MODELING AGGREGATED RETURNS WITH APPLICATION TO SEGREGATED FUND GUARANTEES

by

Gurbakhshash Singh

B.Sc., University of Calgary, 2003.

A PROJECT SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE
in the School
of
Statistics and Actuarial Science

SIMON FRASER UNIVERSITY
Summer 2006

APPROVAL

Name: Gurbakhshash Singh

Degree: Master of Science

Title of project: Modeling Aggregated Returns with Application to Segregated Fund Guarantees

Examining Committee: Dr. Carl Schwarz
Chair

____

Dr. Gary Parker
Senior Supervisor
Simon Fraser University

____

Dr. Richard Lockhart
Simon Fraser University

____

Dr. Yi Lu
External Examiner
Simon Fraser University

Date Approved: ______________________________

Abstract

In guarantee valuation for a segregated fund, the simulation process can be time-consuming. When simulation calculations are based upon weekly or monthly return models, the computation can be quite lengthy for contracts that extend over decades. Simulation run time can be reduced by decreasing the number of calculations. This is accomplished through an aggregated return model.

We study models for the aggregated returns when the estimated model is Lognormal, an AR(1), two-state regime switching and a Multivariate Lognormal. As an illustration of the aggregate models, we use a conditional tail expectation for valuation of a segregated funds guarantee.

Key words: Regime Switching Model, Multivariate Lognormal Model, Segregated Fund Guarantees, Conditional Tail Expectation, Aggregated Returns.

Dedication

To my Mother and Father, without you I don't think that I would be alive!

Acknowledgements

I would like to begin by thanking my supervisor Dr. Gary Parker. I thank him for his patience and his guidance. I especially enjoyed hearing about his experiences. Thank you for taking me on as your graduate student.

I would like to thank my Examining Committee for being part of a wonderful moment in my life. Many thanks to my fellow graduate students for sharing the last two years with me, you are all amazing. Special thanks to Kelly, Charlene and Sadika, your work is priceless. I thank the professors in both the Mathematics and Statistics Department at Simon Fraser University and the University of Calgary.

While I could continue to list the countless people that have helped me, I would like to address the following to them instead. No matter how long our interactions were, you have impacted me. Your presence has helped me, your thoughts have inspired me, your courage has enthralled me, your strength has guided me, your jokes have made me laugh and your laughter has made me smile.

I would like to finish by thanking my family. To my sisters, I am lucky to be your brother. To my Mother and Father, I am lucky to be your son.

Expectation is the framework for most of the modeling that I do but I have learned that life is about the journey and not the expectation of the journey.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

"The Office of the Superintendent of Financial Institution (OSFI) was created to contribute to public confidence in the Canadian financial system" [1].

With such a mandate, OSFI is obviously involved in regulating and determining adequate reserve levels in the insurance industry. This ensures the insured that when a claim is made, there will (more than likely) be adequate funds available for the insurance company to pay its financial obligations. OSFI is essentially trying to reduce the risk of insolvency of an insurance company and hence bolster confidence in one of the many parts of Canada's financial system.

The area to be studied is the reserve calculation of an insurance product known as a segregated fund with a reset option. While OSFI has created a methodology [2] for such calculations, they do allow companies to propose their own models which, upon approval by OSFI, may be used to determine the reserve. Initially the main intent of this project was to use different models to describe the growth of a segregated fund and to determine the reserve using a Conditional Tail Expectation (CTE(95))[3]. Even though the computer speed has greatly increased over the last decade, there are still issues with the run time of a simulation process. In an academic setting, simulation can be run for weeks and months but in industry the results of the simulation may be required in days or even hours. The

[1] www.osfi-bsif.gc.ca/osfi/index_e.aspx?DetailID=2

[2] www.osfi-bsif.gc.ca/app/DocRepository/1/eng/guidelines/capital/guidelines/MCCSR_2004_e.pdf

[3] Definition in Section 2.3.

main intent of this project developed into reducing run time of simulations by reducing the number of computations.

The general approach taken is to model the returns of the underlying mutual fund of the segregated fund and the reset option. In modeling the returns, we start with the lognormal model which assumes that all returns are independent and identically normally distributed. The second model is an AR(1) which assumes that the return of the next period depends on the current return plus some normally distributed noise. The third model is a two-state regime switching lognormal (RSLN(2)) model which assumes that the returns during any period are generated from one of two distributions used to represent a two-state financial market. The last model is the multivariate lognormal, where the returns for different funds are correlated.

Upon simulating these models to estimate the CTE(95), it became apparent that the simulations were quite time consuming. One of the ways to reduce the run time is to reduce the number of computations by evaluating the segregated fund less frequently. To do this, a model for the rate of return can be constructed for aggregated returns which describe returns for longer time periods. To estimate the parameters of these new models, difficulties may arise if data sets are too small and aggregated returns are too few to give quality estimates of parameters for the model. We propose building models that estimate the aggregated return model parameters based on the original return model parameters. These new models are referred to as aggregated return models. The original data model and the aggregated return models are used to estimate the average of the worst 5% losses that the insurance company would take under the specified model. This estimate of the CTE(95) is used as a measure of the reserve. The model results are then compared for $\widehat{CTE(95)}$ and run times. To conclude, an assessment of the models and the interpretation of the results will be given.

Chapter 2 presents an introduction to segregated funds which entails the product design, modeling the guarantees and a simulation method to value the CTE(95). Chapters 3, 4, 5, and 6 introduce the Lognormal model, the AR(1) model, the RSLN(2) model, and the Multivariate Lognormal model respectively. Within each of the chapters 3 through 6, an associated aggregated return model is proposed for each return model. In each of these chapters, the results for the $\widehat{CTE}(95)$ and the simulation run time are also presented and discussed for both the return model and the aggregated return model. Chapter 7 discusses some final points about the techniques and procedures in this project.

Chapter 2

Segregated Funds

2.1 Introduction

This chapter introduces segregated funds and then introduces models for the guarantees associated with the products in this project. An overview of the simulation method for determining an estimate for the CTE(95) is also presented.

Segregated funds are a distinctly Canadian product which has grown in popularity since the 90's. The increase in their use is partly because of low interest rates in fixed income products. Investors are looking for higher returns but are concerned about the risk of losing part of their investments. Segregated funds are similar to American Variable Annuities and British Unit Linked insurance.

A segregated fund is an insurance product with the growth potential of a mutual fund tied with a set of guarantees. Insurance companies carry a variety of segregated funds which allow investors to have different options when building their portfolio. Companies design segregated funds which profit the company but are also considered by OSFI as good consumer products. The following list are features that a company defines when making a segregated fund product:

1. Term of the Contract and Time to Maturity

2. Types of Guarantees

3. The Guarantee Value

4. Management Expense Ratios (MER)

5. Special Features

While a large discourse can be provided for the terms above, focus is given to a particular product. We consider a product where the investor's guarantee value(GV) is 100% of the Account Value at time zero (AV_0). For this product the $AV_0 = 100$ dollars and the investor is guaranteed this amount when the contract matures at time T. The product also guarantees the beneficiaries of the investor 100% of (AV_0) if the investor dies during the lifetime of the contract. These two *types of guarantees* are called a guaranteed minimum maturity benefit ($GMMB$) and a guaranteed minimum death benefit ($GMDB$) respectively.

At the beginning of the contract, 100 dollars are invested in an account which grows according to an underlying mutual fund. Denote the Account Value at time t as AV_t. For this product the *term of the contract* is 10 years. If no death occurs or *special feature* is used, the contract will mature at the end of 10 years.

The $GMMB$ and the $GMDB$ both pass any loss of the investor to the insurance company. The death benefit states that if the contract holder dies prior to the maturity of the contract at time t, the beneficiaries of the contract receive $max(AV_t, GV)$. The minimum maturity benefit states that if the contract holder survives to the maturity of the contract, the investor will receive $max(AV_T, GV)$. If the underlying mutual fund performs so poorly that the account value is below the guarantee value when the contract matures or death occurs, the insurance company pays the difference. This implies that upon the expiration of the contract or death of the investor, the insurance company pays the AV_t to the investor or the beneficiaries plus an additional payoff where *payoff* $= \max(GV - AV_t, 0)$.

The *Management Expense Ratio*(MER) is a percentage or rate charged on the investors account. The money gathered by the company through this charge is used to cover the expenses associated with the fund such as fund managers, taxes, and guarantees. If the MER charges collected from the account over the lifetime of the contract is insufficient to cover the payoff, then there is a loss for the company. Because of the guarantees, segregated funds charge an MER that is larger than those of a similar mutual fund.

The product also comes with a *special feature* called a *reset*. A reset allows the investor to change the guarantee value to the current account value with the condition that upon resetting a new 10-year term of the contract starts. The reset option is offered for a limited time period. We will consider a product which allows resets for 20 years after the start of the contract.

As an example, suppose a contract is issued today and that five years from today the investor chooses to reset. If no reset occurs in the 10 years following the reset, then the contract will mature at time 15 years from now. Therefore, the longest lifetime of the contract is 30 years from issue if resets occur in such a way that the contract is still in force at year 20 and a reset is then exercised.

One final terminology to define is the *time to maturity*. It is the time remaining in the term of the contract until it matures, in the absence of resets.

2.2 Modeling Product Guarantees

Having defined the segregated fund and the specific product design that we will study, this section describes models for the GMDB and the GMMB.

2.2.1 GMMB

Whether or not a payoff is made by the insurance company is dependent upon the growth of the underlying mutual fund. What needs to be determined is a model for the growth or return of the account.

A return r for a period of time $[t, t+1)$ can be defined as,

$$r_{[t,t+1)} = \log(\frac{x_{t+1}}{x_t}),$$

where $x_t > 0$ is a realization of a random process X_t representing the value of the fund over time. We denote $R_{[t,t+1)}$ as the return random variable over one time interval.

Models considered for $R_{[t,t+1)}$ include,

1. Lognormal Model

2. Autoregressive Model of order 1, AR(1)

3. Two State Regime Switching Lognormal Model, RSLN(2)

4. Multivariate Normal Distribution Model, MVN

The first three models are discussed in Hardy(2003). Each model will be discussed in detail in the following chapters.

Suppose that an insurance company carries three kinds of segregated funds with the product design described earlier. An assumption is made that the returns on the segregated fund closely follow the returns of an associated investment[1].

1. **Fund A** behaves like Vanguard Long Term Investment (VWESX)

2. **Fund B** behaves like S&P500 (GSPC)

3. **Fund C** behaves like the Nasdaq (IXIC)

Monthly fund values for each of these three funds were collected from http://www.yahoo.finance. Using the data available, parameter estimates can be determined for each return model. The estimated models are used to describe the monthly returns of the underlying mutual fund for each of the Funds available in the insurance firm.

2.2.2 GMDB

The future lifetimes of the investors are based upon a population life table from Statistics Canada[2]. It is the 1995-1997 life table for Canadian males. Following Bowers.et.al (1997), the future lifetime random variable of a life aged x is denoted as $T(x)$. For $t \geq 0$

$$\begin{aligned} {}_tq_x &= Pr[T(x) \leq t], \\ {}_tp_x &= 1 - {}_tq_x \\ &= Pr[T(x) > t], \end{aligned}$$

[1]Stock Market symbols in brackets.

[2]http://www.statcan.ca/english/freepub/84-537-XIE/tables/pdftables/cam.pdf

where ${}_tq_x$ is the probability that a life aged x will die in the next t years and ${}_tp_x$ is the probability that the life will survive the next t years.

Furthermore denote by ${}_t|q_x$ the probability that a life aged x will survive t years but will die within the following year and so ${}_t|q_x = {}_tp_x - {}_{t+1}p_x$, where ${}_tp_x = \prod_{i=0}^{t-1} {}_1p_{x+i}$.

In the document provided by Statistics Canada p_x is given for all ages x. For a given life aged 55 the probability of their death at times $t = 0, 1, \ldots$ in the future can be calculated. The mortality parameter estimates determined in the Statistics Canada document are yearly values. However, the simulation is evaluated every month based on monthly returns simulated from the model for returns. To determine the monthly probabilities of death based on yearly probabilities a Uniform Distribution of Death (UDD) assumption was made. That is, for $0 < s < 1$, ${}_sp_x = 1 - s \cdot q_x$. The probability of dying in any month of a given year is the same.

2.3 Simulation Method

This section describes the simulation methodology used to determine the estimate for the CTE(95). The definition of the CTE is given as follows (see Hardy(2003)).

Definition 1. *Let L be the continuous loss random variable, then given α between 0 and 100, the CTE is defined as the expected value of the loss given that the loss falls in the upper $(1-\alpha)\%$ tail of the distribution. For quantile risk measure Q_α, i.e. $P(L > Q_\alpha) = \alpha\%$, $CTE(100-\alpha) = E[L|L > Q_\alpha]$.*

2.3.1 CTE(95)

Since it is not readily apparent how to explicitly determine the CTE(95) for a 10-year term with a 20-year reset period, a minimum maturity and a death benefit, simulation is used to estimate the CTE(95). The following is a general overview of the algorithm used to determine $\widehat{CTE}(95)$.

1. Randomly generate monthly return $r_{[k-1,k)}$. Here we assume that k is a time measured in months $k = 1, \ldots, T$. The generation of $r_{[k-1,k)}$ depends on the return model being

simulated.

2. Determine if the investor dies in month k

3. Account Value grows as follows, $AV_k = (1 - MER)AV_{k-1}e^{r_{(k-1,k)}}$

4. If $(AV_k > ResetThreshold)$ and the investor is alive, then a reset occurs

5. If the investor dies, end the current simulation path

6. The Useable MER in period $k-1$, $UMER_{k-1} = AV_{k-1}(MER/(1+GST\%) - ME)$, is the portion of the charge to the Account (during period $k-1$) used to pay the guarantee. ME are the Maintenance Expenses associated with the segregated fund.

7. $Loss = PV(max(0, GV - AV_T)) - \sum_{i=1}^{T} PV(UMER_{i-1})$, where PV represents the present value based upon a discount factor.

8. Repeat 1 - 7, N times to get a vector of losses

9. Calculate the average of the 5% worst losses. This is an estimate of the CTE(95)

Note that δ, ME, and MER must be measured per simulation period; here monthly. Since they are usually quoted per annum, we must divide them by the periods per year to get the equivalent values for the simulation period.

In step 2, a random uniform (0,1) number generator was used to generate a value u. If ${}_{k-\frac{1}{12}}q_x < u \leq {}_kq_x$, for $k = \frac{1}{12}, \frac{2}{12}, \frac{3}{12}, \ldots$ the investor dies in the interval $(k - \frac{1}{12}, k]$. This is for 12 periods per year and can be adjusted for other periods per year.

Reset Threshold

The reset threshold is the value at which the investor will reset their guarantee value. For the Lognormal, the MVN and the AR(1) model a fixed level of Reset Threshold was used. That is, if the account value was greater than the guarantee value then a reset occurred. We also investigate scenarios with a reset threshold where resets occur if the account value is greater than $1.2 \times GV$ and $1.4 \times GV$. For the RSLN(2), resets depended upon the regime that the mutual fund returns were growing at. There was a reset threshold unique to each regime. Resets are assumed more frequent whenever the regime was in a high volatility state and less frequent when in a low volatility state. The reasoning for this is that when things are uncertain, people watch their money more carefully. The level associated with the high volatility state is the $AV > GV$ and for the low volatility state it is the $AV > 1.2 \times GV$.

Time Steps

With monthly return models, it seems natural for the simulation to be evaluated monthly during the lifetime of the contract. Initially $N = 50000$ and the simulation was repeated many times to get a vector of CTE(95) estimates. From this vector, the average and standard deviation of the CTE(95) estimates were determined. This required the program to run for several hours. In an industry setting, a run time of several hours for one contract on one life is not feasible as the company carries multiple contracts. It was found that the standard deviation for estimates of the CTE(95) based upon twenty CTE(95) estimates from 50000 losses was small enough to warrant only using 50000 losses instead of one million. While this saves computation time, a second idea proposed is to aggregate the returns to get quarterly and yearly returns. Building aggregated return models means evaluating the account value four times a year or once a year which reduces the number of calculations as well as the run time. In the following chapters, a return model and an aggregated return model are presented. The simulation run time and $\widehat{CTE(95)}$ are compared for the two return models.

Time to Maturity

Another consideration is the time to maturity. Not everyone starts their contracts at the same point in time and over time the reserve is adjusted to accommodate for changes in the segregated fund. The CTE(95) is calculated for contracts where 8, 3, and 0 years have passed without a reset or equivalently time to maturity is 2, 7, and 10 years. On the valuation date, the investor is assumed to be aged 55. For example, an individual with a contract with two years to maturity was issued a contract at age 47. The $\widehat{CTE(95)}$ are computed for Account Values of 50, 80 and 100 when time to maturity is 2 and 7 years. For 10 years to maturity the initial account value is 100.

MER calculation

Recall that the MER is a charge to the investor's account which the insurance company uses to pay expenses related to the segregated fund. The MER is determined during the product design creation by the insurance company. For the purpose of the simulation it is a value that must be determined and input into the CTE(95) simulation program.

The MER is split into a portion paying taxes, a portion paying maintenance expenses (ME), and a portion paying off the guarantee (X). It is given by

$$MER = (X + ME)(1 + GST\%).$$

Given ME and GST%, we would like to determine X such that the average total present value of X is enough to pay the average present value payoff at the end of the contract. This is denoted as the average present value (PV) of Total Useable MER (TUMER). Since the company will pay $max(GV - AV_T, 0)$ the average present value (PV) of the payoff is $E[PV(max(GV - AV_T, 0))]$. This rule implies that X will be chosen such that

$$E[PV(TUMER)] = E[PV(max(GV - AV_T, 0))].$$

To calculate the MER for the reset product, we first determined X satisfying the above equation for a 10-year term contract with no reset and no death benefit. The MER for the corresponding product with a reset is set at

$$MER_{reset} = (1.5X + ME)(1 + GST\%) + ProfitMargin,$$

where the Profit Margin is an assumed 5 basis points for Fund A and 10 basis points for funds B and C. Since X is determined for a no reset and no death benefit product, an arbitrary adjustment of 50% is made to X to apply it to a product with a reset and death benefit.

Under certain assumptions $E[PV(max(GV-AV_T, 0))]$ can be calculated explicitly using the Black Scholes Formula while $E[PV(TUMER)]$ is estimated through simulation. Using the bisection method, X can be determined for $E[PV(TUMER)] - E[PV(max(GV - AV_T, 0))] = 0$. Since $E[PV(TUMER)]$ is simulated, the random seed for the simulation must be fixed so that $E[PV(TUMER)]$ can be viewed as a function and not a random value.

This is done to have the bisection method converge to a solution. One further assumption is that the continuous rate δ is 5 % per annum.

The calculation of $E[PV(TUMER)]$ is done by the following steps:

1. Randomly generate return $r_{[k-1,k)}$ from a $N(\mu, \sigma^2)$ where $k = 1, \ldots, 10$ years.
2. Account Value grows as follows, $AV_k = (1 - MER)AV_{k-1}e^{r_{[k-1,k)}}$
3. $UMER_{k-1} = AV_{k-1}(MER/(1 + GST\%) - ME)$, where $MER = (X + ME)(1 + GST\%)$
4. $PV(TUMER) = \sum_{i=1}^{T} PV(UMER_{i-1})$
5. Repeat 1 - 4 N times to get a vector of PV(TUMER) to pay the guarantee
6. Calculate the average of $PV(TUMER)$, this is an estimate of $E[PV(TUMER)]$.

Now determine $E[PV(max(GV - AV_T, 0))]$ explicitly using the following idea. Assume that the account value at maturity T, AV_T, is lognormally distributed with $\ln\left(\frac{AV_T}{AV_0}\right) \sim N\left((\delta - MER)T - \frac{\sigma^2 T}{2}, \sigma^2 T\right)$. This assumption coincides with the lognormal model for the rates of return. Making this assumption allows $E[PV(max(GV - AV_T, 0))]$ to be the price of a put option paying a dividend equal to MER according to the Black Scholes formula. For more about the Black Scholes formula refer to Chapter 12 of Hull(2003).

We can now apply the bisection method to find X such that

$$E[PV(UMER)] - E[PV(max(GV - AV_T, 0))] = 0.$$

If X is smaller than 0.05% then X is set as 0.05% or 5 Basis Points (bp). Having determined X, we can use $MER_{reset} = (1.5X + ME)(1 + GST\%) + ProfitMargin$ to determine the MER for the product with a reset feature and death benefit. The final MER is rounded down to the nearest ten. The results are found in Table 2.1.

Table 2.1: MER Determining Values (bp)

Fund	Annual ME	X	Profit Margin	MER_{reset}
A	110	5	5	130
B	125	44	10	210
C	150	100	10	330

2.4 Remarks

Having discussed models for the segregated fund, the following chapters are dedicated to describe the return models, the aggregated return models and the $\widehat{CTE(95)}$ [3].

[3]Simulations were run on a Pentium 4, 3.20 GHz processor and all codes were written and implemented in R.

Chapter 3

Lognormal Model

3.1 Introduction

The *Lognormal Model* assumes that the returns are independent and at time t the return random variable $R_{[t,t+1)} \sim N(\mu, \sigma^2)$. In Hardy(2003), given a data set of returns $r_1, \ldots, r_n$, the Maximum Likelihood (ML) estimates of parameters μ and σ are:

$$\hat{\mu} = \sum_{i=1}^{n} \frac{r_i}{n}$$
$$\hat{\sigma} = \sqrt{\frac{\sum_{i=1}^{n}(r_i - \hat{\mu})^2}{n}}$$

Because monthly return data was modeled, the estimated model is for predicting monthly returns. The simulation runs under this model are time consuming. To reduce the run time, a model for the aggregated returns was determined:

$$\begin{aligned} r_{[t+1,t+2)} + r_{[t,t+1)} &= \log\left(\frac{x_{t+2}}{x_t}\right) \\ &= r_{[t,t+2)}. \end{aligned}$$

The parameters for this longer return or aggregated return will be modeled using non-overlapping rates $r_1 + r_2, r_3 + r_4, \cdots, r_{N-1} + r_N$ resulting in half as much data to estimate the parameters of the aggregated model. For some financial data sets this could prove troublesome since they consist of a small number of data points. Another method may be employed to build aggregate models with parameter estimates from the monthly rate model parameter estimates. The following theorem from Hogg.et.al(2005) page 166 will be quite useful.

Theorem 1. *Let $Y_1, \ldots, Y_n$ be independent random variables such that, for $i = 1, \ldots, n, Y_i$ has a $N(\mu_i, \sigma_i^2)$ distribution. Let $Y = \sum_{i=1}^n a_i Y_i$, where $a_1, \ldots, a_n$ are constants. Then $Y \sim N(\sum_{i=1}^n a_i\mu_i, \sum_{i=1}^n a_i^2\sigma_i^2)$.*

3.2 Aggregated Model

Since $R_{[i,i+1)}$ are independent identically distributed for $i = 1, \ldots, n$ and by Theorem 1, $(R_{[i+1,i+2)} + R_{[i,i+1)}) \sim N(2\mu, 2\sigma^2)$. Using this idea, models for the return of a longer period $[t, t+n+1)$ are based on $R_{[t,t+n+1)} = \sum_{i=t}^{t+n} R_{[i,i+1)} \sim N((n+1)\mu, (n+1)\sigma^2)$.

3.3 Results

The results of the parameter estimation under the lognormal model are used to determine the MER charged for the corresponding fund. The parameters' estimates are,

Table 3.1: Lognormal Model Parameter Estimates based on Monthly Data

Fund	Observation Period	Maximum Likelihood Estimates	
A	Nov87–May06	$\hat{\mu} = 0.0070$	$\hat{\sigma} = 0.0206$
B	Jan56–Dec01	$\hat{\mu} = 0.0059$	$\hat{\sigma} = 0.0422$
C	Feb71–May06	$\hat{\mu} = 0.0074$	$\hat{\sigma} = 0.0642$

These estimates were used to determine $\widehat{CTE(95)}$.

3.4 Discussion

Tables 3.2 - 3.4 contain the results of the simulation runs. A record is made for the run time, $\widehat{CTE}(95)$, number of 50000 losses that had no reset, and the number of GMDB payments. In Tables 3.2, 3.3, and 3.4, it appears that the aggregated return model produces $\widehat{CTE}(95)$ results that are close (within 2 dollars) to those obtained under the monthly return model for

1. a low volatility fund, e.g. fund A,
2. a short Time to Maturity,
3. a current Account Value less than 100,
4. few Resets, either from a
   (a) high reset threshold, or
   (b) small Account Value and short time to maturity making resets unlikely to occur.

It is also apparent that the calculation time for the aggregated model is 10 times faster than the monthly models computation time. This is a significant reduction in run time.

A question that arises from Tables 3.2-3.4 is why there is a difference for the CTE(95) estimates under the two models. To assess what the differences are, a batch of simulations were run with a reset threshold set high enough so that no reset occurred. The results in Table 3.5 show that the CTE(95) estimate from the two models are quite similar. The reset activities seem to be the significant factor in the difference of the two CTE(95) estimates. While the aggregate model appears to capture the growth of the account value very well, it does not appear to capture the reset behavior of the monthly return model. As expected under the UDD assumption, the mortality experience is estimated quite closely between the two simulations. There are a few simulations in which the number of deaths are quite different between the models. It is unclear why this occurred, perhaps because the resets extended the duration of the contract and allowed for more deaths to be observed.

A question remaining is what to do with parameter sets which do not give close CTE(95) estimates between the models. One possibility is to introduce a correction coefficient of 1.2

times the aggregate CTE(95) for time to maturity of 10 to estimate the monthly CTE(95). The coefficient was chosen based on the observed values of the ratio $\frac{\widehat{CTE(95)}_{monthly}}{\widehat{CTE(95)}_{yearly}}$ for a time to maturity of 10. Another possibility is to use aggregated returns to model quarterly returns.

We note that in Table 3.5 there are negative values for the $\widehat{CTE}(95)$. This is a possibility if enough of the 5% worse scenarios have more MER revenues than the amount needed to pay the payoff at the maturity of the contract. This case means that there is a very small chance that the company will take a loss.

Table 3.2: Fund A: Run Time and $\widehat{CTE}(95)$, Lognormal Model

Current Account Value	Period Per Year	Time to Maturity	Reset Threshold	Run Time (s)	$\widehat{CTE}(95)$	Number No Reset Paths	Number of Deaths
50	1	2	1	7.5	48.076	50000	784
50	12	2	1	54.43	48.197	50000	707
50	1	2	1.2	7.53	48.049	50000	740
50	12	2	1.2	54.48	48.233	50000	712
50	1	2	1.4	7.46	48.153	50000	728
50	12	2	1.4	54.56	48.190	50000	665
80	1	2	1	14.69	22.520	39627	5078
80	12	2	1	153.61	22.604	37669	5354
80	1	2	1.2	7.67	22.593	49780	853
80	12	2	1.2	56.92	22.586	49735	816
80	1	2	1.4	7.51	22.487	49999	730
80	12	2	1.4	54.52	22.674	50000	690
100	1	2	1	39.83	6.705	3057	19825
100	12	2	1	448.55	7.519	705	19139
100	1	2	1.2	18.81	6.014	32664	7080
100	12	2	1.2	204.34	6.044	30020	7407
100	1	2	1.4	8.35	5.496	48707	1192
100	12	2	1.4	67.7	5.646	48475	1190
50	1	7	1	22.45	36.128	43016	7492
50	12	7	1	238	37.004	41971	7547
50	1	7	1.2	18.56	35.938	48970	3991
50	12	7	1.2	189.27	36.699	48796	3699
50	1	7	1.4	18.06	35.859	49899	3576
50	12	7	1.4	180.87	36.937	49862	3259
80	1	7	1	47.59	11.040	4441	28531
80	12	7	1	513.95	12.531	3462	27443
80	1	7	1.2	37.31	10.605	16739	20357
80	12	7	1.2	417.89	11.509	14753	19859
80	1	7	1.4	27.91	10.154	32243	11734
80	12	7	1.4	307.17	11.399	30503	11634
100	1	7	1	48.28	5.478	515	30235
100	12	7	1	529.97	9.342	100	28343
100	1	7	1.2	44.75	2.585	3059	26711
100	12	7	1.2	498.77	3.880	2224	25177
100	1	7	1.4	39.5	1.431	10739	21298
100	12	7	1.4	443	2.043	9163	20546
100	1	10	1	51.37	6.685	490	35920
100	12	10	1	568.78	10.919	98	34230
100	1	10	1.2	48.51	3.287	1745	32870
100	12	10	1.2	544.5	4.554	1223	31274
100	1	10	1.4	46.03	1.912	4428	29531
100	12	10	1.4	517.69	2.431	3611	28016

Table 3.3: Fund B: Run Time and $\widehat{CTE}(95)$, Lognormal Model

Current Account Value	Period Per Year	Time to Maturity	Reset Threshold	Run Time (s)	$\widehat{CTE}(95)$	Number No Reset Paths	Number of Deaths
50	1	2	1	7.66	57.148	49898	822
50	12	2	1	55.92	57.170	49873	761
50	1	2	1.2	7.56	56.901	49993	821
50	12	2	1.2	54.92	57.222	49993	684
50	1	2	1.4	7.48	57.164	50000	743
50	12	2	1.4	54.53	57.165	50000	667
80	1	2	1	16.96	37.895	35104	5943
80	12	2	1	196.75	38.942	30592	6807
80	1	2	1.2	9.63	37.290	46475	1819
80	12	2	1.2	87.58	37.618	45179	2067
80	1	2	1.4	7.9	37.012	49352	970
80	12	2	1.4	60.95	37.373	49132	908
100	1	2	1	31.92	33.394	11707	13841
100	12	2	1	392.58	39.155	3172	15448
100	1	2	1.2	18.44	29.662	31200	6294
100	12	2	1.2	219.74	32.225	25590	7336
100	1	2	1.4	11.54	26.461	43409	2551
100	12	2	1.4	110.26	26.869	41403	2874
50	1	7	1	23.86	47.383	39733	8326
50	12	7	1	257.01	47.970	37898	8547
50	1	7	1.2	20.39	47.415	45280	5427
50	12	7	1.2	215.46	47.756	44221	5407
50	1	7	1.4	19.03	47.524	47965	4222
50	12	7	1.4	195.87	47.734	47384	4022
80	1	7	1	37.89	35.733	13891	19770
80	12	7	1	426.93	38.802	10570	19554
80	1	7	1.2	30.85	34.617	25300	13186
80	12	7	1.2	366.37	35.451	21678	13487
80	1	7	1.4	25.72	33.803	33768	9338
80	12	7	1.4	286.07	34.311	31020	9474
100	1	7	1	42.45	36.671	3818	23404
100	12	7	1	473.59	44.519	967	22623
100	1	7	1.2	36.31	34.566	11711	17654
100	12	7	1.2	410.94	37.809	8758	17308
100	1	7	1.4	31.43	31.192	20694	13198
100	12	7	1.4	354.07	32.811	17291	13369
100	1	10	1	45.9	38.762	2687	27871
100	12	10	1	501.94	46.782	702	26754
100	1	10	1.2	40.75	35.622	8125	22221
100	12	10	1.2	457.05	38.935	5695	21580
100	1	10	1.4	37.63	32.200	14726	18167
100	12	10	1.4	421.75	33.985	11932	17983

Table 3.4: Fund C: Run Time and $\widehat{CTE}(95)$, Lognormal Model

Current Account Value	Period Per Year	Time to Maturity	Reset Threshold	Run Time (s)	$\widehat{CTE}(95)$	Number No Reset Paths	Number of Deaths
50	1	2	1	8.63	62.770	48406	1270
50	12	2	1	70.93	63.113	47776	1346
50	1	2	1.2	7.77	62.753	49610	892
50	12	2	1.2	58.4	63.024	49517	813
50	1	2	1.4	7.61	62.576	49912	785
50	12	2	1.4	55.59	62.667	49887	796
80	1	2	1	19.94	54.087	29667	7485
80	12	2	1	241.45	61.474	23547	8535
80	1	2	1.2	12.88	51.098	40787	3425
80	12	2	1.2	152.06	54.141	37573	4128
80	1	2	1.4	9.75	48.114	46011	1850
80	12	2	1.4	90.5	49.424	44598	2102
100	1	2	1	29.72	59.492	13510	12702
100	12	2	1	377.59	76.490	3817	13985
100	1	2	1.2	20.91	55.388	26612	7386
100	12	2	1.2	255.3	63.667	19839	8755
100	1	2	1.4	15	47.380	37001	4191
100	12	2	1.4	165.88	52.738	32750	5043
50	1	7	1	27.73	56.963	32498	11080
50	12	7	1	311.79	61.192	29432	11346
50	1	7	1.2	23.99	55.811	38274	7965
50	12	7	1.2	262.14	57.684	36137	8290
50	1	7	1.4	21.68	54.194	42365	6261
50	12	7	1.4	234.6	55.417	40497	6420
80	1	7	1	37.16	62.608	12812	18694
80	12	7	1	418.34	79.743	9216	18573
80	1	7	1.2	32.02	59.811	20793	13980
80	12	7	1.2	363.53	67.663	16973	14259
80	1	7	1.4	28.59	55.943	27392	11066
80	12	7	1.4	321.43	59.725	23791	11421
100	1	7	1	43.21	77.492	5275	20934
100	12	7	1	450.42	99.514	1433	20331
100	1	7	1.2	35.41	73.147	11557	16767
100	12	7	1.2	402.48	84.637	7723	16615
100	1	7	1.4	32.22	64.920	17751	13730
100	12	7	1.4	362.91	73.054	14139	13841
100	1	10	1	43.86	85.850	3914	24890
100	12	10	1	477.36	112.379	1055	23738
100	1	10	1.2	39.75	78.871	8332	20684
100	12	10	1.2	445.34	97.553	5479	20183
100	1	10	1.4	37.46	73.592	13253	18063
100	12	10	1.4	418.37	81.369	10254	17490

Table 3.5: All Funds: $\widehat{CTE}(95)$, Lognormal Model-No Reset Feature

Fund	Current Account Value	Period Per Year	Time to Maturity	Reset Thres- -hold	Run Time (s)	$\widehat{CTE}(95)$	Number No Reset Paths	Number of Deaths
A	50	1	7	∞	17.81	35.931	50000	3441
A	50	12	7	∞	178.88	36.986	50000	3244
A	80	1	7	∞	17.81	10.168	50000	3451
A	80	12	7	∞	179.47	10.976	50000	3103
A	100	1	7	∞	17.83	0.085	50000	3436
A	100	12	7	∞	178.35	0.162	50000	3069
A	100	1	10	∞	23.64	-0.124	50000	5678
A	100	12	10	∞	248.77	-0.096	50000	5236
B	50	1	7	∞	17.82	47.434	50000	3502
B	50	12	7	∞	178.72	47.557	50000	3044
B	80	1	7	∞	17.92	31.533	50000	3457
B	80	12	7	∞	180.8	31.816	50000	3086
B	100	1	7	∞	17.79	21.554	50000	3452
B	100	12	7	∞	178.53	21.629	50000	3168
B	100	1	10	∞	23.64	17.484	50000	5836
B	100	12	10	∞	249.93	17.087	50000	5231
C	50	1	7	∞	17.87	52.081	50000	3431
C	50	12	7	∞	178.55	52.464	50000	3219
C	80	1	7	∞	17.84	40.288	50000	3395
C	80	12	7	∞	178.2	40.510	50000	3165
C	100	1	7	∞	17.84	31.945	50000	3476
C	100	12	7	∞	178.36	32.737	50000	3151
C	100	1	10	∞	23.81	27.044	50000	5755
C	100	12	10	∞	249.5	27.434	50000	5273

Chapter 4

AR(1) Model

4.1 Introduction

The *AR(1) Model* assumes that the return $R_{[t,t+1)}$, has some dependence upon $R_{[t-1,t)}$, in the form of

$$R_{[t,t+1)} - \mu = \phi(R_{[t-1,t)} - \mu) + \epsilon_t,$$

where $\epsilon_t \sim N(0, \sigma^2)$ are independent error terms. The parameter ϕ represents the measure of dependence between successive returns and the parameter μ is the long-term mean of the process. The ML estimates of the parameters ϕ, μ, and σ are found in Hardy(2003). The estimates of the parameters based on the monthly return data of the funds are in Table 4.1.

Table 4.1: AR(1) Model Parameter Estimates based on Monthly Data

Fund	Estimates		
A	$\hat{\mu} = 0.0070$	$\hat{\sigma} = 0.0205$	$\hat{\phi} = 0.0621$
B	$\hat{\mu} = 0.0059$	$\hat{\sigma} = 0.04221$	$\hat{\phi} = 0.0264$
C	$\hat{\mu} = 0.0074$	$\hat{\sigma} = 0.06374$	$\hat{\phi} = 0.1233$

4.2 Aggregated Model

Assume that the AR(1) is an appropriate model for the monthly data. Let $R_t = R_{[t,t+1)} - \mu_1$ be centered about zero and given r_0 and r_1, the aggregate return model is as follows,

$$\begin{aligned} R_t + R_{t-1} &= \phi(R_{t-1} + R_{t-2}) + \epsilon_t + \epsilon_{t-1} \\ &= \phi^2(R_{t-2} + R_{t-3}) + (\epsilon_t + \epsilon_{t-1}) + \phi(\epsilon_{t-1} + \epsilon_{t-2}) \\ &= \dots \\ &= \phi^{t-1}(r_1 + r_0) + \sum_{j=0}^{t-2} \phi^j \left(\epsilon_{t-j} + \epsilon_{t-j-1}\right), \qquad t \geq 2. \end{aligned}$$

While this is not an AR(1), it is not apparent what sort of model this is. In an effort to make it familiar, we use an idea similar to that found in Telser(1967). Let ϵ_t and ϵ_{t-1} be uncorrelated and identically distributed as $N(0, \sigma_o^2)$. Let $\rho_{t,i} = \sum_{j=0}^{i} R_{t-j}$ and build an AR(1) relating two successive sums of non overlapping returns $\rho_{t,1}$ and $\rho_{t-2,1}$ as

$$\begin{aligned} R_t + R_{t-1} &= \phi^2(R_{t-2} + R_{t-3}) + (\epsilon_t + \epsilon_{t-1}) + \phi(\epsilon_{t-1} + \epsilon_{t-2}) \\ &= \rho_{t,1} \\ &\approx \psi\rho_{t-2,1} + \epsilon_t^*. \end{aligned}$$

Here we approximate the process as an AR(1) with parameter ψ and independent noise terms $\epsilon_t^* \sim N(0, \sigma^2)$. Based on the above formulation, the estimate of the parameter σ^2 for the aggregated return process is $\hat{\sigma}^2 = (1 + (1 + \hat{\phi})^2 + \hat{\phi}^2)\hat{\sigma_o}^2$, where $\hat{\sigma_o}$ is the estimate of σ_o. An estimate of the μ associated with the aggregated process is $2\hat{\mu_1}$.

The estimate of the dependence parameter ψ is

$$\hat{\psi} = \frac{Cov(\rho_{t,1}, \rho_{t-2,1} | \rho_{1,1})}{\sqrt{Var(\rho_{t,1} | \rho_{1,1}) Var(\rho_{t-2,1} | \rho_{1,1})}}$$

where for $t \geq 2$,

$$\begin{aligned} Cov(\rho_{t,1}, \rho_{t-2,1} | \rho_{1,1}) &= Cov(R_t + R_{t-1}, R_{t-2} + R_{t-3} \,| r_1 + r_0) \\ &= Cov\left(\sum_{j=0}^{t-2} \phi^j(\epsilon_{t-j} + \epsilon_{t-j-1}), \sum_{k=0}^{t-4} \phi^k(\epsilon_{t-k-2} + \epsilon_{t-k-3})\right) \end{aligned}$$

with

$$\lim_{t\to\infty} Cov(\rho_{t,1}, \rho_{t-2,1} | \rho_{1,1}) = \frac{(\phi + 2\phi^2 + \phi^3)\sigma_o^2}{1 - \phi^2},$$

and

$$\begin{aligned} Cov(\rho_{t,1}, \rho_{t,1}|\rho_{1,1}) &= Cov(R_t + R_{t-1}, R_t + R_{t-1}|r_1 + r_0) \\ &= Cov\left(\sum_{j=0}^{t-2} \phi^j(\epsilon_{t-j} + \epsilon_{t-j-1}), \sum_{k=0}^{t-2} \phi^k(\epsilon_{t-k} + \epsilon_{t-k-1})\right) \end{aligned}$$

with

$$\lim_{t\to\infty} Cov(\rho_{t,1}, \rho_{t,1}|\rho_{1,1}) = \frac{2(1+\phi)\sigma_o^2}{1-\phi^2}.$$

Since the process is assumed to be stationary, $Cov(\rho_{t,1}, \rho_{t,1}|\rho_{1,1}) = Cov(\rho_{t-2,1}, \rho_{t-2,1}|\rho_{1,1})$. The estimate of the dependence parameter in terms of ϕ is $\hat{\psi} = \frac{\hat{\phi}+\hat{\phi}^2}{2}$ for the aggregated AR(1).

Note that the above aggregation technique can be generalized to any number of aggregated returns. For aggregating n ($n > 2$) returns, we have

$$\begin{aligned} R_t + \ldots + R_{t-(n-1)} &= \phi^n(R_{t-n} + \ldots + R_{t-2n+1}) + \sum_{l=0}^{n-1} \phi^l(\epsilon_{t-l} + \ldots + \epsilon_{t-(n-1)-l}) \\ &= \rho_{t,n-1} \\ &\approx \psi\rho_{t-n,n-1} + \epsilon_t^*, \qquad t \geq n, \end{aligned}$$

in which

$$\begin{aligned} Cov(\rho_{t,n-1}, \rho_{t-n,n-1}|\rho_{n-1,n-1}) &= Cov\left(\sum_{j=0}^{n-1} R_{t-j}, \sum_{j=0}^{n-1} R_{t-j-n}|r_{n-1} + \ldots + r_0\right) \\ &= Cov\left(\sum_{j=0}^{t-n} \phi^j \sum_{i=0}^{n-1} \epsilon_{t-j-i}, \sum_{k=0}^{t-2n} \phi^k \sum_{i=0}^{n-1} \epsilon_{t-k-n-i}\right) \end{aligned}$$

with

$$\lim_{t\to\infty} Cov(\rho_{t,n-1}, \rho_{t-n,n-1}|\rho_{n-1,n-1}) = \frac{\left(\sum_{i=1}^{n} i\phi^i + \sum_{i=1}^{n-1}(n-i)\phi^{n+i}\right)\sigma_o^2}{1-\phi^2},$$

and

$$\begin{aligned} Cov(\rho_{t,n-1}, \rho_{t,n-1}|\rho_{n-1,n-1}) &= Cov\left(\sum_{j=0}^{n-1} R_{t-j}, \sum_{j=0}^{n-1} R_{t-j}|r_{n-1} + \ldots + r_0\right) \\ &= Cov\left(\sum_{j=0}^{t-n} \phi^j \sum_{i=0}^{n-1} \epsilon_{t-j-i}, \sum_{k=0}^{t-n} \phi^k \sum_{i=0}^{n-1} \epsilon_{t-k-i}\right) \end{aligned}$$

with

$$\lim_{t\to\infty} Cov(\rho_{t,n-1}, \rho_{t,n-1}|\rho_{n-1,n-1}) = \frac{\left(n + \sum_{i=1}^{n-1} 2(n-i)\phi^i\right)\sigma_o^2}{1-\phi^2},$$

which gives the estimate of the dependence parameter

$$\hat{\psi} = \frac{\sum_{i=1}^{n} i\hat{\phi}^i + \sum_{i=1}^{n-1}(n-i)\hat{\phi}^{n+i}}{n + \sum_{i=1}^{n-1} 2(n-i)\hat{\phi}^i}.$$

A similar generalization can be written for the σ estimates. For aggregating n ($n > 2$) returns,

$$\hat{\sigma}^2 = \left(\sum_{j=0}^{n-1}\left(\sum_{i=0}^{j}\hat{\phi}^i\right)^2 + \sum_{j=1}^{n-1}\left(\sum_{i=1}^{j}\hat{\phi}^{n-i}\right)^2\right)\hat{\sigma_o}^2.$$

The estimates of the μ for aggregating n ($n > 2$) returns is $n\hat{\mu}_1$. Aggregating 3 and 12 returns or 4 and 1 periods per year (PPY) respectively, the AR(1) parameter estimates are given in Table 4.2.

Table 4.2: AR(1) Aggregate Return Model Parameter Estimates

Fund	PPY	Estimates		
A	1	$\hat{\mu} = 0.084$	$\hat{\sigma} = 0.0755$	$\hat{\psi} = 0.0052$
A	4	$\hat{\mu} = 0.028$	$\hat{\sigma} = 0.0371$	$\hat{\psi} = 0.0217$
B	1	$\hat{\mu} = 0.0708$	$\hat{\sigma} = 0.1498$	$\hat{\psi} = 0.0022$
B	4	$\hat{\mu} = 0.0236$	$\hat{\sigma} = 0.0744$	$\hat{\psi} = 0.0090$
C	1	$\hat{\mu} = 0.0888$	$\hat{\sigma} = 0.2493$	$\hat{\psi} = 0.0107$
C	4	$\hat{\mu} = 0.0296$	$\hat{\sigma} = 0.1206$	$\hat{\psi} = 0.0454$

To test the estimates of the parameters for the aggregated AR(1), a simulation was conducted.

1. A large data set $(r_0, \ldots, r_N)$ is simulated from an AR(1) with $\phi \in [0, 1)$ and $\sigma = 0.01$.

2. Estimates of ϕ and σ are obtained.

3. A new data set is created $\vec{\rho} = (\rho_0, \ldots, \rho_{N/2}) = (r_0 + r_1, r_2 + r_3 \ldots, r_{N-1} + r_N)$.

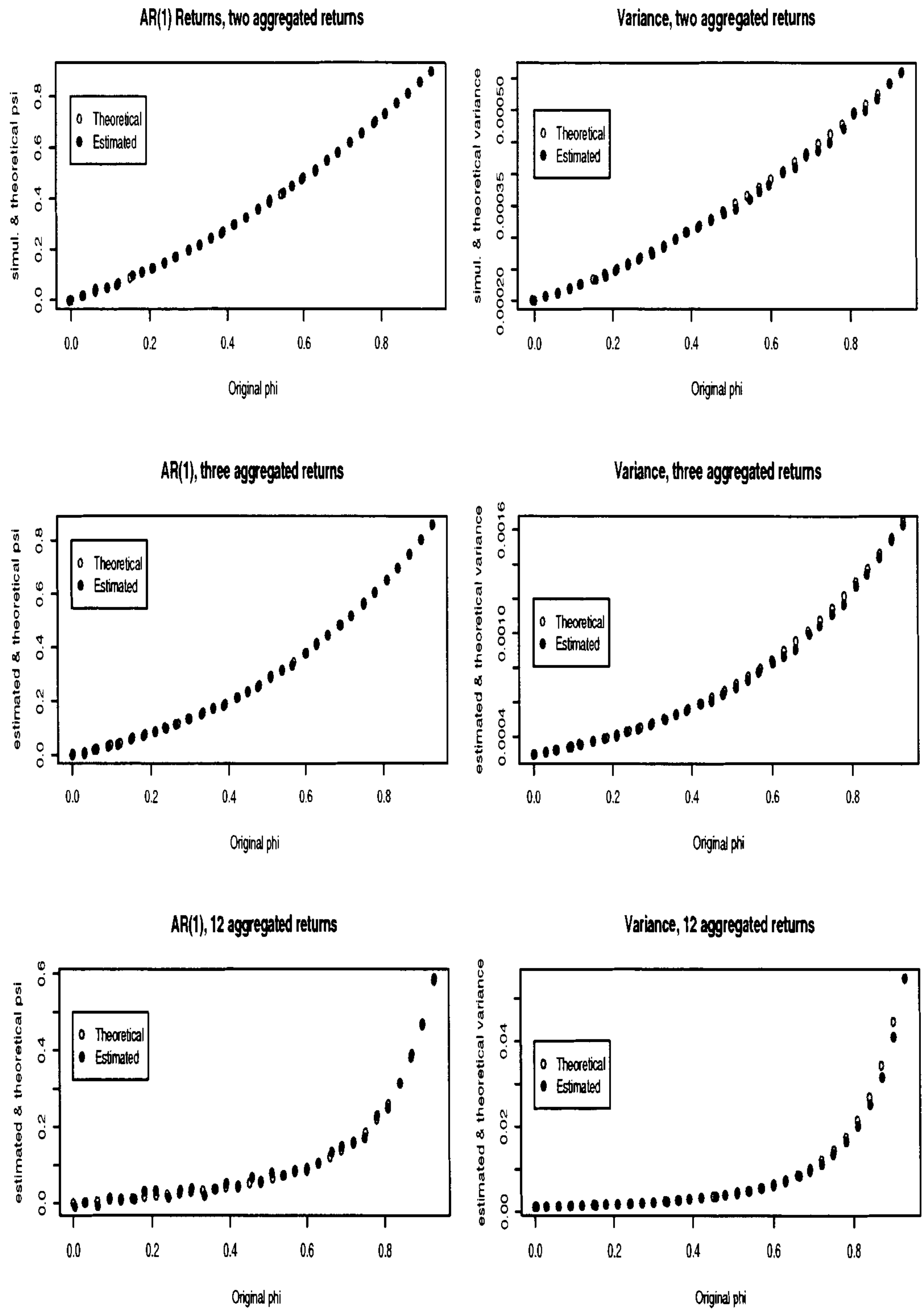
4. An AR(1) is fit to $\vec{\rho}$, estimating a ϕ_ρ and σ_ρ.

5. The parameter estimates of the AR(1) for $\vec{\rho}$ and the theoretical parameter estimates of the AR(1) given above are shown in the set of graphs in Figure 4.1.

The x-axis of the graphs are the ϕ used to simulate the AR(1). On the left hand side of Figure 4.1, the plots are the ϕ on the x-axis and theoretical and estimated $\hat{\psi} = \hat{\phi}_\rho$ on the y-axis for various aggregated returns. The plots on the right hand side are ϕ on the x-axis and the estimated and theoretical $\hat{\sigma}^2_\rho$ on the y-axis.

The simulation is repeated for sums of three consecutive returns and 12 consecutive returns. The graphs seems to indicate that the theoretical parameter estimates and the simulated parameter estimates are quite close. This means that the monthly AR(1) parameter estimates can be used to estimate the corresponding AR(1) for the aggregated returns.

Figure 4.1: Aggregated Returns and associated AR(1) Parameter Estimates



4.3 Results

In tables 4.3-4.5, we set the reset threshold at 1 when simulating the $\widehat{CTE}(95)$. A second consideration was to simulate an AR(1) so that the process started when the initial value of the simulation was no longer influencing the current return. Since the dependence parameter is close to zero for the funds, it was felt that the initial value would have no influence after about 50 simulated AR(1) returns. For each CTE(95) loss simulation path, 50 AR(1) simulated values were discarded before beginning the simulation of the loss.

Table 4.3: Fund A: Run Time and $\widehat{CTE}(95)$, AR(1)

Market Value	Period Per Year	Time to Maturity	Run Time (s)	$\widehat{CTE}(95)$	Number No Reset Paths	Number of Deaths
50	1	2	46.56	48.492	50000	750
50	4	2	63.06	48.693	50000	774
50	12	2	83.06	48.717	50000	721
80	1	2	57.17	23.223	38797	5206
80	4	2	100.36	23.453	37711	5370
80	12	2	189.25	23.396	37028	5537
100	1	2	83.24	8.159	3634	19540
100	4	2	223.15	8.373	1541	19043
100	12	2	476.22	8.798	817	19200
50	1	7	64.25	36.630	42220	7885
50	4	7	137.77	37.176	41448	7735
50	12	7	276.95	37.675	41138	7906
80	1	7	90.73	12.297	4891	28185
80	4	7	247.89	13.069	4034	27289
80	12	7	541.93	13.589	3883	27240
100	1	7	93.73	7.103	533	30058
100	4	7	260.01	9.607	203	28815
100	12	7	562	10.873	103	28390
100	1	10	97.52	8.030	508	35636
100	4	10	279.25	11.342	141	34505
100	12	10	599.92	12.439	94	34128

Table 4.4: Fund B: Run Time and $\widehat{CTE}(95)$, AR(1)

Market Value	Period Per Year	Time to Maturity	Run Time (s)	CTE(95)	Number No Reset Paths	Number of Deaths
50	1	2	47.05	57.432	49885	804
50	4	2	62.4	57.530	49838	770
50	12	2	85.93	57.522	49833	779
80	1	2	59	38.825	34761	5956
80	4	2	114.06	39.640	32086	6537
80	12	2	228.11	39.683	30320	6881
100	1	2	73.28	35.510	11829	13781
100	4	2	196.17	39.183	5748	14803
100	12	2	423.11	41.381	3216	15421
50	1	7	64.26	47.936	39255	8439
50	4	7	144.49	48.445	37990	8615
50	12	7	293.58	48.554	37279	8724
80	1	7	80.78	37.167	13968	19646
80	4	7	211.14	39.425	11777	19333
80	12	7	462.11	40.561	10729	19517
100	1	7	86.41	39.395	3943	23038
100	4	7	233.42	45.111	1787	22516
100	12	7	541.19	46.885	979	22364
100	1	10	90.28	41.700	2772	27788
100	4	10	247.42	46.933	1271	26573
100	12	10	536.43	50.607	733	26381

Table 4.5: Fund C: Run Time and $\widehat{CTE}(95)$, AR(1)

Market Value	Period Per Year	Time to Maturity	Run Time (s)	$\widehat{CTE}(95)$	Number No Reset Paths	Number of Deaths
50	1	2	48.71	65.953	47446	1661
50	4	2	68.56	66.459	47014	1725
50	12	2	108.83	66.611	46571	1683
80	1	2	61.39	67.585	28554	7696
80	4	2	133.9	76.788	24457	8104
80	12	2	276.42	80.522	22707	8529
100	1	2	74.44	82.990	14140	12147
100	4	2	183.36	98.633	7238	13157
100	12	2	402.07	105.498	4152	13652
50	1	7	69.11	67.031	30820	11651
50	4	7	167.05	74.811	28424	11724
50	12	7	348.06	76.968	27669	11838
80	1	7	80.07	90.898	12988	18182
80	4	7	212.75	108.833	10252	17757
80	12	7	444.66	119.293	9291	18006
100	1	7	84.18	110.207	5769	20348
100	4	7	219.89	138.133	2808	19625
100	12	7	481	150.903	1548	19720
100	1	10	89.06	134.128	4437	24190
100	4	10	237.35	164.341	2172	23022
100	12	10	503.26	176.403	1174	22915

4.4 Discussion

The findings for modeling the returns using the AR(1) are quite similar to the Lognormal Model. In Table 4.6, the aggregated models appear to estimate the CTE(95) as well as the monthly model for a no reset product. So the aggregated models seem to capture the growth of the account value just as well as the original model.

When the reset feature is introduced, in Tables 4.3-4.5, for certain cases it appears that the aggregated return models give a close estimate of the CTE(95) under the monthly return model. The cases for which this happens are simulations with less resets and cases similar to those listed in Chapter 3 for the Lognormal Model. The quarterly return model appears to give a very good estimate of the CTE(95) under the monthly return model and it runs twice as fast.

Table 4.6: All Funds: $\widehat{CTE}(95)$, AR(1)-No Reset Feature

Fund	Current Account Value	Period Per Year	Time to Maturity	Reset Thres- -hold	Run Time (s)	$\widehat{CTE}(95)$	Number No Reset Paths	Number of Deaths
A	50	1	7	∞	59.99	36.419	50000	3413
A	50	4	7	∞	110.5	37.134	50000	3218
A	50	12	7	∞	209.25	37.479	50000	3172
A	80	1	7	∞	59.92	11.279	50000	3440
A	80	4	7	∞	111.55	11.936	50000	3201
A	80	12	7	∞	208.45	11.995	50000	3218
A	100	1	7	∞	60.33	0.423	50000	3456
A	100	4	7	∞	110.29	0.443	50000	3327
A	100	12	7	∞	208.21	0.514	50000	3231
A	100	1	10	∞	65.89	-0.016	50000	5657
A	100	4	10	∞	145.68	0.012	50000	5284
A	100	12	10	∞	280.14	0.001	50000	5212
B	50	1	7	∞	60.39	47.558	50000	3456
B	50	4	7	∞	111.3	47.692	50000	3198
B	50	12	7	∞	208.89	48.121	50000	3138
B	80	1	7	∞	59.87	32.150	50000	3474
B	80	4	7	∞	110.65	32.339	50000	3266
B	80	12	7	∞	208.64	32.507	50000	3166
B	100	1	7	∞	60.15	22.455	50000	3465
B	100	4	7	∞	110.47	22.550	50000	3361
B	100	12	7	∞	208.64	22.552	50000	3229
B	100	1	10	∞	66.34	18.626	50000	5841
B	100	4	10	∞	140.36	18.565	50000	5313
B	100	12	10	∞	279.86	18.320	50000	5327
C	50	1	7	∞	60.39	54.515	50000	3421
C	50	4	7	∞	110.91	54.533	50000	3195
C	50	12	7	∞	209.66	54.486	50000	3170
C	80	1	7	∞	59.82	43.686	50000	3359
C	80	4	7	∞	110.45	43.921	50000	3207
C	80	12	7	∞	208.95	43.890	50000	3176
C	100	1	7	∞	60.68	37.125	50000	3473
C	100	4	7	∞	110.92	37.256	50000	3259
C	100	12	7	∞	209.21	37.245	50000	3208
C	100	1	10	∞	66.51	31.054	50000	5705
C	100	4	10	∞	139.87	32.038	50000	5541
C	100	12	10	∞	280.37	31.636	50000	5217

Chapter 5

RSLN(2) Model

5.1 Introduction

The *Two-State Regime Switching Lognormal Model* assumes that the returns for an investment are distributed according to a normal distribution with state dependent parameters μ_{S_t} and σ_{S_t}, where S_t is the state at time t, and

$$R_{[t,t+1)}\,|S_t \sim N(\mu_{S_t}, \sigma^2_{S_t}).$$

At the end of each time unit, switching between the states is determined by a Markovian process. This means that the determination of the current state only depends upon the previous state and no further history. A major issue is that S_t is a latent variable. In estimating the ML estimates for the parameters of the model, there is a special consideration for these latent variables. Denote the transition probabilities between S_t and S_{t+1} as $p_{1,1} = Pr[S_{t+1} = 1\,|S_t = 1]$ and $p_{2,2} = Pr[S_{t+1} = 2\,|S_t = 2]$. This yields a transition Matrix P, given by

$$\mathbf{P} = \begin{pmatrix} p_{1,1} & p_{1,2} \\ p_{2,1} & p_{2,2} \end{pmatrix}.$$

It should be noted that $p_{1,1} + p_{1,2} = 1, p_{2,1} + p_{2,2} = 1$. The Markovian property implies that $Pr[S_{t+1}\,|S_t, S_{t-1}, \ldots, S_1] = Pr[S_{t+1}\,|S_t]$. The parameter space for the model is $\Theta = \{\mu_1, \mu_2, \sigma_1, \sigma_2, p_{1,1}, p_{2,2}\}$. The methodology to determine the ML estimates of the parameters is described in Hardy(2003). The estimates of parameters for monthly returns are given in Table 5.1.

Table 5.1: Parameter Estimates based on Monthly Data

Fund	Estimates RSLN(2)		
A	$\hat{\mu}_1 = 0.0034$	$\hat{\sigma}_1 = 0.0227$	$\hat{p}_{1,1} = 0.7968$
	$\hat{\mu}_2 = 0.0158$	$\hat{\sigma}_2 = 0.0096$	$\hat{p}_{2,2} = 0.4956$
B	$\hat{\mu}_1 = 0.0097$	$\hat{\sigma}_1 = 0.0351$	$\hat{p}_{1,1} = 0.9476$
	$\hat{\mu}_2 = -0.0183$	$\hat{\sigma}_2 = 0.0685$	$\hat{p}_{2,2} = 0.6542$
C	$\hat{\mu}_1 = 0.0142$	$\hat{\sigma}_1 = 0.0432$	$\hat{p}_{1,1} = 0.9701$
	$\hat{\mu}_2 = -0.0178$	$\hat{\sigma}_2 = 0.1081$	$\hat{p}_{2,2} = 0.8932$

5.2 Aggregated Model

By an application of Theorem 1, we get

$$(R_{[t+1,t+2)} + R_{[t,t+1)}|S_{t+1}, S_t) \sim N(\mu_{S_{t+1}} + \mu_{S_t}, \sigma^2_{S_{t+1}} + \sigma^2_{S_t}). \tag{5.1}$$

While the parameters from the two-state model may be used to model this, it is a little unclear how to estimate the transition probabilities, perhaps as a three-state regime switching model. In a three-state regime switching model the parameter set is $\Theta = \{\mu_1, \mu_2, \mu_3, \sigma_1, \sigma_2, \sigma_3, p_{1,2}, p_{1,3}, p_{2,1}, p_{2,3}, p_{3,1}, p_{3,2}\}$. Relative to the RSLN(2), the RSLN(3) has more parameters to compute. Since the return data is aggregated there is less data to estimate the 6 extra parameters. For small data sets this may be a problem. To avoid this an alternate model is presented below to model the aggregate returns based on the RSLN(2).

As an alternative suppose that the RSLN(2) is the best model for the original data then build a Markov Chain whose state transition occurs at time $t-1$ and the next state transition occurs at time $t+1$. Given S_{t-1} and S_{t+1} and a visit to state 1 at time t determines S_{t-1}, S_t, S_{t+1}. For example if we know $S_{t-1} = 1$, $S_{t+1} = 1$ and there are no visits to state 1 in the time between then $(S_{t-1}, S_t, S_{t+1}) = (1, 2, 1)$. The probability associated with travelling this path is given by:

$$
\begin{aligned}
Pr(S_{t+1}=1,S_t=2|S_{t-1}=1) &= Pr(S_{t+1}=1|S_t=2,S_{t-1}=1)Pr(S_t=2|S_{t-1}=1)\\
&= Pr(S_{t+1}=1|S_t=2)Pr(S_t=2\,|S_{t-1}=1)\\
&= p_{2,1}p_{1,2}.
\end{aligned}
$$

The aggregated return corresponding to this path has the following distribution:

$$
\left(R_{[t+1,t+2)}+R_{[t,t+1)}|S_{t+1}=1,S_t=2\right)\sim N(\mu_1+\mu_2,\sigma_1^2+\sigma_2^2).
$$

The idea is that knowing where one starts and where one ends plus information about the number of times one visits state 1 in the time between can determine the probability and distribution of returns associated with a state having that path.

To formalize this, build a state space where each state consists of a 2-tuple (a,b) where a is the current state and b is the number of times state 1 was visited in the previous time steps, for this example only one previous time step must be considered. In this manner the transition matrix in Table 5.2 was constructed, where for simplicity, $p_{1,1}=p$ and $p_{2,2}=q$.

Table 5.2: Markov Chain Transition Probabilities for Two Aggregated Returns

	Next State			
Current State	$S_{t+1}=(1,0)$	$S_{t+1}=(1,1)$	$S_{t+1}=(2,0)$	$S_{t+1}=(2,1)$
$S_{t-1}=(1,0)$	$(1-p)(1-q)$	p^2	$(1-p)q$	$p(1-q)$
$S_{t-1}=(1,1)$	$(1-p)(1-q)$	p^2	$(1-p)q$	$p(1-q)$
$S_{t-1}=(2,0)$	$q(1-p)$	$(1-q)p$	q^2	$(1-q)(1-p)$
$S_{t-1}=(2,1)$	$q(1-p)$	$(1-q)p$	q^2	$(1-q)(1-p)$

The state space of this new Markov chain is $\{(1,0),(1,1),(2,0),(2,1)\}$. This means that a transition from state $(1,0)$ to state $(1,0)$ is equivalent to traveling along the path $(S_{t-1},S_t,S_{t+1})=(1,2,1)$ and the probability associated with this path is $p_{1,2}p_{2,1}=(1-p)(1-q)$.

It is apparent that transitions from state $(1,0)$ to $(1,0)$ and $(1,1)$ to $(1,0)$ have the same probability since S_{t-2} has no impact on the probability calculation. In this Markov chain the rows of the transition matrix $(1,0)$ and $(1,1)$ have the same transition probabilities. Similarly rows $(2,0)$ and $(2,1)$ have the same transition probabilities.

Noting that there are now 4 states at time $t+1$, by (5.1) we have that
State $(1,0)$ has a return with distribution $N(\mu_2+\mu_1, \sigma_1^2+\sigma_2^2)$,
State $(1,1)$ has a return with distribution $N(2\mu_1, 2\sigma_1^2)$,
State $(2,0)$ has a return with distribution $N(2\mu_2, 2\sigma_2^2)$,
State $(2,1)$ has a return with distribution $N(\mu_1+\mu_2, \sigma_1^2+\sigma_2^2)$.

Now based on the estimates of the RSLN(2) it is possible to build a Markov chain to approximate the aggregated returns. This process can be extended to larger sums and results for the sum of 4 and 12 log returns are presented in the Appendix. To formalize the idea a second derivation for the sum of 3 log returns is presented below.

Build a Markov chain whose state transitions occur at time $t-1$ and the next state occurs at time $t+2$. Given S_{t-1} and S_{t+2} and the number of visits to state 1 in the time between, the path $S_{t-1}, S_t, S_{t+1}, S_{t+2}$ has an associated probability and distribution of returns. It should be noted that some paths share the same probability and distribution of returns. For the distribution of returns, apply Theorem 1,

$$\left(R_{[t+2,t+3)}+R_{[t+1,t+2)}+R_{[t,t+1)}|S_{t+2},S_{t+1},S_t\right)\sim N(\mu_{S_{t+2}}+\mu_{S_{t+1}}+\mu_{S_t},\sigma^2_{S_{t+2}}+\sigma^2_{S_{t+1}}+\sigma^2_{S_t}).$$

Note that there are now 6 states where
State $(1,0)$ has a return with distribution $N(\mu_1+2\mu_2, \sigma_1^2+2\sigma_2^2)$,
State $(1,1)$ has a return with distribution $N(2\mu_1+\mu_2, 2\sigma_1^2+\sigma_2^2)$,
State $(1,2)$ has a return with distribution $N(3\mu_1, 3\sigma_1^2)$,
State $(2,0)$ has a return with distribution $N(3\mu_2, 3\sigma_2^2)$,
State $(2,1)$ has a return with distribution $N(\mu_1+2\mu_2, \sigma_1^2+2\sigma_2^2)$,
State $(2,2)$ has a return with distribution $N(2\mu_1+\mu_2, 2\sigma_1^2+\sigma_2^2)$.

Table 5.3 gives the associated transition matrix for the above state process. As an example of the calculation of the transition matrix, we look at the transition from $(1,0)$ to $(2,1)$. This transition is equivalent to $(S_{t-1}, S_t, S_{t+1}, S_{t+2}) = (1, S_t, S_{t+1}, 2)$ where either S_t or S_{t+1} is equal to 1 but not both. The transition probability associated with this is

Table 5.3: Markov Chain Transition Probabilities for Three Aggregated Returns

Next State	Current State (1,0)	Current State (2,0)
(1,0)	$(1-p)q(1-q)$	$q^2(1-q)$
(1,1)	$2p(1-p)(1-q)$	$pq(1-q)+(1-p)(1-q)^2$
(1,2)	p^3	$p^2(1-q)$
(2,0)	$(1-p)q^2$	q^3
(2,1)	$p(1-p)q+(1-p)^2(1-q)$	$2(1-p)q(1-q)$
(2,2)	$p^2(1-p)$	$p(1-p)(1-q)$

$$Pr(S_{t+2}=2,S_{t+1}=1,S_t=2|S_{t-1}=1)+Pr(S_{t+2}=2,S_{t+1}=2,S_t=1|S_{t-1}=1)$$

$$=Pr(S_{t+2}=2|S_{t+1}=1,S_t=2,S_{t-1}=1)$$

$$\times Pr(S_{t+1}=1|S_t=2,S_{t-1}=1)Pr(S_t=2|S_{t-1}=1)$$

$$+Pr(S_{t+2}=2|S_{t+1}=2,S_t=1,S_{t-1}=1)$$

$$\times Pr(S_{t+1}=2|S_t=1,S_{t-1}=1)Pr(S_t=1|S_{t-1}=1)$$

$$=Pr(S_{t+2}=2|S_{t+1}=1)Pr(S_{t+1}=1|S_t=2)Pr(S_t=2|S_{t-1}=1)$$

$$+Pr(S_{t+2}=2|S_{t+1}=2)Pr(S_{t+1}=2|S_t=1)Pr(S_t=1|S_{t-1}=1)$$

$$=p_{1,2}p_{2,1}p_{1,2}+p_{2,2}p_{1,2}p_{1,1}$$

$$=(1-p)^2(1-q)+q(1-p)p$$

From lines 2 to 5 we use the Markov property from the underlying log return model. Using this technique the transition matrix can be calculated. After all these calculations, there is a Markov chain which uses the underlying RSLN(2) model for data of time step 1 to make a model for log returns of time step 2, 3, 4 and 12. See Appendix B for the transition matrices for time steps 4 and 12.

5.3 Results

In running the simulations to determine the CTE(95), for the monthly return model the reset threshold for state 1 (rt1) was 1.2 and the reset threshold for state 2 (rt2) was 1. As discussed earlier, state 2 represents a higher volatility state and investors are assumed to reset to a higher guarantee whenever they get the chance because the markets are more volatile. For p periods per year return models, the reset threshold was the average of the path reset thresholds. For example if there is only 1 period per year and the aggregate model is in a state with 2 visits to state 1 and 10 visits to state 2 in the year, then the reset threshold for the year is $\frac{2\times rt1+10\times rt2}{12}$. A second consideration was to start the simulation as if the model was stationary. The results are given in Tables 5.4-5.6.

Table 5.4: Fund A: Run Time and $\widehat{CTE}(95)$,RSLN(2)

Market Value	Period Per Year	Time to Maturity	Run Time (s)	$\widehat{CTE}(95)$	Number No Reset Paths	Number of Deaths
50	1	2	11.42	48.602	50000	731
50	4	2	32.89	48.723	50000	702
50	12	2	69.24	48.780	50000	728
80	1	2	13.1	23.221	48575	1282
80	4	2	48.94	23.423	46472	1935
80	12	2	163.04	23.456	40895	4146
100	1	2	43.72	7.352	21535	11386
100	4	2	200.25	7.810	11747	14398
100	12	2	547.96	8.093	2180	18450
50	1	7	31.8	36.599	47700	4601
50	4	7	119.31	37.329	46039	5244
50	12	7	293.67	37.516	43180	6791
80	1	7	67.97	11.657	12003	23010
80	4	7	273.16	12.347	7687	24423
80	12	7	642.25	13.087	4476	26527
100	1	7	77.56	4.760	1955	28054
100	4	7	295.14	6.536	893	27281
100	12	7	672.56	8.983	259	28171
100	1	10	82.45	5.484	1291	33509
100	4	10	313.05	7.721	630	32852
100	12	10	723.7	10.459	201	34082

Table 5.5: Fund B: Run Time and $\widehat{CTE}(95)$,RSLN(2)

Market Value	Period Per Year	Time to Maturity	Run Time (s)	$\widehat{CTE}(95)$	Number No Reset Paths	Number of Deaths
50	1	2	11.66	59.470	49991	756
50	4	2	32.9	59.395	49997	746
50	12	2	69.98	59.399	49978	744
80	1	2	16.75	41.483	45677	2076
80	4	2	57.25	42.029	43863	2548
80	12	2	153.6	42.355	40777	3501
100	1	2	33.8	38.150	27981	7458
100	4	2	142.89	39.892	22041	8704
100	12	2	391.67	42.570	14181	11331
50	1	7	34.56	49.368	44281	5916
50	4	7	129.09	49.658	42808	6181
50	12	7	306.36	49.817	40843	7127
80	1	7	53.03	41.188	22786	14416
80	4	7	206.61	41.942	18989	15243
80	12	7	502.61	43.386	14965	17020
100	1	7	62.9	43.469	10274	18449
100	4	7	240.14	46.205	6813	18732
100	12	7	591.52	49.111	3530	20558
100	1	10	69.19	45.892	6999	22859
100	4	10	265.07	50.112	4560	23245
100	12	10	618.01	52.328	2517	24217

Table 5.6: Fund C: Run Time and $\widehat{CTE}(95)$,RSLN(2)

Market Value	Period Per Year	Time to Maturity	Run Time (s)	$\widehat{CTE}(95)$	Number No Reset Paths	Number of Deaths
50	1	2	12.11	69.793	49631	897
50	4	2	35.41	70.017	49404	877
50	12	2	76.6	70.163	49270	873
80	1	2	23.38	70.284	38813	4229
80	4	2	90.61	76.266	35028	4955
80	12	2	227.68	81.266	32011	5796
100	1	2	39.98	89.953	21763	9300
100	4	2	165.35	101.965	15544	10487
100	12	2	407.75	110.304	10788	11808
50	1	7	43.15	70.506	35039	9654
50	4	7	163.45	75.059	32181	10098
50	12	7	382.51	78.747	30028	10810
80	1	7	57.39	99.150	17118	15990
80	4	7	220.07	108.684	13228	16294
80	12	7	519.55	121.788	10849	16995
100	1	7	63.56	125.527	8953	17942
100	4	7	240.79	142.406	5430	18340
100	12	7	553.03	157.482	3276	18816
100	1	10	67.86	139.785	7033	21565
100	4	10	262.38	171.729	4071	21590
100	12	10	588.63	187.241	2525	21783

5.4 Discussion

The findings for modeling the returns using the RSLN(2) are quite similar to the Lognormal Model. In Table 5.7, the aggregated models appear to estimate the CTE(95) as well as the monthly model for a no reset product. The aggregated model appears to capture the growth of the account value as well as the RSLN(2).

In Tables 5.4-5.6, it appears that the aggregated return model give a close estimate of the CTE(95) under the monthly return model for all the same cases as those in the AR(1) model. Again the quarterly return model appears to produce $\widehat{CTE}(95)$ that are quite close to the monthly return model in half the computing time.

Table 5.7: All Funds: $\widehat{CTE}(95)$ for RSLN(2)-No Reset Feature

Fund	Current Account Value	Period Per Year	Time to Maturity	Reset Thres- -hold	Run Time (s)	$\widehat{CTE}(95)$	Number No Reset Paths	Number of Deaths
A	50	1	7	∞	28.99	36.522	50000	3486
A	50	4	7	∞	102.69	37.014	50000	3166
A	50	12	7	∞	228.31	36.778	50000	3008
A	80	1	7	∞	29.62	10.950	50000	3396
A	80	4	7	∞	101.09	11.517	50000	3275
A	80	12	7	∞	226.77	11.650	50000	3144
A	100	1	7	∞	28.97	0.396	50000	3484
A	100	4	7	∞	102.12	0.363	50000	3116
A	100	12	7	∞	227.01	0.494	50000	3099
A	100	1	10	∞	38.9	-0.0287	50000	5647
A	100	4	10	∞	141.47	0.029	50000	5520
A	100	12	10	∞	316.34	0.008	50000	5186
B	50	1	7	∞	28.96	49.147	50000	3398
B	50	4	7	∞	103.11	49.373	50000	3320
B	50	12	7	∞	228.83	49.435	50000	3157
B	80	1	7	∞	28.87	35.016	50000	3472
B	80	4	7	∞	101.87	34.962	50000	3189
B	80	12	7	∞	227.65	34.935	50000	3130
B	100	1	7	∞	28.86	25.944	50000	3432
B	100	4	7	∞	101.58	25.688	50000	3202
B	100	12	7	∞	226.48	26.177	50000	3280
B	100	1	10	∞	39.2	21.630	50000	5675
B	100	4	10	∞	142.36	21.431	50000	5318
B	100	12	10	∞	318.42	21.435	50000	5413
C	50	1	7	∞	32.57	58.377	50000	3424
C	50	4	7	∞	101.05	58.597	50000	3299
C	50	12	7	∞	226.75	58.722	50000	3261
C	80	1	7	∞	29.29	50.254	50000	3444
C	80	4	7	∞	102.17	50.536	50000	3226
C	80	12	7	∞	227.75	50.633	50000	3183
C	100	1	7	∞	28.96	45.053	50000	3430
C	100	4	7	∞	101.54	45.355	50000	3164
C	100	12	7	∞	226.17	46.045	50000	3158
C	100	1	10	∞	39.17	39.238	50000	5713
C	100	4	10	∞	142.31	39.097	50000	5434
C	100	12	10	∞	319.58	39.316	50000	5259

Chapter 6

Multivariate Normal Model

6.1 Introduction

Thus far each fund has been studied independently without considering the dependence in the returns of the funds. In this chapter, the returns are studied with the assumption that there is some correlation between them. The three fund returns are modeled using a Multivariate Normal (MVN), for $i = 1, \ldots, n$ and returns (R)

$$R_i = \begin{pmatrix} R_{A,[i,i+1)} \\ R_{B,[i,i+1)} \\ R_{C,[i,i+1)} \end{pmatrix} \sim N_3 \left(\mu = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{pmatrix}, \Sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_{22} & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_{33} \end{pmatrix} \right).$$

where $R_{Y,[i,i+1)}$ represents the returns for fund Y and period $[i, i+1)$, $\mu = (\mu_1, \mu_2, \mu_3)^t$ and Σ is the variance covariance matrix.

Under the MVN model a single life builds a portfolio with three different segregated funds. Three accounts are established for the funds and returns for each account come from the MVN. Simulating this requires the mortality experience of a single loss simulation path to be the same for each segregated fund.

From page 171 of Johnson(2002) ML estimates for the MVN model parameters based upon observed values $r_1, r_2, \ldots, r_n$ are simply the sample mean and sample variance - covariance matrix, given by

$$\hat{\mu_j} = \bar{r_j} = \sum_{i=1}^{n} \frac{r_{i,j}}{n},$$
$$\hat{\sigma_{jk}} = \frac{\sum_{i=1}^{n}(r_{i,j} - \bar{r_j})(r_{i,k} - \bar{r_k})}{n}.$$

To estimate $\hat{\mu}$ and $\hat{\Sigma}$ the same length of data is used from November 1987 to May 2006. The estimates of the parameters are

$$\hat{\mu} = (0.00698, 0.00766, 0.00904)^t,$$

$$\hat{\boldsymbol{\Sigma}} = \begin{pmatrix} 0.000425 & 0.000162 & 0.000122 \\ 0.000162 & 0.001622 & 0.002247 \\ 0.000122 & 0.002247 & 0.004771 \end{pmatrix},$$

and the correlation matrix is

$$\hat{\rho} = \begin{pmatrix} 1.0000 & 0.1954 & 0.0854 \\ 0.1954 & 1.0000 & 0.8077 \\ 0.0854 & 0.8077 & 1.0000 \end{pmatrix}.$$

From the correlation matrix, note that Fund B and C have highly correlated returns. The MVN estimated here is based upon monthly returns and again the run time for $\widehat{CTE(95)}$ is significant. We propose using an aggregate return model to reduce the computations and hence the run time.

6.2 Aggregated Model

The following theorem from page 165 Johnson(2002) is used in aggregating the MVN model for returns.

Theorem 2. *Let $X_1, \ldots, X_n$ be mutually independent with X_j distributed as $N_p(\mu_j, \Sigma)$. Then $V = c_1X_1 + \ldots + c_nX_n$ is distributed as $N_p(\sum_{j=1}^{n} c_j\mu_j, \left(\sum_{j=1}^{n} c_j^2\right) \Sigma)$.*

This implies that assuming $R_1, \ldots, R_n$ are independent identically distributed random variables with distribution $N_3(\mu, \Sigma)$ then $R = \sum_{i=1}^{n} R_i \sim N_3(n\mu, n\Sigma)$ is the aggregated return model.

6.3 Results

For a portfolio of 3 funds, an individual invests 100 at time 0 into each fund account. The investor is allowed to reset a segregated fund without resetting the other segregated funds in the investor's portfolio. For each fund in the portfolio, a reset threshold of one and a guarantee value of 100 was assumed. The returns were simulated using the same time step for each fund. The $\widehat{CTE}(95)$ is computed for a covariance matrix with zeros in the off-diagonals as well as for the estimated variance-covariance above. From this, we are able to determine if the correlation of the fund returns affects the CTE(95) estimates.

A single run time is determined for each portfolio and the results are given in Table 6.1.

Table 6.1: Run Time and $\widehat{CTE}(95)$:MVN model

Covariance Included	Fund	Period Per Year	Time to Maturity	Run Time (s)	$\widehat{CTE}(95)$	Number No Reset Paths	Number of Deaths
1	A	1	10	272.98	6.510	476	35895
1	B	1	10	-	32.051	1345	32065
1	C	1	10	-	124.980	2951	27177
1	A	4	10	989.71	9.682	166	34479
1	B	4	10	-	37.762	528	30692
1	C	4	10	-	148.279	1325	25898
1	A	12	10	3208.61	10.513	86	34083
1	B	12	10	-	41.367	314	30402
1	C	12	10	-	171.124	759	25707
0	A	1	10	277.55	6.536	483	35841
0	B	1	10	-	31.923	1314	32011
0	C	1	10	-	125.222	2984	27203
0	A	4	10	1060.34	9.722	159	34594
0	B	4	10	-	38.360	524	30822
0	C	4	10	-	154.609	1366	26092
0	A	12	10	3290.26	10.937	81	34179
0	B	12	10	-	41.964	308	30429
0	C	12	10	-	167.643	779	25758

6.4 Discussion

In Table 6.1, there appears to be no difference between CTE(95) estimates using a zero and non-zero covariance for Funds A and B. For Fund C there appears to be a slight increase in the estimate of the CTE(95) for the non-zero covariance model. Again, the quarterly return model gives a similar estimate of the CTE(95) under the monthly model.

Further, Table 6.2 shows that when the covariance is zero and each fund is treated individually, estimates are similar to the zero covariance CTE(95) estimates but run significantly faster. For one period per year the CTE(95) for all three funds treated individually runs in 145.03 seconds while the zero covariance model runs in 277.55 seconds. It appears that the estimates are quite similar for the funds as well.

Table 6.2: Run Time and $\widehat{CTE(95)}$:Lognormal Model

Fund	Period Per Year	Time to Maturity	Run Time (s)	$\widehat{CTE(95)}$	Number No Reset Paths	Number of Deaths
A	1	10	51.37	6.685	490	35920
B	1	10	48.29	32.000	1334	32405
C	1	10	45.37	120.905	2981	27332
A	12	10	568.78	10.919	98	34230
B	12	10	539.39	41.023	281	30188
C	12	10	502.97	178.292	745	25658

Chapter 7

Conclusion

Having studied several models for the rate of return of the account value, it appears that the aggregated return model captures the growth of the account but the reset behavior is not completely captured. This leads to smaller CTE(95) estimates than those given by the monthly return model. It appears that the quarterly return model is sufficient to capture the reset behavior. The computation time is at least twice as fast as the monthly model.

It is quite interesting to see that all the models for returns give a similar value for the CTE(95). The AR(1) and RSLN(2) are quite similar in magnitude for their estimates of the CTE(95). The MVN and Lognormal are very similar as well.

The new ideas presented in this project include the aggregated AR(1), the aggregated return Markov Chain for the RSLN(2), and associating different investor behavior with each state of the RSLN(2). The aggregated models reduce the computation time of the $\widehat{CTE}(95)$ for a segregated fund with a reset feature. However the models did not capture the reset behavior as well as it captured the growth of the account value. This suggests that these techniques can be applied to other problems where computational time can be reduced for modeling the growth rate of a financial instrument.

Based on the values of $\widehat{CTE}(95)$ for fund C, a company may question offering a reset feature on the highly volatile fund as the reserve is quite large.

7.1 Further Work

Further work in this area include deriving aggregated models for higher order AR processes and an ARCH or GARCH model. Another possibility is to use a health index, developed in Health Economics to assess an individual's self evaluation of her or his health status. This information can be quite useful to insurance companies who are faced with the adverse selection problem. It may also allow hypotheses about reset behavior of investors. One final possibility is to use a stochastic model for the interest rate.

Appendix A

Maximum Likelihood Estimation

Both MLE estimation methods are described in Hardy(2003).

A.1 MLE estimate for AR(1)

The basic result is that Y_t depends on Y_{t-1} plus some noise. Specifically,

$$Y_t - \mu = \phi(Y_{t-1} - \mu) + \sigma\epsilon_t,$$

where $\epsilon_t \sim N(0,1)$ are independent error terms. The parameter set for this model is $\Theta = \{\mu, \sigma, \phi\}$ and conditionally we have

$$Y_t|Y_{t-1} \sim N(\mu(1-\phi) + \phi Y_{t-1}, \sigma^2), \qquad t = 2, 3, \ldots, n$$

and furthermore,

$$Y_t \sim N(\mu, \frac{\sigma^2}{1-\phi^2}).$$

The likelihood function is,

$$L(\Theta|\underline{y}) = f(y_1; \Theta) \prod_{t=2}^{n} (f(y_t; \Theta\,|y_{t-1}))$$

Finally the following function is maximized for the parameter set:

$$l(\Theta) = \frac{n}{2}\log(2\pi) + \frac{\log(1-\phi^2)}{2} - n\log(\sigma) \\ - \frac{1}{2}\left\{ \left(\frac{(y_1-\mu)^2(1-\phi^2)}{\sigma^2} \right) + \sum_{t=2}^{n} \left(\frac{(y_t - (1-\phi)\mu - \phi y_{t-1})^2}{\sigma^2} \right) \right\}.$$

A.2 MLE estimate for RSLN(2)

Let S_t denote the state in the time interval $[t, t+1)$. It is assumed that the log returns Y_t depend upon some underlying two state Markov process. In each state or regime the log returns are normally distributed with parameters specific to that state, that is,

$$Y_t|S_t \sim N(\mu_{S_t}, \sigma^2_{S_t}), \qquad S_t = 1, 2.$$

There are 6 parameters to estimate, $\mu_1, \mu_2, \sigma_1, \sigma_2, p_{1,2}, p_{2,1}$, where $p_{1,2}, p_{2,1}$ are the transition probabilities from state 1 to 2 and 2 to 1, respectively. We want $f(y_t|y_{t-1}, \ldots, y_1, \Theta)$ which we get from the following:

$$f(S_t = j, S_{t-1} = i, y_t|y_{t-1}, \ldots, y_1) = Pr(S_{t-1} = i|y_{t-1}, \ldots, y_1)Pr(S_t = j|S_{t-1} = i)f(y_t|S_t),$$

where $Pr(S_t = j|S_{t-1} = i) = p_{i,j}$ and i,j=1,2 and

$$Pr(S_{t-1} = i|y_{t-1}, \ldots, y_1) = \sum_{S_{t-2}=1}^{2} \frac{f(S_{t-1} = i, S_{t-2}, y_{t-1}|y_{t-2}, \ldots, y_1)}{f(y_{t-1}|y_{t-2}, \ldots, y_1)}, \qquad i = 1, 2.$$

Now

$$f(y_t|y_{t-1}, \ldots, y_1, \Theta) = \sum_{S_t=1}^{2} \sum_{S_{t-1}=1}^{2} f(S_t, S_{t-1}, y_t|y_{t-1}, \ldots, y_1).$$

To start this recursion, we need to determine $Pr(S_0)$ which can be found from the stationary distribution $\pi = (\pi_1, \pi_2)$ of the regime switching Markov chain. Under the stationary distribution π and the transition matrix P, we know that $\pi P = \pi$. Note that

$$\mathbf{P} = \begin{pmatrix} p_{1,1} & p_{1,2} \\ p_{2,1} & p_{2,2} \end{pmatrix},$$

then $\pi P = \pi$ gives the following equations:

$$\pi_1 p_{1,1} + \pi_2 p_{2,1} = \pi_1$$

$$\pi_1 p_{1,2} + \pi_2 p_{2,2} = \pi_2$$

Since $p_{1,1} + p_{1,2} = 1$, we get

$$\pi_1 = \frac{p_{2,1}}{p_{1,2} + p_{2,1}}$$

Based on this, the recursion can be computed and the likelihood and parameter estimates are obtained.

Appendix B

Aggregated Returns for the RSLN(2)

B.1 Sum of four returns

When aggregating four returns, we build a Markov chain whose states occur at time $t-1$ and the next state occurs at time $t+3$. In this manner, the state path can be created for the interval $[t-1, t+3]$ if a visit to state 1 at time t, $t+1$, and $t+2$ is also noted. The idea is that if it is known where one starts and where one ends plus information about the number of times one visits state 1 in the time between, then all possible paths from time $t-1$ to $t+3$ can be determined. In this manner applying Theorem 1 yields,

$$\sum_{i=t}^{t+3} Y_i \Bigg| S_{t+3}, \ldots, S_t \sim N(\sum_{i=t}^{t+3} \mu_{S_i}, \sum_{i=t}^{t+3} \sigma^2_{S_i})$$

It is noted that there are now 8 states and for $g = \{0, 1, 2, 3\}$:
State $(1, g)$ has an associated return from the distribution

$$N((1+g)\mu_1 + (3-g)\mu_2, (1+g)\sigma_1^2 + (3-g)\sigma_2^2),$$

State $(2, g)$ has an associated return from the distribution

$$N(g\mu_1 + (4-g)\mu_2, g\sigma_1^2 + (4-g)\sigma_2^2).$$

The transition probabilities for a transition from current state $(1, g)$ to any state one period ahead are the same for all g. Similarly the transition probability for a transition from current state $(2, g)$ to any state one period ahead are the same for all g.

Table B.1: Markov Chain Transition Probabilities for Four Aggregated Returns

Next State	Current State (1,g)	Current State (2,g)
(1,0)	$(1-p)q^2(1-q)$	$q^3(1-q)$
(1,1)	$2p(1-p)q(1-q)+(1-p)^2(1-q)^2$	$2(1-p)q(1-q)^2+pq^2(1-q)$
(1,2)	$3p^2(1-p)(1-q)$	$2p(1-p)(1-q)^2+p^2q(1-q)$
(1,3)	p^4	$p^3(1-q)$
(2,0)	$(1-p)q^3$	q^4
(2,1)	$2(1-p)^2q(1-q)+p(1-p)q^2$	$3(1-p)q^2(1-q)$
(2,2)	$2p(1-p)^2(1-q)+p^2(1-p)q$	$2p(1-p)q(1-q)+(1-p)^2(1-q)^2$
(2,3)	$p^3(1-p)$	$p^2(1-p)(1-q)$

B.2 Sum of 12 log returns

When aggregating twelve returns, we build a Markov Chain whose states occur at time $t-1$ and the next state occurs at time $t+11$. In this manner, the state path can be created for the interval $[t-1, t+11]$ if a visit to state 1 at time $t, \ldots, t+11$, is also noted. The idea is that if it is known where one starts and where one ends plus information about the number of times one visits state 1 in the time between, then all possible paths from time $t-1$ to $t+11$ can be determined. In this manner applying Theorem 1 yields,

$$\sum_{i=t}^{t+11} Y_i \bigg| S_{t+11}, \ldots, S_t \sim N(\sum_{i=t}^{t+11} \mu_{S_i}, \sum_{i=t}^{t+11} \sigma_{S_i}^2)$$

It is noted that there are now 24 states and for $h = \{0, \ldots, 11\}$:
State $(1, h)$ has an associated return from the distribution

$$N((1+h)\mu_1 + (11-h)\mu_2, (1+h)\sigma_1^2 + (11-h)\sigma_2^2),$$

State $(2, h)$ has an associated return from the distribution

$$N(h\mu_1 + (12-h)\mu_2, h\sigma_1^2 + (12-h)\sigma_2^2).$$

The transition probabilities for a transition from current state $(1, h)$ to any state one period ahead are the same for all h. Similarly the transition probability for a transition from current state $(2, h)$ to any state one period ahead are the same for all h.

Table B.2: Markov Chain Transition Probabilities for Twelve Aggregated Returns

Next State	Current State (1,h)
(1,0)	$(1-p)q^{10}(1-q)$
(1,1)	$2p(1-p)q^9(1-q)+9(1-p)^2(1-q)^2q^8$
(1,2)	$3p^2(1-p)q^8(1-q)+24p(1-p)^2q^7(1-q)^2+28(1-p)^3q^6(1-q)^3$
(1,3)	$4p^3(1-p)q^7(1-q)+42p^2(1-p)^2q^6(1-q)^2+84p(1-p)^3q^5(1-q)^3+35(1-p)^4q^4(1-q)^4$
(1,4)	$5p^4(1-p)q^6(1-q)+60p^3(1-p)^2q^5(1-q)^2+150p^2(1-p)^3q^4(1-q)^3+100p(1-p)^4q^3(1-q)^4+15(1-p)^5q^2(1-q)^5$
(1,5)	$6p^5(1-p)q^5(1-q)+75p^4(1-p)^2q^4(1-q)^2+200p^3(1-p)^3q^3(1-q)^3+150p^2(1-p)^4q^2(1-q)^4+30p(1-p)^5q(1-q)^5+(1-p)^6(1-q)^6$
(1,6)	$7p^6(1-p)q^4(1-q)+84p^5(1-p)^2q^3(1-q)^2+210p^4(1-p)^3q^2(1-q)^3+140p^3(1-p)^4q(1-q)^4+21p^2(1-p)^5(1-q)^5$
(1,7)	$8p^7(1-p)q^3(1-q)+84p^6(1-p)^2q^2(1-q)^2+168p^5(1-p)^3q(1-q)^3+70p^4(1-p)^4(1-q)^4$
(1,8)	$9p^8(1-p)q^2(1-q)+72p^7(1-p)^2q(1-q)^2+84p^6(1-p)^3(1-q)^3$
(1,9)	$10p^9(1-p)q(1-q)+45p^8(1-p)^2(1-q)^2$
(1,10)	$11p^{10}(1-p)(1-q)$
(1,11)	p^{12}
(2,0)	$(1-p)q^{11}$
(2,1)	$p(1-p)q^{10}+10(1-p)^2q^9(1-q)$
(2,2)	$p^2(1-p)q^9+18p(1-p)^2q^8(1-q)+36(1-p)^3q^7(1-q)^2$
(2,3)	$p^3(1-p)q^8+24p^2(1-p)^2q^7(1-q)+84p(1-p)^3q^6(1-q)^2+56(1-p)^4q^5(1-q)^3$
(2,4)	$p^4(1-p)q^7+28p^3(1-p)^2q^6(1-q)+126p^2(1-p)^3q^5(1-q)^2+140p(1-p)^4q^4(1-q)^3+35(1-p)^5q^3(1-q)^4$
(2,5)	$p^5(1-p)q^6+30p^4(1-p)^2q^5(1-q)+150p^3(1-p)^3q^4(1-q)^2+200p^2(1-p)^4q^3(1-q)^3+75p(1-p)^5q^2(1-q)^4+6(1-p)^6q(1-q)^5$
(2,6)	$p^6(1-p)q^5+30p^5(1-p)^2q^4(1-q)+150p^4(1-p)^3q^3(1-q)^2+200p^3(1-p)^4q^2(1-q)^3+75p^2(1-p)^5q(1-q)^4+6p(1-p)^6(1-q)^5$
(2,7)	$p^7(1-p)q^4+28p^6(1-p)^2q^3(1-q)+126p^5(1-p)^3q^2(1-q)^2+140p^4(1-p)^4q(1-q)^3+35p^3(1-p)^5(1-q)^4$
(2,8)	$p^8(1-p)q^3+24p^7(1-p)^2q^2(1-q)+84p^6(1-p)^3q(1-q)^2+56p^5(1-p)^4(1-q)^3$
(2,9)	$p^9(1-p)q^2+18p^8(1-p)^2q(1-q)+36p^7(1-p)^3(1-q)^2$
(2,10)	$p^{10}(1-p)q+10p^9(1-p)^2(1-q)$
(2,11)	$p^{11}(1-p)$

Table B.3: Markov Chain Transition Probabilities for Twelve Aggregated Returns

Next State	Current State (2,h)
(1,0)	$q^{11}(1-q)$
(1,1)	$q^{10}(1-q)p+10q^9(1-q)^2(1-p)$
(1,2)	$q^9(1-q)p^2+18q^8(1-q)^2p(1-p)+36q^7(1-q)^3(1-p)^2$
(1,3)	$q^8(1-q)p^3+24q^7(1-q)^2p^2(1-p)+84q^6(1-q)^3p(1-p)^2+56q^5(1-q)^4(1-p)^3$
(1,4)	$q^7(1-q)p^4+28q^6(1-q)^2p^3(1-p)+126q^5(1-q)^3p^2(1-p)^2+140q^4(1-q)^4p(1-p)^3+35q^3(1-q)^5(1-p)^4$
(1,5)	$q^6(1-q)p^5+30q^5(1-q)^2p^4(1-p)+150q^4(1-q)^3p^3(1-p)^2+200q^3(1-q)^4p^2(1-p)^3+75q^2(1-q)^5p(1-p)^4+6q(1-q)^6(1-p)^5$
(1,6)	$q^5(1-q)p^6+30q^4(1-q)^2p^5(1-p)+150q^3(1-q)^3p^4(1-p)^2+200q^2(1-q)^4p^3(1-p)^3+75q(1-q)^5p^2(1-p)^4+6(1-q)^6p(1-p)^5$
(1,7)	$q^4(1-q)p^7+28q^3(1-q)^2p^6(1-p)+126q^2(1-q)^3p^5(1-p)^2+140q(1-q)^4p^4(1-p)^3+35(1-q)^5p^3(1-p)^4$
(1,8)	$q^3(1-q)p^8+24q^2(1-q)^2p^7(1-p)+84q(1-q)^3p^6(1-p)^2+56(1-q)^4p^5(1-p)^3$
(1,9)	$q^2(1-q)p^9+18q(1-q)^2p^8(1-p)+36(1-q)^3p^7(1-p)^2$
(1,10)	$q(1-q)p^{10}+10(1-q)^2p^9(1-p)$
(1,11)	$(1-q)p^{11}$
(2,0)	q^{12}
(2,1)	$11q^{10}(1-q)(1-p)$
(2,2)	$10q^9(1-q)p(1-p)+45q^8(1-q)^2(1-p)^2$
(2,3)	$9q^8(1-q)p^2(1-p)+72q^7(1-q)^2p(1-p)^2+84q^6(1-q)^3(1-p)^3$
(2,4)	$8q^7(1-q)p^3(1-p)+84q^6(1-q)^2p^2(1-p)^2+168q^5(1-q)^3p(1-p)^3+70q^4(1-q)^4(1-p)^4$
(2,5)	$7q^6(1-q)p^4(1-p)+84q^5(1-q)^2p^3(1-p)^2+210q^4(1-q)^3p^2(1-p)^3+140q^3(1-q)^4p(1-p)^4+21q^2(1-q)^5(1-p)^5$
(2,6)	$6q^5(1-q)p^5(1-p)+75q^4(1-q)^2p^4(1-p)^2+200q^3(1-q)^3p^3(1-p)^3+150q^2(1-q)^4p^2(1-p)^4+30q(1-q)^5p(1-p)^5+(1-q)^6(1-p)^6$
(2,7)	$5q^4(1-q)p^6(1-p)+60q^3(1-q)^2p^5(1-p)^2+150q^2(1-q)^3p^4(1-p)^3+100q(1-q)^4p^3(1-p)^4+15(1-q)^5p^2(1-p)^5$
(2,8)	$4q^3(1-q)p^7(1-p)+42q^2(1-q)^2p^6(1-p)^2+84q(1-q)^3p^5(1-p)^3+35(1-q)^4p^4(1-p)^4$
(2,9)	$3q^2(1-q)p^8(1-p)+24q(1-q)^2p^7(1-p)^2+28(1-q)^3p^6(1-p)^3$
(2,10)	$2q(1-q)p^9(1-p)+9(1-q)^2(1-p)^2p^8$
(2,11)	$(1-q)p^{10}(1-p)$

Bibliography

[1] Bowers,N., Gerber,H., Hickman,J., Jones,D., Nesbitt,C. (1997) *Actuarial Mathematics.* Society of Actuaries, Schaumburg, Illinois.

[2] Hardy,M. (2003) *Investment Guarantees: modeling and risk management for equity-linked life insurance.* John Wiley and Sons, New Jersey.

[3] Hogg,R., McKean,J., Craig,A. (2005) *Introduction to Mathematical Statistics.* Prentice Hall, New Jersey.

[4] Hull,J. (2003) *Options, Futures, and Other Derivatives.* Prentice Hall, New Jersey.

[5] Johnson,R., Wichern,D. (2002) *Applied Multivariate Statistical Analysis.* Prentice Hall, New Jersey.

[6] Telser, L. (1967) Discrete Samples and Moving sums in Stationary Stochastic Processes. *Journal of the American Statistical Association*, 62:484-499