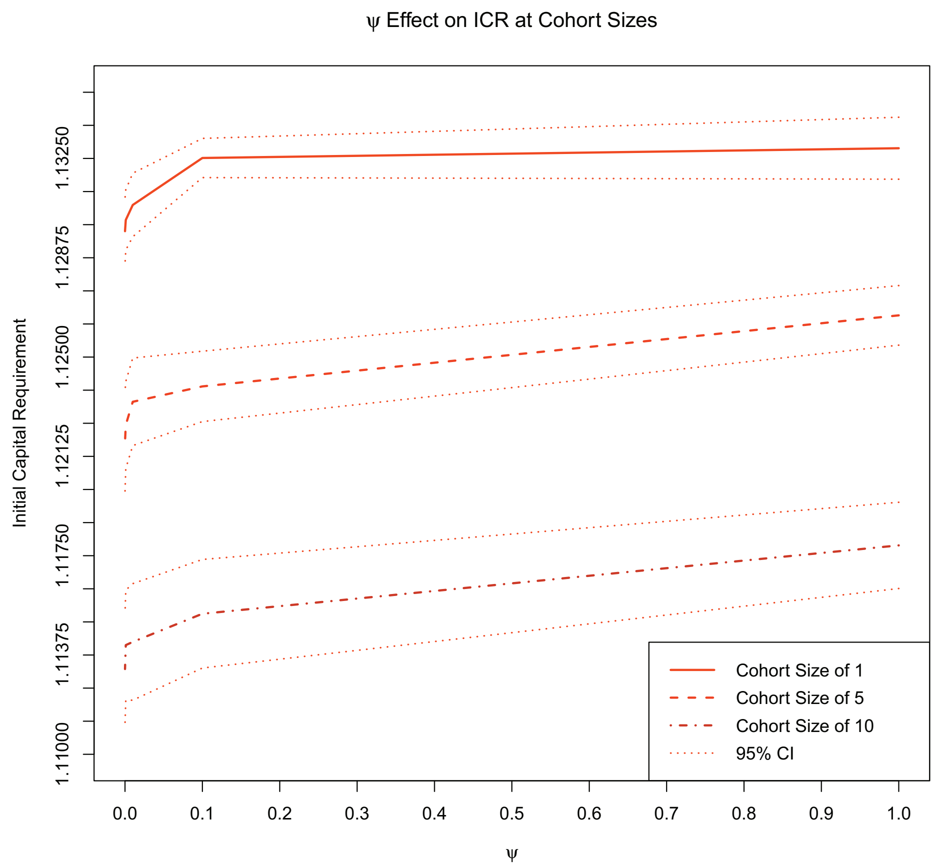

Figure 5.17: Effect of ψ with Differently Sized Cohorts

cyl_bessel_i function in the *Boost C++ Libraries*. This implementation returns accurate results for low non-centrality as noted by van Aubel and Gawronski (2003). The fact that this is not a closed form function and its reliance on recursion makes it very slow. In fact this routine needs to be run m times at every m node at every $N-1$ time step, more over this needs to be done 2 times since the interest rate process is actually a mixture of two non-central χ^2 distributions. During code analysis it was shown that this routine takes $\approx 42\%$ of CPU time, by far the largest for a single routine. This is the main reason why in the above examples m was restricted to 1,000, if the weights were all given by closed form functions this could be substantially increased.

Secondly, the framework presented relies on linear programming, for which the *GNU Linear Programming Kit* (GLPK) was opted for. It should be noted that without any modifications, the size of the constraint matrix is $(m+1) \times (m+4)$. While given large enough available memory this can be done for fairly large values of m, it can be extremely slow for the purposes in this thesis. Even for $m = 1,000$ which is used for the examples above this can be slow since this needs to be done for every m node at each $N-1$ time steps. Note that in Chapter 4 the nodes for time $t + \Delta_1$ are generated from the initial node and not a particular node at t, hence it is possible that the probability of going from Y_{tj} to $Y_{t+\Delta_1,k}$ can be very low. In fact depending on the progression of mortality and surrender in each path, this probability can be zero. Therefore instead of using all m nodes, a subset is used by sampling l nodes from the m nodes using the probabilities calculated from the weights in Chapter 4. In the examples above l was set to 100. (See Figure 5.18) A second note is that, GLPK package version 4.50 which was used here is not re-enterable and should be recompiled with certain modifications in order to be used in a concurrent manner.

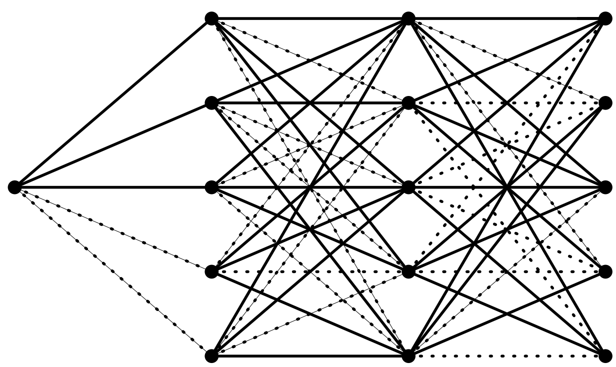

Figure 5.18: Sampled Stochastic Mesh ($m = 5$, $l = 3$)

The *C/C++* code used for the numerical examples is not included in the Appendix of this thesis due to it being written in an object-oriented manner with multiple levels of dependencies. That is, one piece of code would not give the complete picture of the computation and due to the cumbersomeness of the code it is infeasible to include every part. For this reason the source files are available upon request. Finally it should be noted that due to the nature of the mesh, specifically the independence of the calculations for the nodes at a particular time, it is a prime candidate for general-purpose computing on graphics processing units (GPGPU). However to do this, the GLPK package needs to be recompiled (or a different package altogether) either to work with *OpenCL* or Nvidia Corp's *CUDA*. This was not done in these numerical examples since fast enough times were achieved for the intended purposes and the recompiling of the GLPK package is outside the scope of this author.

Conclusions

In this thesis the main objective was to demonstrate the use of iterative risk measures for partial hedging of an equity-linked product with surrender options. The model that is presented in Chapter 1 is rather complex and most practitioners would perhaps see it as cumbersome. The choice of this model however was to demonstrate the flexibility of the valuation method. The method presented in Chapter 3 makes very little assumption permitting the use of realistic surrender schemes like the one presented in Chapter 4.2. In turn, using this surrender scheme it was possible to show that when hedging the riskiest and costliest portion of future liabilities, the capital requirements drop when the policyholders react optimally to market conditions. This result is profound since a company may be selling these products assuming optimal behaviour under another valuation method thinking that doing so would reduce the tail risk incurred by the business. However in this thesis it was demonstrated that if policyholders react non-optimally the risk is actually increased and in fact assuming optimal behaviour would actually leave a business unprepared. It should also be noted that this method is a partial hedging strategy which hedges a particular percentage of the worse cases.

Another advantage of the presented surrender scheme coupled with this particular valuation method is the fact that it is robust with respect to the choice of risk-neutral measure used in the Bermudian option valuation. That is, while the impact of this choice is very large in the evaluation of the price of an option, it is not the case for the initial capital requirement of the partial hedging portfolio.

Due to the flexibility of the methods presented in this thesis, a lot more work can be envisioned. One area where improvements can be made is specifically the surrender scheme and the choice of probability. In Section 4.2.1, the probabilities were a linear piecewise function of the moneyness ratio. It could be argued that instead of using a linear function other functions could be used in order to better capture policyholder behaviour. One method for finding these better functions could be through the use of utility theory. In addition, in order to reduce the initial capital requirements, more investment options can be added to the partial hedging portfolios such as a call option.

Due to the complex nature of the model, the use of simulations was employed. While this method gives a good opportunity to calculate numerical examples, it may also make more subtle behaviours more fuzzy. Due to computing requirements/limits and time constraints, the numerical examples in Chapter 5 were limited to a subset of parameter interactions. In order to better observe these behaviours, one could use a simpler model which avoids the use of cumbersome simulations.

Other areas of improvement are valuation of more complex products that may be path-dependent. Unfortunately the stochastic mesh presented in Chapter 4 is inadequate for this work. While a random tree can be used with minimal change to the valuation method and surrender scheme, this author found that it was difficult to get any meaningful numerical results using this method.

Bibliography

Alfonsi, A. (2005). On the discretization schemes for the cir (and bessel squared) processes. *Monte Carlo Methods and Applications (MCMA)*, 11(4):355–384.

Artzner, P., Delbaen, F., Eber, J.-M., and Heath, D. (1997). Thinking coherently. *Risk Magazine*, 10:68–71.

Artzner, P., Delbaen, F., Eber, J.-M., and Heath, D. (1999). Coherent measures of risk. *Mathematical Finance*, 9(3):203–228.

Bacinello, A. R. (2004). Modelling the surrender conditions in equity-linked life insurance. *Quaderni del Dipartimento di Matematica Applicata alle Scienze Economiche Statistiche e Attuariali Bruno de Finetti, Università degli Studi di Trieste*, (2).

Barber, B. M. and Odean, T. (2011). The behavior of individual investors. *Handbook of the Economics of Finance*, 2:1533–1570.

Bauer, D. K., Kling, A., and Russ, J. (2008). A universal pricing framework for guaranteed minimum benefits in variable annuities. *ASTIN Bulletin*, 38(2):621.

Black, F. and Scholes, M. S. (1973). The pricing of options and corporate liabilities. *Journal of Political Economy*, 81:637–654.

Bollen, N. P. (1998). Valuing options in regime-switching models. *The Journal of Derivatives*, 6(1):38–49.

Bowers, N. (1997). *Actuarial Mathematics*. Society of Actuaries.

Boyle, P. P. and Hardy, M. R. (1997). Reserving for maturity guarantees: Two approaches. *Insurance: Mathematics and Economics*, 21(2):113–127.

Broadie, M. and Glasserman, P. (1997). Pricing american-style securities using simulation. *Journal of Economic Dynamics and Control*, 21(8):1323–1352.

Broadie, M. and Glasserman, P. (2004). A stochastic mesh method for pricing high-dimensional american options. *Journal of Computational Finance*, 7:35–72.

Cox, J. C., Ingersoll Jr, J. E., and Ross, S. A. (1985). A theory of the term structure of interest rates. *Econometrica: Journal of the Econometric Society*, pages 385–407.

Cox, J. C., Ross, S. A., and Rubinstein, M. (1979). Option pricing: A simplified approach. *Journal of Financial Economics*, 7(3):229–263.

Embrechts, P. (2009). Copulas: A personal view. *Journal of Risk and Insurance*, 76(3):639–650.

Embrechts, P., McNeil, A., and Straumann, D. (2002). Correlation and dependence in risk management: properties and pitfalls. In Dempster, M. A. H., editor, *Risk Management: Value at Risk and Beyond*, pages 176–223. Cambridge University Press, Cambridge.

Feller, W. (1951). Two singular diffusion problems. *Annals of Mathematics*, 54:173–182.

Föllmer, H. and Leukert, P. (1999). Quantile hedging. *Finance and Stochastics*, 3(3):251–273.

Föllmer, H. and Sondermann, D. (1986). Hedging of non-redundant contingent claims. *Contributions to Mathematical Economics*.

Forsyth, P. and Vetzal, K. (2014). An optimal stochastic control framework for determining the cost of hedging of variable annuities. *Journal of Economic Dynamics and Control*, 44:29–53.

Gaillardetz, P. and Hachem, S. (2014). Risk minimization for equity-linked products. *Concordia University (Working Paper).*

Gaillardetz, P. and Moghtadai, M. (2014). Equity-linked product evaluation using iterated risk measures. *Concordia University (Working Paper).*

Genest, C., Rémillard, B., and Beaudoin, D. (2009). Goodness-of-fit tests for copulas: A review and a power study. *Insurance: Mathematics and Economics*, 44(2):199–213.

Gerber, H. U. and Shiu, E. S. (1994). Option pricing by esscher transforms. *Transactions of the Society of Actuaries*, 46(99):140.

Grosen, A. and Jørgensen, P. L. (1997). Valuation of early exercisable interest rate guarantees. *Journal of Risk and Insurance*, 64(3):481–503.

Grosen, A. and Jørgensen, P. L. (2000). Fair valuation of life insurance liabilities: the impact of interest rate guarantees, surrender options, and bonus policies. *Insurance: Mathematics and Economics*, 26(1):37–57.

Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica: Journal of the Econometric Society*, 57:357–384.

Hardy, M. (2003). *Investment guarantees: modeling and risk management for equity-linked life insurance*, volume 215. John Wiley & Sons.

Hardy, M. R. (2001). A regime-switching model of long-term stock returns. *North American Actuarial Journal*, 5(2):41–53.

Hardy, M. R. and Wirch, J. L. (2004). The iterated cte: a dynamic risk measure. *North American Actuarial Journal*, 8(4):62–75.

Jondeau, E. and Rockinger, M. (2006). The copula-garch model of conditional dependencies: An international stock market application. *Journal of International Money and Finance*, 25(5):827–853.

Kijima, M. and Wong, T. (2007). Pricing of ratchet equity-indexed annuities under stochastic interest rates. *Insurance: Mathematics and Economics*, 41(3):317–338.

Kurpiel, A. and Roncalli, T. (1998). Option hedging with stochastic volatility. http://ssrn.com/abstract=1031927. [Online; accessed 19-July-2013].

Laeven, R. J. and Goovaerts, M. J. (2004). An optimization approach to the dynamic allocation of economic capital. *Insurance: Mathematics and Economics*, 35(2):299–319.

Lee, H. (2003). Pricing equity-indexed annuities with path-dependent options. *Insurance: Mathematics and Economics*, 33(3):677–690.

Li, D. X. (2000). On default correlation: a copula function approach. *The Journal of Fixed Income*, 9(4):43–54.

LIMRA (2012). Limra: Deferred income annuity sales reach $1 billion; fixed indexed annuity sales hit record high in 2012. http://www.limra.com/Posts/PR/News_Releases/LIMRA__Deferred_Income_Annuity_Sales_Reach_$1_Billion;_Fixed_Indexed_Annuity_Sales_Hit_Record_High_in_2012.aspx. [Online; accessed 1-July-2012].

Lin, S. X. and Tan, K. S. (2003). Valuation of equity-indexed annuities under stochastic interest rates. *North American Actuarial Journal*, 7(4):72–91.

Lin, X. S. (2006). *Introductory Stochastic Analysis for Finance and Insurance*. Wiley-Interscience.

Lin, X. S., Tan, K. S., and Yang, H. (2009). Pricing annuity guarantees under a regime-switching model. *North American Actuarial Journal*, 13(3):316–332.

Melnikov, A. and Skornyakova, V. (2005). Quantile hedging and its application to life insurance. *Statistics & Decisions*, 23(4/2005):301–316.

Melnikov, A. and Smirnov, I. (2012). Dynamic hedging of conditional value-at-risk. *Insurance: Mathematics and Economics*, 51(1):182–190.

Moller, T. (1998). Risk-minimizing hedging strategies for unit-linked life insurance contracts. *ASTIN Bulletin*, 28(1):17–47.

Packham, N. and Schmidt, W. M. (2008). Latin hypercube sampling with dependence and applications in finance. http://ssrn.com/abstract=1269633. [Online; accessed 20-april-2014].

Panjer, H. H., Boyle, P. P., Cox, S. H., Gerber, H. U., and Mueller, H. H. (1998). *Financial Economics: With Applications to Investments, Insurance and Pensions*. Society of Actuaries.

Patton, A. J. (2006). Estimation of multivariate models for time series of possibly different lengths. *Journal of Applied Econometrics*, 21(2):147–173.

Quandt, R. E. (1958). The estimation of the parameters of a linear regression system obeying two separate regimes. *Journal of the American Statistical Association*, 53(284):873–880.

Rockafellar, T. R. and Uryasev, S. (2000). Optimization of conditional value-at-risk. *Journal of Risk*, 2:21–42.

Sklar, A. (1959). *Fonctions de Répartition À N Dimensions Et Leurs Marges*. Université Paris 8.

Tasche, D. (2000). Risk contributions and performance measurement. Technical report, Research paper, Zentrum Mathematik (SCA).

Tiong, S. (2000). Valuing equity-indexed annuities. *North American Actuarial Journal*, 4(4):149–163.

van Aubel, A. and Gawronski, W. (2003). Analytic properties of noncentral distributions. *Applied Mathematics and Computation*, 141(1):3–12.

Vasicek, O. (1977). An equilibrium characterization of the term structure. *Journal of Financial Economics*, 5(2):177–188.

Wang, S. (2002). A set of new methods and tools for enterprise risk capital management and portfolio optimization. pages 43–78.

Wang, T. (1999). A class of dynamic risk measures. http://web.cenet.org.cn/upfile/57263.pdf. [Online; accessed 1-June-2012].

Wirch, J. L. and Hardy, M. R. (1999). A synthesis of risk measures for capital adequacy. *Insurance: mathematics and Economics*, 25(3):337–347.

Appendix A

Code

A.1 R Code for Regime-Switching CIR

A.1.1 Discrete CIR Density

```
CIR.d<-function(par,x0,x,dt){
        cc<-2*par[1]/(par[3]^2*(1-exp(-par[1]*dt)))
        uu<-cc*x0*exp(-par[1]*dt);vv<-cc*x
        qq<-2*par[1]*par[2]/par[3]^2-1
        ll<-cc*exp(-uu-vv)*(vv/uu)^(qq/2)*besselI(2*(uu*vv)^.5,qq)
        return(ll)
}
```

A.1.2 Regime-Switching ML Estimation

```
est.1<-rep(0,8)
like.1<-0

rs.cir.mle<-function(dd,x)
{
        for(i in 1:length(x))
                if(x[i]<0) return(Inf)
        for(i in 1:2)
                if(x[i]>1) return(Inf)
        if(sqrt(2*x[3]*x[4])<x[5] || sqrt(2*x[6]*x[7])<x[8]) return(Inf)

        p11<-x[1]; p22<-x[2]; p12<-1-x[1]; p21<-1-x[2]
```

```
        pi1<-p21/(p12+p21); pi2<-1-pi1

        f1<-function(d,x0){      CIR.d(c(x[3:5]),x0,d,1/52)      }
        f2<-function(d,x0){      CIR.d(c(x[6:8]),x0,d,1/52)      }

        f.1.s<-c((pi1*p11+pi2*p21)*f1(dd[2],dd[1]),(pi1*p12+pi2*p22)
                *f2(dd[2],dd[1]))
        f.1<-sum(f.1.s)
        r.1<-f.1.s/f.1

        l<-log(f.1)
        for(i in 3:(length(dd)))
        {
                f.1.s<-c(p11,p21,p12,p22)*c(f1(dd[i],dd[i-1]),f1(dd[i],dd[i-1]),
                        f2(dd[i],dd[i-1]),f2(dd[i],dd[i-1]))*rep(r.1,2)
                f.1<-sum(f.1.s)
                r.1<-c((f.1.s[1]+f.1.s[2])/f.1,(f.1.s[3]+f.1.s[4])/f.1)
                l<-l+log(f.1)
        }
        if( is.nan(l)||l==Inf) return(0) else return(-l)
}

o1=c(est.1,-like.1)
for(i in 1:100)
{
        par1<-c(runif(1,0,1),runif(1,0,1))
        par2a<-c(runif(1,0,10),runif(1,0,1))
        par2b<-sqrt(2*par2a[1]*par2a[2])-.05

        par3a<-c(runif(1,0,10),runif(1,0,1))
        par3b<-sqrt(2*par3a[1]*par3a[2])-.05
        par<-c(par1,par2a,par2b,par3a,par3b)

        ts=optim(par,rs.cir.mle,dd=rr.w)
        o1<-rbind(o1,c(ts$par,ts$value))
}

est.1=o1[which(o1[,9]==min(o1[,9])),1:8]
like.1=-o1[which(o1[,9]==min(o1[,9])),9]
AIC.1=2*8+2*o1[which(o1[,9]==min(o1[,9])),9]
BIC.1=8*log(length(rr.w))-2*like.1
```

A.2 C/C++ Code

A.2.1 Weight Comparison

```
#include <ppl.h>
#include <boost/math/distributions/normal.hpp>
#include <boost/random/variate_generator.hpp>
#include <boost/random/normal_distribution.hpp>
#include <boost/random/mersenne_twister.hpp>
#include <chrono>
#include "Stoch_Mesh_Test.h"

using namespace std;
using namespace concurrency;

boost::mt19937 gen_test(chrono::system_clock::now().time_since_epoch().count());


double dN(double x, int n, double mu, double sig){
  boost::math::normal norm(mu*double(n), sig*sqrt(double(n)));
  return(pdf(norm, x));
}

double rN(double mu, double sig){
  boost::random::normal_distribution<> norm(mu, sig);
  boost::variate_generator<boost::mt19937&, boost::random::normal_distribution<> >
    next_value(gen_test, norm);
  return(next_value());
}

void Stoch_Mesh(int tt, int m, int n, double *th, double *br){
  double r0 = .03; double s = .4; double div = 0; double s0 = 1;  double K = 1;
  double dt = 1.0 / double(n);  int nt = tt*n;
  double mu = (r0-div - pow(s, 2) / 2.0) * dt;  double sig = s*sqrt(dt);

  //Stock Matrix
  double **Ms = new double*[m];
  for (int i = 0; i < m; i++)
    Ms[i] = new double[nt + 1];

  for (int i = 0; i < m; i++)
    Ms[i][0] = s0;
```

```
for (int i = 0; i < m; i++)
  for (int j = 1; j < (nt + 1); j++)
    Ms[i][j] = Ms[i][j - 1] * exp(rN(mu, sig));

//Denominator Matrix
double **Qp = new double*[m];
for (int i = 0; i < m; i++)
  Qp[i] = new double[nt + 1];

for (int i = 0; i < m; i++)
  Qp[i][0] = 1.0;

for (int i = 0; i < m; i++)
  for (int j = 1; j < (nt + 1); j++)
    Qp[i][j] = dN(log(Ms[i][j] / s0), j, mu, sig);

double *V_th = new double[m]; double *V_br = new double[m];

//Ending value
for (int i = 0; i < m; i++) {
  double temp = max(Ms[i][nt]-K, double(0));
  V_th[i] = temp;    V_br[i] = temp;
}

double v = exp(-r0*dt);
for (int jj = (nt - 1); jj >= 0; jj--)  {
  double *P_th = new double[m];    double *P_br = new double[m];
  if (jj == 0) m = 1;

  parallel_for(0, m, [&](int ii)  {
    double *qq = new double[m];
    double *W_th = new double[m];    double *W_br = new double[m];

    for (int i = 0; i < m; i++)
      qq[i] = dN(log(Ms[i][jj + 1] / Ms[ii][jj]), 1, mu, sig)
        / Qp[i][jj + 1];

    qq[ii] = 1.0;
    double W_sum = 0;
    for (int i = 0; i < m; i++)
      W_sum += qq[i];
```

```
      for (int i = 0; i < m; i++) {
        W_th[i] = qq[i] / W_sum;  W_br[i] = qq[i] / double(m);
      }

      double P0 = max(Ms[ii][jj]-K, double(0));
      double S_th = 0; double S_br = 0;
      for (int i = 0; i < m; i++) {
        S_th += V_th[i] * W_th[i];  S_br += V_br[i] * W_br[i];
      }

      P_th[ii] = max(P0, v*S_th); P_br[ii] = max(P0, v*S_br);
      delete[] qq; delete[] W_th; delete[] W_br;
    });
    for (int i = 0; i < m; i++) {
      V_th[i] = P_th[i];  V_br[i] = P_br[i];
    }
    delete[] P_th; delete[] P_br;
  }

  th[0] = V_th[0]; br[0] = V_br[0];
  delete[] V_th; delete[] V_br;

  for (int i = 0; i < m; i++)
    delete[] Ms[i];
  delete[] Ms;

  for (int i = 0; i < m; i++)
    delete[] Qp[i];
  delete[] Qp;
}
```

Appendix B

Numerical Example Tables

Table B.1: Initial Interest Rate Impact

		r(0) (%)			
		0.5	1	2	3
Type	Stochastic	1.152262	1.150239	1.144494	1.130885
	Constant	1.148525	1.128504	1.092091	1.028354

Table B.2: Dependence Parameter Impact

Factor of ρ_i				
0	0.5	1	2	4
1.148975	1.14849	1.149711	1.150639	1.151366

Table B.3: Esscher Parameter Impact

h	$W(0,0)$	$\hat{P}(0)$
-20	1.056079	1.004692
-10	1.094632	1.027577
-5	1.13684	1.068361
-1	1.14365	1.182629
0	1.145488	1.224168
1	1.144686	1.276933
5	1.146851	1.50764
10	1.148008	1.701284
20	1.148607	1.991466

Table B.4: Participation Rate Impact on Initial Capital Requirement (95% CVaR)

		$\gamma\%$			
		40	60	80	100
$(\beta\%, g\%)$	(90,0)	1.019801	1.059424	1.101603	1.145687
	(90,1)	1.041379	1.08389	1.127228	1.173797
	(90,3)	1.102541	1.143433	1.187944	1.232055
	(100,0)	1.072927	1.115879	1.159206	1.205113
	(100,1)	1.108452	1.148178	1.193363	1.236584
	(100,3)	1.191035	1.227865	1.268617	1.310984

Table B.5: Cap Impact on Initial Capital Requirement (95% CVaR)

		$\zeta\%$			
		5	10	15	∞
$(\beta\%, g\%)$	(90,0)	1.09614	1.13299	1.145358	1.146595
	(90,1)	1.112389	1.158535	1.169813	1.172334
	(90,3)	1.153769	1.209892	1.227589	1.231798
	(100,0)	1.13481	1.188025	1.201479	1.203614
	(100,1)	1.158096	1.215913	1.232514	1.23742
	(100,3)	1.208355	1.282498	1.303699	1.31184

Table B.6: Surrender Parameter Effects on Initial Capital Requirement

		(Fee (%), ϕ (%))					
		(0,0)	(0,1)	(0,5)	(5,0)	(5,1)	(5,5)
ψ	0	1.13242	1.130872	1.130033	1.131299	1.128893	1.124694
	0.001	1.13246	1.131708	1.129219	1.131972	1.129225	1.125511
	0.01	1.133497	1.134444	1.130784	1.132005	1.129614	1.127031
	0.1	1.133418	1.133907	1.131367	1.132204	1.132472	1.127045
	∞	1.132632	1.132794	1.130895	1.132473	1.132765	1.128328

Table B.7: Surrender Method Comparison

		Forsyth	Presented
ψ	0	1.129942	1.130316
	0.001	1.130561	1.130768
	0.01	1.132772	1.130747
	0.1	1.132722	1.132488
	∞	1.133393	1.13338

Table B.8: Impact of Threshold (ψ) on Initial Capital Requirement at Different Cohort Sizes

		l_0		
		1	5	10
ψ	0	1.129753	1.121804	1.11322
	0.001	1.130179	1.122436	1.11412
	0.01	1.130744	1.123311	1.114244
	0.1	1.132518	1.123893	1.115308
	∞	1.132887	1.126575	1.117889