Monte Carlo Simulation Approaches to the Valuation and Risk Management of Unit-Linked Insurance Products with Guarantees

Mark J. Cathcart

Thesis submitted
for the degree of
Doctor of Philosophy

School of Mathematical and Computer Sciences
Heriot-Watt University
July 2012

The copyright in this thesis is owned by the author. Any quotation from the thesis or use of any of the information contained in it must acknowledge this thesis as the source of the quotation or information.

Abstract

With the introduction of the Solvency II regulatory framework, insurers face the challenge of managing the risk arising from selling unit-linked products on the market. In this thesis two approaches to this problem are considered:

Firstly, an insurer could project the value of their liabilities to some future time using Monte Carlo simulation in order to reserve adequate capital to cover these with a high level of confidence. However, the complex nature of many liabilities means that valuation is a task requiring further simulation. The resulting 'nested-simulation' is computationally inefficient and a regression-based approximation technique known as least-squares Monte Carlo (LSMC) simulation is a possible solution. In this thesis, the problem of configuring the LSMC method to efficiently project complex insurance liabilities is considered. The findings are illustrated by applying the technique to a realistic unit-linked life insurance product.

Secondly, an insurer could implement a hedging strategy to mitigate their exposure from such products. This requires the calculation of market risk sensitivities (or 'Greeks'). For complex, path-dependent liabilities, these sensitivities are typically estimated using simulation. Standard practice is to use a 'bump and revalue' method. As well as requiring multiple valuations, this approach can be unreliable for higher order Greeks. In this thesis some alternative estimators are developed. These are implemented for a realistic unit-linked life insurance product within an advanced economic scenario generator model, incorporating stochastic interest rates and stochastic equity volatility.

Acknowledgements

Firstly, I would like to thank Professor Alexander McNeil for providing guidance on the research conducted in this PhD program. The discussions with him helped produce the results achieved and conclusions made over the last three years. Also, his comments on the initial draft contributed to an improved final thesis.

Secondly, I would like to thank Dr. Steven Morrison of Barrie and Hibbert. Our regular meetings were of great benefit in aiding my understanding of the finer details of my studies. Also, his knowledge and appreciation of the current technical challenges facing insurers helped shape the direction of my research. I would also like to thank the rest of the staff at Barrie and Hibbert for their hospitality and for providing an inspiring atmosphere in which to work.

I also wish to acknowledge the discussions with many of the participants of the Scottish Financial Risk Academy colloquium on Solvency II at which I presented some of my research. This helped guide the final aspects of the work undertaken in the PhD program.

Finally, I would like to thank my mum and dad for their constant love, support and encouragement throughout the last three years.

ACADEMIC REGISTRY
Research Thesis Submission

Name*:*	MARK JAMES CATHCART		
School/PGI:	MATHEMATICAL AND COMPUTER SCIENCES		
Version: *(i.e. First, Resubmission, Final)*	FINAL	Degree Sought (Award **and** Subject area)	DOCTOR OF PHILIOSOPHY IN FINANCIAL AND ACTUARIAL MATHEMATICS

Declaration

In accordance with the appropriate regulations I hereby submit my thesis and I declare that:

1) the thesis embodies the results of my own work and has been composed by myself
2) where appropriate, I have made acknowledgement of the work of others and have made reference to work carried out in collaboration with other persons
3) the thesis is the correct version of the thesis for submission and is the same version as any electronic versions submitted*.
4) my thesis for the award referred to, deposited in the Heriot-Watt University Library, should be made available for loan or photocopying and be available via the Institutional Repository, subject to such conditions as the Librarian may require
5) I understand that as a student of the University I am required to abide by the Regulations of the University and to conform to its discipline.

* *Please note that it is the responsibility of the candidate to ensure that the correct version of the thesis is submitted.*

Signature of Candidate*:*		Date:	

Submission

Submitted By *(name in capitals):*	
Signature of Individual Submitting:	
Date Submitted:	

For Completion in the Student Service Centre (SSC)

Received in the SSC by *(name in capitals):*			
Method of Submission *(Handed in to SSC; posted through internal/external mail):*			
*E-thesis Submitted (**mandatory for final theses)***			
Signature:		Date:	

Please note this form should bound into the submitted thesis.

Updated February 2008, November 2008, February 2009, January 2011

Contents

Chapter 1

Introduction to the thesis

This thesis is the culmination of research on the topic of the risk-management of unit-linked insurance products which feature an embedded guarantee. In the section which follows, an overview of the research of this thesis and how it relates to the existing literature will be given. But before moving on to this, I feel it is important to give some background to the PhD opportunity from which this thesis comes. This research was funded jointly by the Engineering and Physical Sciences Research Council (EPSRC) and Barrie and Hibbert Ltd. through an industrial CASE studentship. The purpose of such initiatives is to help encourage collaboration between academia and industry through the research of a PhD student. Barrie and Hibbert are a world leader in the provision of economic scenario generation solutions and related consultancy. Therefore, the research in this PhD will have Monte Carlo methodologies at its core. Furthermore, the research the company conducts through its role as a consultant is of both a technical and practical nature and the research in this PhD shares this philosophy.

1.1 Literature review and contributions of thesis

Before discussing some background topics which are relevant to the later chapters of this thesis, a literature review of the previous work on which this thesis builds and an outline of the original contributions of this thesis will be given.

In Part I of the thesis the least-squares Monte Carlo (LSMC) method for projecting insurance liabilities will be investigated. This approximation technique could prove very useful for practitioners in the insurance industry looking for an efficient approach to calculating a solvency capital requirement (SCR) under the Solvency II regulatory framework. The natural simulation approach to such calculations leads to a computational set-up known as nested simulation, where a number of inner valuation scenarios branch out from a number of scenarios projecting future states of the economy. The nested simulation set-up has been discussed previously in the finance literature: Gordy and Juneja [Gor00] investigate how a fixed computational

budget may be optimally allocated between the outer and inner scenarios, given realisations of the relevant risk factors up to some time horizon for a portfolio of derivatives. They also introduce a jack-knife procedure within this set-up for reducing bias levels in estimated values. Bauer, Bergmann and Reuss [Bau11] perform similar analysis for a nested simulation set-up in the context of calculating a SCR. In this paper a mathematical framework for the calculation of a SCR is developed and the nested simulation set-up is shown to result naturally from this framework. In a similar manner to Gordy and Juneja the optimal allocation of outer and inner scenarios within this nested simulation set-up is also investigated, as is the reduction in bias from implementing a jack-knife style procedure. Another line of research investigated in this article is the construction of a confidence interval for the SCR within this nested simulation framework, based on the approach of Lan, Nelson and Staum [Lan07]. Finally, they consider the implementation of screening procedures in the calculation of a SCR. The idea here is to perform an initial simulation run and use the results of this to disregard those outer scenarios which are 'unlikely' to belong to the tail of the liability distribution when performing the final simulation run (which is used to calculate the SCR). This approach follows the paper of Lan, Nelson and Staum [Lan10a]. Bauer, Bergmann and Reuss conclude their article by testing the analysis on a hypothetical insurer selling a single participating fixed-term contract.

Another area in financial mathematics where a nested simulation set-up occurs is the valuation of American options. This will be discussed further in Section 2.1, however we note that calculating the price of an American option by simulation is impractical unless some sort of approximation method is used. One such technique is known as least-squares Monte Carlo (LSMC) simulation and was developed by Carriere [Car96], Tsitsiklis and Roy [Tsi99] and Longstaff and Schwartz [Lon01]. It essentially aims to improve the accuracy of the estimate of the continuation value of the option at each timestep by performing a regression on the key economic variables on which this value depends. This approach has become very popular with practitioners looking to efficiently price American-type financial products in recent years. Some papers which investigate the convergence of the LSMC algorithm for American options are Clément, Lamberton and Protter [Clé02], Stentoft [Ste03], Zanger [Zan09] and Cerrato and Cheung [Cer05]. Such theoretical results of conver-

gence will extend to the case where the LSMC method is applied in the context of calculating an insurance SCR. This alternative context for the LSMC method will now be introduced.

Bauer, Bergmann and Reuss [Bau10] and [Bau11] propose taking this LSMC methodology and applying it to the challenge of calculating a SCR, which also naturally yields a nested simulation set-up. They find the nested simulation set-up is "very time-consuming and, moreover, the resulting estimator is biased" [Bau10], and this is despite some of the extensive analysis given in optimising the allocation of the outer and inner scenarios and reducing levels of bias within this framework. Whereas, they note the LSMC approach is "more efficient and provides good approximations of the SCR". This article does warn, however, of the significance of the choice of the regression model on the success of this approach.

Part I of this thesis will also consider the LSMC approach in a capital adequacy context. In Chapter 3 some analysis will be given regarding the key outstanding issues in the implementation of the technique for calculating a projected insurance liability. In order to make progress we introduce the similar problem of estimating the projected value of a European put option, where the valuation scenarios are performed under the Black-Scholes model. As this alternative problem yields analytical valuations for each outer scenario, the success of the LSMC method under different configurations is far easier to investigate. The results of the investigation of such issues include finding that a stepwise AIC algorithm is a reasonably good approach for selecting the regression model and one which is robust to statistical over-fitting (which is shown to be a problematic issue in the LSMC technique). It is also shown that if the outer fitting scenarios, used to calibrate the regression model, are sampled from the real-world distribution, the fit to the projected value distribution can be somewhat poor in the upper tail. This obviously has consequences in insurance risk-management, where it is the upper tail of the liability distribution which is of key concern. On the other hand, if the outer fitting scenarios are sampled in an alternative manner, based on a quasi-random sampling scheme, it is shown that this gives a significant improvement in the fit in the upper tail of this distribution. Evidence is also presented in Chapter 3 which suggests that some improvement in accuracy may be possible by using orthogonal polynomials in the LSMC regression model. Finally, results are presented indicating that when implementing the LSMC

algorithm, only one pair of antithetic valuation scenarios should be performed, with the remainder of the available computational budget used to generate as large a number of outer fitting scenarios as possible. Some of these issues are discussed by Bauer, Bergmann and Reuss for the nested simulation set-up SCR calculation, thus the analysis for the LSMC framework given in this thesis is complementary to their analysis.

In Chapter 4 the LSMC method is applied to estimate the projected liability distribution of a unit-linked variable annuity contract. This product, which offers equity participation coupled with an embedded guarantee, is typical of the type of insurance product which has become popular with consumers in recent years. Many of the findings from Chapter 3 are used in configuring the LSMC set-up in this insurance context and a thorough analysis of how the ideas developed in this earlier chapter extend to the insurance context is presented. Investigating the issues and ultimate success in applying the LSMC method to this type of VA contract is another original contribution of this thesis. It is found that the LSMC method performs well in estimating percentiles in the upper tail and centre of the liability distribution projected one year into the future. The approach is also found to perform reasonably well in approximating the projected liability distribution at year five, however the fit in the upper tail is somewhat less accurate due to difficulties in implementing quasi-random sampled fitting scenarios in this case. Some lines of promising further research which could help improve the fit in the upper tail for the five year (and also a one year) liability projection are outlined in Section 4.6. Overall, the analysis of Chapter 4 demonstrates the LSMC technique to be a successful method in the challenge of estimating projected insurance liabilities and, hence, in the calculation of a SCR.

As well as being able to accurately value and project complex insurance liabilities, many insurance companies wish to employ a hedging strategy to mitigate some of the risk they are exposed to from selling unit-linked products featuring guarantees on the market. Investigating how such hedging strategies can be developed is the main theme of Part II of this thesis.

In order to construct an effective hedging strategy for an option, one needs to know the sensitivities of the option value to the key risk-drivers on which this quantity depends. These sensitivities are often known collectively as the Greeks, as each

sensitivity is denoted by a different Greek letter. Some references which give an introduction to hedging strategies for options are Baxter and Rennie [Bax96], Wilmott [Wil00] and Bingham and Kiesel [Bin04]. To hedge some of the risk faced in selling unit-linked insurance products, practitioners must similarly determine the sensitivity of the value of the liability to the key risk-drivers on which this depends. Calculating these insurance Greeks would be an easier task if we were to assume the underlying asset and economy were described by the Black-Scholes model. However, if we want a realistic valuation of an insurance liability, we need a more sophisticated description of the underlying equity dynamics and economy. Two equity models which offer this are introduced in Sections 5.1 and 5.2. The structure of both these models was introduced and developed by Cox, Ingersoll and Ross [Cox85] in the context of describing short-term interest rates. Heston [Hes93] later applied this form of model to describe the volatility of equity returns and showed that under this model a semi-analytical formula for the value of a European option could be found. Many years earlier, Merton [Mer76] proposed an extension to the Black-Scholes equity model to include random, discontinuous jumps, in order to give a better fit to observed equity asset dynamics. Bates [Bat96] then combined the Heston model with this Merton model to give a model which is sometimes known as Bates' model, but which we will refer to as the stochastic volatility jump diffusion (SVJD) model. In Section 8.3, we combine the Heston model with the CIR model to give an economic model describing equity, volatility and short-term interest rate dynamics. This model, which we denote as the Heston-CIR model, has not been widely used in the literature. Indeed, it was only after developing this model for the analyses of this thesis that this author became aware of further references in the literature. Grzelak and Oosterlee [Grz10] investigate finding an affine approximation to the Heston-CIR model. This form of approximation will be very useful in efficiently calibrating this model to market observables, as it can be used for very fast pricing of European options.

In Chapter 6 of this thesis the theoretical framework and derivation of the semi-analytical value for a European option under the Heston model is given a complete introduction. This follows the derivation given in Gatheral [Gat06], however the treatment given in this thesis expands on this explanation and also gives some relevant background theory. This should provide greater clarity in illustrating how the semi-analytical formula is constructed. Furthermore, some errors in Gatheral are

highlighted and corrected. In the later sections of Chapter 6, this semi-analytical formula is extended to calculate the liabilities on some simple unit-linked insurance contracts. These are found using the approach of Hardy [Har03] who derived these liability formulae under the Black-Scholes model. Obtaining semi-analytical values and sensitivities for these simple unit-linked insurance products under the Heston and SVJD models is another original contribution of this thesis. For more complex insurance products, however, such analytical formulae are not available. Such products' liabilities must then be valued by numerical techniques, such as Monte Carlo simulation. In Section 5.3 an overview of the discretisation approaches for simulating realisations from the Heston model for equity asset returns is given. Lord et al. [Lor08] introduce and compare some of the simple discretisation schemes for the Heston model. Andersen [And07] proposes a more sophisticated approach for this discretisation which claims to reduce levels of discretisation bias compared to standard discretisation approaches. Other possible discretisation schemes have been proposed by Kahl and Jäckel [Kah05a], Zhu [Zhu08], Halley, Malham and Wiese [Hal09] and Glasserman and Kim [Gla09]. Broadie and Kaya [Bro06] discuss a sampling approach which can simulate realisations from the Heston model without any discretisation bias. This technique is relatively slow to simulate paths, however, and thus may not be a practical approach in an insurance risk-management context where a large number of real-world scenarios are required.

In Chapter 7 an overview of the main approaches for estimating option price sensitivities by Monte Carlo simulation is given. Three standard approaches are reviewed: the bump and revalue method, which is the natural finite difference approach often used in practice; the pathwise method, which was developed in the context of option pricing by Brodie and Glasserman [Bro96]; the likelihood ratio method, developed in the context of option pricing by Broadie and Glasserman [Bro96] and Glasserman and Zhao [Gla00]. Mixed hybrid estimators, introduced by Broadie and Glasserman [Bro96], which combine the latter two of these standard approaches to construct an efficient estimator for second-order sensitivities, will also be reviewed. In Chapter 7, these estimators will be calculated under a Black-Scholes model with fixed withdrawals being subtracted from the equity fund at regular intervals. This model has not, to this author's knowledge, been considered in the literature before. Thus, the development of these estimators for this model is an original contribution of

this thesis. This model captures some of the features of a GMWB variable annuity contract, thus this analysis provides some guidance to the challenge of calculating sensitivity estimators for unit-linked insurance products under the more sophisticated Heston-CIR model. Investigating this problem is the purpose of Chapter 8.

In Section 8.3 the likelihood ratio method is extended to the setting of the Heston-CIR economic model. This is an original innovation and builds on the work of Broadie and Kaya [Bro04], who discuss how the likelihood ratio method can be applied under a Heston model. In Section 8.4 the standard approaches of Section 7.1 and the extension of the likelihood ratio method in Section 8.3 are developed for the sensitivities to the liability of a stylised variable annuity product. The pathwise approach for these sensitivities, derived in Section 8.4.2, follows a similar approach to the article of Hobbs et al. [Hob09], except that this thesis considers a more complex product and a stochastic model for volatility and interest rates. The likelihood ratio method is then extended to find the sensitivities to the liability of our stylised VA product, in Section 8.4.3. In Section 8.4.4, a mixed estimator is constructed for the VA liability gamma sensitivity. Finally, Section 8.5 compares all the estimators developed for the stylised VA product in terms of numerical efficiency. The mixed gamma sensitivity estimator is found to be particularly efficient, which is appealing as this is the sensitivity for which the standard approach performs worst. The development of all these estimators in the context of a variable annuity life insurance contract are original contributions of this thesis, although the pathwise estimator is based on the methodology of Hobbs et al. [Hob09].

1.2 Solvency II insurance directive

Before beginning to introduce and develop the main ideas of this thesis, some context describing where this research will be of interest within the insurance industry will be given. According to the European Commission Solvency II website a general definition of a solvency margin is the amount of "regulatory capital an insurance undertaking is obliged to hold against unforeseen events" [EUSD]. Some form of requirements on such an amount have been in place since the 1970s, with the European Commission (EC) reviewing the solvency rules for European Union member states in the 1990s. This led to some reform of the insurance regulatory framework in Europe known as Solvency I. During the process of developing and implement-

ing Solvency I, however, it became clear that more fundamental regulation with greater scope was necessary. With insurance companies now large, multi-national companies with investments in many different asset-classes in a large number of markets, a regulatory framework which would consider the "overall financial position of the insurance undertaking" and take into account "current developments in insurance, risk management, finance techniques, international financial reporting and prudential standards, etc" has been developed over the last ten years [EUSD]. This framework has become known as Solvency II and European insurance companies have been actively preparing to operate under these new rules and guidelines from the beginning of 2013.

The following summary of the framework will largely follow the Solvency II introductory document of the consultancy firm EMB. The directive is based on three categories of requirements, or pillars. The first pillar is concerned with the quantitative requirements of the framework. There are two levels of capital requirement defined under the regulations: the solvency capital requirement (SCR) and the minimum capital requirement (MCR). Failure to meet each of these requirements will result in differing levels of supervisory intervention. The SCR is "intended to reflect all quantifiable risks" that an insurer could face. The Solvency II directive gives two possible methodologies for calculating this amount: either using a European standard formula or using a firms own internal model of its assets and liabilities. The SCR should also take into account "any risk mitigation techniques" that an insurer may use to minimise its exposure. If the SCR is not met by an insurer, then they "must submit a recovery plan to the supervisor and will be closely monitored to ensure compliance with this plan." The MCR, on the other hand, is a lower level capital requirement, which if breached could trigger withdrawal of authorisation by the relevant supervisor.

The second pillar in the Solvency II directive contains the qualitative requirements. This essentially concerns the system of governance within insurance firms and on how the risk management function should integrate into the organisational structure of a firm. Through this firms must show that there is "proper processes in place for identifying and quantifying their risks in a coherent framework" and supervisors will require that such an internal assessment "reflects the specific risks faced by the firm based on internal data". [EMB10] This process will encourage insurers to

employ models which realistically capture the risks to which they are exposed, both in their risk management practice and regulatory reporting. As a result firms should make "informed business decisions understanding the impact of risk and capital on the firm". The third pillar of Solvency II is concerned with the disclosure of the solvency and general financial stability of each insurance company. As part of this report a description of "risk exposure, concentration, mitigation and sensitivity by risk category" and the "methods of valuation of assets and technical provisions" should be given [EMB10]. Capital adequacy information, including the SCR and MCR levels, should also be provided in these publications.

In this thesis we will be investigating a technique which can be used in calculating a SCR for complex insurance liabilities and also introducing methodologies for calculating hedging strategies for insurance companies who wish to mitigate some of their exposure to such liabilities. This is firmly in the remit of pillar one of the Solvency II requirements. We will now briefly explore the general process through which an insurer calculates a capital requirement. This will follow the Solvency II introductory slides of McNeil [McN11].

Let us consider an insurance company with current net asset value given by V_t. This is just the total assets of the firm minus the total of all the liabilities to which it is exposed. To ensure the firm remains solvent in one years time with some high probability α, it may need to hold some amount of extra capital x_0 which is determined by

$$x_0 = \inf\{x : \mathbb{P}(V_{t+1} + x \cdot (1+i) \geq 0) = \alpha\}, \tag{1.1}$$

where i is the one-year risk-free rate of interest. If x_0 is negative, this signifies the firm is well capitalised and money could be 'taken out', that is additional liabilities could be taken by the business which are not matched by additional assets. With some simple algebra, this can be written

$$x_0 = \inf\{x : \mathbb{P}(V_t - V_{t+1}/(1+i) \leq x + V_t) = \alpha\}, \tag{1.2}$$

which implies that $V_t + x_0 = q_\alpha(V_t - V_{t+1}/(1+i))$, where q_α denotes a quantile at the level α. The sum $V_t + x_0$ can be thought of as the SCR and is a quantile of the distribution of $L_{t+1} = V_t - V_{t+1}/(1+i)$. In general, capital requirements are calculated by applying a risk measure to the distribution L_{t+1}. In the above analysis

this risk-measure is a value-at-risk (VaR) and this is the method typically proposed under Solvency II. However, alternative risk measures could also be employed here. See McNeil, Frey and Embrechts [McN05] for a complete introduction to different financial risk-measures. One possibility is expected shortfall (sometimes known as tail-VaR), which is just the conditional expectation of L_{t+1}, conditional on being in some upper tail of this distribution. But how would an insurance company determine V_t? Well, the Solvency II directive states that "the calculation of technical provisions shall make use of and be consistent with information provided by the financial markets [...] (market consistency). [Article 76] Furthermore, "where future cash flows associated with [...] obligations can be replicated reliably using financial instruments for which a reliable market value is observable the value of technical provisions [...] shall be determined on the basis of the market value of those instruments." [Article 77(4)]

In practice the market consistent valuation of many liabilities (and assets) has to be done on a mark-to-model basis, because there are no relevant quoted prices available in liquid and transparent markets. Preferably the parameters of such models will be determined using fully observed market inputs, although some economic judgement may have to be used.

For firms with complex assets and liabilities, the calculation of V_t can be difficult enough. Determining the distribution of V_{t+1} is even more challenging. The natural Monte Carlo approach for calculating this is computationally demanding and for many liabilities impractical. This will be discussed further in Section 2.1. Part I of this thesis investigates a technique for approximating such a value using Monte Carlo methodologies. The construction of a hedging strategy for mitigating some of the exposure an insurer faces requires accurate and reliable calculations of the sensitivities of this liability to its key risk drivers. For complex insurance liabilities, numerical techniques such as Monte Carlo simulation are required to calculate these sensitivities. Part II of the thesis develops Monte Carlo estimators for the liabilities which arise from complex unit-linked insurance products.

1.3 Variable annuity (VA) insurance products

A form of financial product which will underlie much of the later analysis in this thesis will be the class of variable annuity (VA) insurance products. This type of product has become very popular in the USA and Japan over the last 10-15 years and many experts believe that this success will extend to the UK and Europe in the foreseeable future [Led10]. Before outlining why these products create problems in the context of risk management, a broad definition of what constitutes a variable annuity will be given. Much of this discussion is based on a Faculty of Actuaries Variable Annuity Working Party paper [Led10].

A general definition of a VA product is "any unit-linked or managed fund vehicle which offers optional guarantee benefits as a choice for the customer". One may think of an annuity as an "insurance contract that can help individuals save for retirement through tax-deferred accumulation of assets" and at some later stage, perhaps during retirement, as a "means of receiving payments ... that are guaranteed to last for a specified period, [perhaps] including the lifetime of the annuitant". Thus, from the payment of money upfront, some annuity products will guarantee periodic payments for the remaining lifetime of the policy holder at some point in the future. The difference between a traditional annuity of the past and a variable annuity product is in the optional benefits available to the customer, which offer guaranteed payments to customers at certain policy anniversaries or perhaps upon the death of the policyholder.

Another common property of VA products is the variety of investment options available to the contract owners. This allows them to put some assets into investment funds, allowing the fund to keep pace with inflation, or to choose safer forms of investment. This is similar to unit-linked retirement savings products available in the UK, however the distinguishing feature of these new products is in some of the guarantees offered to customers by these VA products, as mentioned above.

These guarantees generally fall into 4 main classes and a brief description of each of these will be given at the top of the next page:

- **Guaranteed Minimum Death Benefits (GMDBs)** This option guarantees a return of the principal invested, upon the death of the policyholder. If the underlying unit account is greater than this principal, the amount paid on the death of the policyholder would be the balance in the account. A variation to this, which will be included in the product we will analyse later, is the addition of a 'ratchet' feature. Here the principal invested will be guaranteed to accumulate by "periodically locking into (and thereby guaranteeing) the growth in the account balance".

- **Guaranteed Minimum Accumulation Benefits (GMABs)** The benefits of this option are similar to that of the GMDB, except here the guarantee is not conditional on the death of the policyholder, but will initiate at certain policy anniversaries (or between certain dates while the policy remains in force).

- **Guaranteed Minimum Income Benefits (GMIBs)** This option guarantees a minimum income stream in the form of a life annuity from some specified future time. This could be fixed initially or depend on the account balance at annuitisation. The customer would typically lose access to the fund value by choosing this option.

- **Guaranteed Minimum Withdrawal Benefits (GMWBs)** This feature guarantees regular withdrawals from the account balance. For example, a fixed term GMWB option could guarantee that withdrawals of 5% of the original investment can be made by the policyholder for a period of 20 years. Recently some VA products have allowed a GMWB for the lifetime of the policyholder (even if the account value reduces to zero). With the GMWB option, the remaining fund would be paid to the estate of the policyholder on their death, whereas this is not the case with a GMIB.

In the past, with-profits policies were very popular in the UK and Europe. The Variable Annuity Working Party paper states that these products gave customers an "apparently simple product with the prospects of high investment returns ... coupled with a range of guarantees". However, over the past 15 years the UK with-profits business has "declined sharply ... with little prospect of any recovery". This was a result of sustained periods of poor equity returns, which resulted in poor performance of with-profits products, due to insurers not having been prudent enough in

the previous years of strong equity growth. With large exit penalties and a lack of investment control available to the policyholders, the uptake of such products diminished dramatically. Therefore, there appears to be demand for a product which offers some security through certain guarantees, but whose value will not be completely eroded through inflation. These new VA products could prove to meet the customer's needs, without the apparent disadvantages of with-profits policies.

With this in mind many insurers in the UK and Europe are looking to offer VA type products over the coming years. Unfortunately, as much as they might appeal to customers, they create some problems in the context of risk management. An interesting article in the magazine *Risk* in 2004 discusses the problems US insurers have faced in calculating capital requirements amidst the rapid growth of evermore complex VA products [Rud10]. Indeed one of the largest re-insurers of VA guarantees, Cigna, had to stop its reinsurance operations in 2002 as a result of having underestimated reserve requirements. This challenge of calculating realistic capital requirements for complex insurance products was introduced in Section 1.2. Obtaining accurate approximations for such calculations is even more crucial as Europe enters this new phase in insurance regulation.

The VA class of products will feature in the analysis in this thesis as follows: In Chapters 4 and 8 Monte Carlo estimation techniques will be developed for calculating the projected liability value and the sensitivity of the liability to some key risk-drivers for a GMWB type of VA contract. In Chapter 6.3, analytical values for the liabilities on GMAB and GMDB VA contracts under the Heston stochastic volatility model will be derived.

1.4 Introduction to Monte Carlo valuation

The central mathematical concept which will form the basis of the liability valuation and risk-management techniques developed throughout this thesis is Monte Carlo (MC) simulation. An excellent resource which gives a complete overview on the application of the MC technique in a financial context is the textbook "Monte Carlo Methods in Financial Engineering" by Paul Glasserman [Gla03]. This text guides the reader from the basics of simulation through to applying the technique across a broad range of financial models and products for valuation and managing risk.

A review of some of the fundamental areas of MC simulation covered in this text will now be given. These topics are important in understanding all the subsequent chapters of this thesis and provide a solid background to some of the key concepts in MC simulation. It should also help illustrate how powerful this approach can be for estimating financial quantities and values, which will complement some of the ideas which are developed in later parts of the thesis.

Let us begin by stating succinctly what is meant by MC simulation. These methods are a class of computational algorithms that are based on repeated random sampling, often used when simulating physical and mathematical systems. They are useful for modelling phenomena with significant uncertainty in inputs, for example in finance for the calculation of risk or to value and analyze (complex) derivatives and portfolios. To do this we simulate, or mimic, the various sources of uncertainty that affect the value of the instrument or portfolio we are interested in and then calculate a representative value or risk-level given these possible values of the inputs (which will be described by the model(s) we choose to employ). A MC estimator for the general financial problem $\alpha = \mathbb{E}[p(S(\mathbf{Z}_i))]$ can be expressed as

$$\hat{\alpha} = \frac{1}{n}\sum_{i=1}^{n} p(S(\mathbf{Z}_i)). \tag{1.3}$$

Here, $\mathbb{E}[\cdot]$ represents the mean, or expected value operator. The function p gives the payoff, liability or risk-measure given a realisation of the behaviour of some underlying asset(s) modelled by $S(\mathbf{Z})$, which itself is a function of some source(s) of uncertainty $\mathbf{Z}$. The $\mathbf{Z}_i$ are n independent random vectors needed to evaluate the payoff along each simulation path $i = 1, \ldots, n$. These vectors could consist of uniform random variables, or from some other statistical distribution by simply transforming the uniform variates appropriately. Standard normal random variables, which are very popular in stochastic financial models, can be readily obtained from uniform variates using the Box-Muller transform, for example. To generate uniform random numbers a computer typically employs what is known as a pseudo-random number sequence, which is an algorithm for generating a sequence of numbers (which is deterministic once the initial state or seed value is chosen). The sequence generated mimics the behaviour of a sample drawn from a uniform distribution. There is also the possibility of using quasi-random number (or low discrepancy) sequences. This

is where sample points are systematically chosen so that they evenly fill a Cartesian grid according to a particular algorithm, with the aim of reducing the variance of any estimators calculated. This approach will be introduced in more detail at the end of this section.

One key advantage the MC method has over other techniques is the ability to work with problems consisting of a large number of sources of uncertainty (i.e., of high dimensionality). In such instances we essentially just have to generate an additional stream of random numbers with each additional dimension of the problem (some of which may be correlated to other uncertainty source's generated random number stream). This compares favourably with other numerical integration techniques, such as finite difference approximations, which typically break-down when the dimensionality of a problem becomes too large.

1.4.1 Sampling error and variance reduction

Given a MC estimator there are two issues which concern us. First, of course, there is the numerical value the estimator takes. Equally, importantly, however, is the uncertainty associated with this value. Let us consider a standard MC simulation to value some option or liability. Imagine n trials are performed, and the standard deviation of the n resultant simulated option prices is σ. Then the Central Limit Theorem implies that the standard (sampling) error for this MC simulation is given by

$$\text{SE} = \frac{\sigma}{\sqrt{n}}. \tag{1.4}$$

Notice that in order to reduce the sampling error by half we must quadruple the number of replications performed. This 'law of diminishing returns' means that as we seek greater accuracy using MC simulation, the number of scenarios we need to perform increases rapidly. This relatively slow convergence is one of the weaknesses of the MC simulation technique. Indeed, it has led many people to look for alternative methods which can help reduce the sampling error in MC simulation, without having to increase the number of replications used. These attempts come under the general name of variance reduction techniques and we shall now introduce a few of these approaches.

Antithetic Variates

Perhaps the easiest variance reduction technique to implement is known as antithetic variates. To introduce this method we consider the general financial Monte Carlo estimation problem from the beginning of this section and follow the discussion of Higham [Hig04]. The challenge was to estimate

$$\alpha = \mathbb{E}\left[p(S(\mathbf{Z}))\right], \tag{1.5}$$

where $\mathbf{Z}$ is a vector of standard normal variates. To simplify the illustration of the technique of using antithetic variates, let us assume the risk-driver, or shock, is one-dimensional and set $q(Z) := p(S(Z))$. Then, the natural estimator for α under a Monte Carlo simulation is simply given by

$$\hat{\alpha} = \frac{1}{n}\sum_{i=1}^{n} q(Z_i), \tag{1.6}$$

where Z_i are independent and identically distributed standard normal variates. On the other hand, the alternative antithetic variate estimator is given by

$$\hat{\alpha}^{\star} = \frac{1}{n}\sum_{i=1}^{n} \frac{q(Z_i) + q(-Z_i)}{2}. \tag{1.7}$$

Of course, if Z_i is a standard normal variate, then so to is $-Z_i$. Thus, this estimator is clearly unbiased. But why would using this estimator be likely to reduce the variance as compared to the estimate using the standard MC estimator? Well, the variance of the antithetic estimator is given by

$$\begin{aligned}
\mathrm{Var}\left(\frac{q(Z_i) + q(-Z_i)}{2}\right) &= \frac{1}{4}\Big(\mathrm{Var}(q(Z_i)) + \mathrm{Var}(q(-Z_i) + 2\mathrm{Cov}(q(Z_i), q(-Z_i))\Big) \\
&= \frac{1}{4}\Big(2\mathrm{Var}(q(Z_i)) + 2\mathrm{Cov}(q(Z_i), q(-Z_i))\Big) \\
&= \frac{1}{2}\mathrm{Var}(q(Z_i)) + \frac{1}{2}\mathrm{Cov}(q(Z_i), q(-Z_i)),
\end{aligned} \tag{1.8}$$

where $\mathrm{Cov}(A, B)$, denotes the covariance of the random variables A and B. Now, we assume that it takes approximately twice the computation time to simulate n antithetic paths than it does to simulate n standard paths. This ignores the potential overheads saved by simply multiplying half the shocks generated by -1 rather than

generating new random shocks for all paths. However, this saving will generally be small compared to the time taken to value the payoff function along each path, particularly for complex products, so it is probably fair to claim the antithetic estimator will take twice the time to simulate than the standard estimator. With this assumption, the antithetic estimator will reduce the variance if it has a smaller variance than the standard estimator with double the number of standard paths, i.e., if

$$\mathrm{Var}\left(\frac{q(Z_i)+q(-Z_i)}{2}\right) < \mathrm{Var}\left(\frac{1}{2n}\sum_{i=1}^{2n} q(Z_i)\right), \tag{1.9}$$

which, after substituting the above identity for the left hand side and evaluating the right-hand side, can be expressed as

$$\frac{1}{2}\mathrm{Var}(q(Z_i)) + \frac{1}{2}\mathrm{Cov}(q(Z_i), q(-Z_i)) < \frac{1}{2}\mathrm{Var}(q(Z_i)). \tag{1.10}$$

Thus, the antithetic estimator will have smaller variance than the standard estimator taking the same computation time if $\mathrm{Cov}(q(Z_i), q(-Z_i))$ is negative. A sufficient condition ensuring this is the case is for the payoff function $q(Z)$ to be monotonic in Z. Glasserman [Gla03] provides an argument that the technique will be even more successful in reducing variance for payoff functions which are close to linear in the variable Z. With monotonic payoff functions being commonplace in finance, using antithetic variates gives a fairly straightforward method in which the variance of the estimate from a Monte Carlo simulation can be reduced, whilst maintaining the number of simulations performed.

Example 1.1. By way of an example of applying the antithetic variates technique, let us estimate the price of a simple call option written on a underlying asset whose dynamics are governed by the Black-Scholes model. Firstly we shall approach this using a standard MC simulation, then we will look at also considering the antithetic path and the effect this has on the variance of the estimator of the price. Let $S(0) = 100$, $K = 105$, $\sigma = 0.2$, $r = 0.05$ and $T = 1$. The analytical price for this option is £8.02. In Figure 1.1 a box-plot is given showing the results of 500 different estimates of the option price, found by simulating 500,000 standard simulation paths, and 500 estimates found by simulating 250,000 antithetic pair paths. This should give a fair comparison of the two simulation approaches, as was discussed a moment ago. The results show that using antithetic variates reduces the variance in estimating

the price of this basic option. The spread of the 500 estimates around the mean is smaller for the antithetic variate approach than the standard approach, both in the full range and inter-quantile range. However, the reduction in variance achieved by this approach is generally not as large as other variance reduction methods when these are available. Some of these other variance reduction approaches will now be discussed.

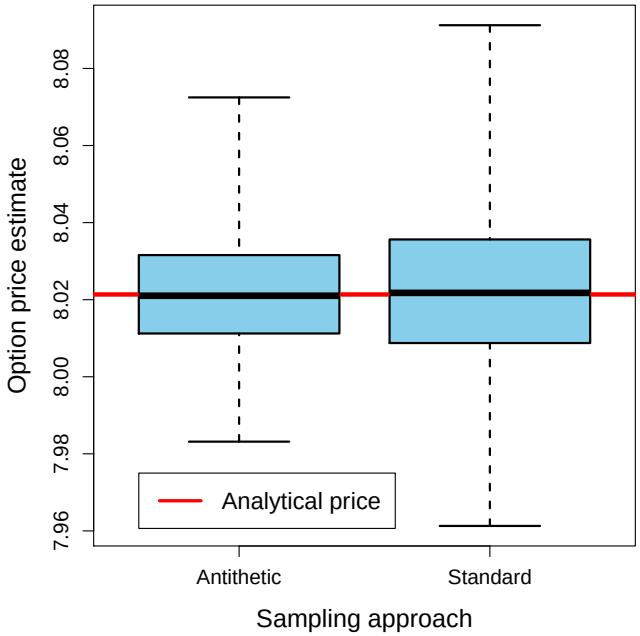

Figure 1.1: Box-plot for 500 estimates of the price of a call option under the Black-Scholes model, with and without the use of antithetic variates. For each of the 500 standard estimates, 500,000 asset price paths were simulated. For each of the antithetic estimates, 250,000 antithetic pair paths were simulated. Analytical price of this option is £8.02. Option parameters are given in the main text.

Control Variates

Another popular variance reduction method is to employ a control variate. This introduction will follow Section 4.1 of Glasserman [Gla03]. Under this approach the error around known exact quantities is used to reduce the error arising in estimating some unknown quantity. To make this method clearer and to explain how one would use the method in practice, let us summarise the basic underlying theory of MC simulation using some simple notation.

Recall, the standard financial MC set-up is to estimate $\alpha = \mathbb{E}[p(S(\mathbf{Z}))]$. Let us define $Y = p(S(\mathbf{Z}))$. We would then proceed by generating values $Y_1, \ldots, Y_n$ sampled from

the underlying density of the random variable Y. Let us then assume that on each replication we can calculate another quantity X_i, which depends on the shock $\mathbf{Z}_i$ and is such that $\mathbb{E}[X]$ is known. Then for any constant b, we can calculate

$$Y_i^{\text{CV}}(b) = Y_i - b(X_i - \mathbb{E}[X]) \tag{1.11}$$

for the i-th replication. Taking the average of these values over all replications,

$$\bar{Y}^{\text{CV}}(b) = \frac{1}{n}\sum_{i=1}^{n}(Y_i - b(X_i - \mathbb{E}[X])) = \bar{Y} - b(\bar{X} - \mathbb{E}[X]), \tag{1.12}$$

gives us the control variate (CV) estimator. This can be shown to be both unbiased and consistent. See Glasserman [Gla03] for details.

We now show how this estimator can reduce the sampling error in our MC simulation and how we choose the arbitrary constant b to minimise this error. Firstly we note that each $Y_i(b)$ has variance

$$\sigma^2_{\text{CV}}(b) = \text{Var}(Y_i - b(X_i - \mathbb{E}[X])) = \sigma_Y^2 - 2b\sigma_X\sigma_Y\rho_{XY} + b^2\sigma_X^2. \tag{1.13}$$

The CV estimator then has variance $\frac{\sigma^2_{\text{CV}}(b)}{n}$ and the standard MC estimator has variance $\frac{\sigma_Y^2}{n}$. The CV estimator then has lower variance than the standard estimator if

$$b^2\sigma_X^2 < 2b\sigma_X\sigma_Y\rho_{XY}. \tag{1.14}$$

Therefore, inequality 1.14 gives us a condition which ensures the CV estimator will reduce variance. But what if we have a CV estimator which satisfies this condition for many different values of b? In this case we can find the value of b which minimises the variance of the CV estimator. Minimising Equation 1.13 with respect to b using simple calculus yields the value b^* which minimises the variance of this estimator as

$$b^* = \frac{\sigma_X\sigma_Y\rho_{XY}}{\sigma_X^2} = \frac{\text{cov}[X,Y]}{\text{Var}[X]}. \tag{1.15}$$

Setting $b = b^*$ gives the minimum variance which can be achieved using the CV

estimator involving the random variable X as

$$\mathrm{Var}[b^*] = \mathrm{Var}[Y] - \frac{(\mathrm{cov}[X,Y])^2}{\mathrm{Var}[X]}. \tag{1.16}$$

The ratio of the minimum variance CV estimator to the standard MC estimator variance is then $1 - \rho_{XY}^2$.

Therefore, the success of a CV estimator in reducing sampling error is related to the correlation between the control variable and the quantity of interest Y. This, of course, makes sense as we are using the knowledge about the error in estimating $\mathbb{E}[X]$ to refine our estimate of $\mathbb{E}[Y]$, thus the stronger the correlation (or anti-correlation) between X and Y, the more success the control will have. To find the coefficient b^*, however, we need to know σ_Y and ρ_{XY}, which is unlikely if we do not know $\mathbb{E}[Y]$ (and finding this is the goal of MC simulation). One should still, however, obtain a reasonably successful control variate if we estimate b using sample estimates for σ_Y and ρ_{XY}, despite this estimation introducing some bias.

Example 1.2. Let us consider an example which demonstrates a practical application of the control variate approach, which was originally proposed by Kemna and Vorst [Kem90]. This example is also relevant to the research in the thesis, as it demonstrates an application of a variance reduction technique for a path-dependent option payoff. This is illustrative in thinking about how one could construct a variance reduction technique for complex insurance liability estimators.

For this example, let us introduce an arithmetic and a geometric Asian option. The discounted payoff on an arithmetic Asian call option is given by

$$e^{-rT} \max\left(\frac{1}{T}\int_0^T S(t)dt - K, 0\right) \tag{1.17}$$

and the discounted payoff of a geometric Asian call option is given by

$$e^{-rT} \max\left(\exp\left(\frac{1}{T}\int_0^T \ln(S(t))dt\right) - K, 0\right). \tag{1.18}$$

We note that there does not exist a closed-form solution for the price of an arithmetic Asian call option (under the Black-Scholes model), thus some numerical method must be used to price this class of derivative. One such approach is MC simulation, but what could one use as a good control variate in this case? Although no analytical

formula for the price of an arithmetic Asian option is available, there does exist an analytical formula for the price of a geometric Asian call option. Furthermore, the prices of arithmetic and geometric Asian options are likely to be highly correlated, which suggests the difference between the simulation estimate and analytical value of a geometric Asian option price will be a successful control variate in finding the price of an arithmetic Asian option. Table 1.1 gives some results of implementing a MC algorithm to find the price of this type of option with and without the use of the proposed control variate. We shall not discuss the intricacies of the method further here. However, the results show that for this type of option, employing a CV can significantly reduce the sampling error in the MC estimates.

$S(0)$	K	σ	r	T	95% C.I. without CV	95% C.I. with CV
100	100	0.2	0.05	1	(5.60,5.91)	(5.80,5.82)
100	110	0.2	0.05	1	(1.89,2.09)	(2.02,2.03)
100	90	0.2	0.05	1	(12.37,12.78)	(12.63,12.64)
100	100	0.4	0.05	1	(9.79,10.43)	(10.18,10.24)
100	100	0.2	0.05	2	(8.50,8.97)	(8.73,8.76)

Table 1.1: Estimates of the price of an Arithmetic Asian call option with and without the use of the proposed control variate in the example. All option settings use 10,000 paths and 100 timesteps. Results shown as 95% confidence intervals.

Importance Sampling

The final variance reduction method which we shall outline is slightly more complex than the previous techniques. This method is known as importance sampling and the following introduction is based on Section 4.6 of Glasserman [Gla03]. The idea of importance sampling is to reduce variance by changing the probability measure which the paths are sampled from. Essentially this method places more weight on the paths which are 'important' and thus the efficiency of sampling will be increased. Successfully utilising this method requires a good understanding of the dynamics of the model for the underlying asset and the option payoff function, therefore it can often be difficult to apply for complex problems.

To introduce the technique, consider the general problem of estimating

$$\alpha = \mathbb{E}[p(S(\mathbf{Z}))] = \mathbb{E}[q(\mathbf{Z})] = \int q(\mathbf{z}) f(\mathbf{z}) d\mathbf{z} \tag{1.19}$$

with $\mathbf{Z} \in \mathbb{R}^d$ a random vector (perhaps a random variable) with density f and $q : \mathbb{R}^d \to \mathbb{R}$ some function. The standard MC estimator we introduced earlier in this section would then be

$$\hat{\alpha} = \frac{1}{n} \sum_{i=1}^{n} q(\mathbf{Z}_i) \tag{1.20}$$

with the $\mathbf{Z}_i$ independent samples from the density f. Now consider an alternative density g on $\mathbb{R}^d$ which is absolutely continuous with respect to f, i.e., $g(\mathbf{z}) = 0 \implies f(\mathbf{z}) = 0$, for all $\mathbf{z} \in \mathbb{R}^d$. Then α can be expressed in the form

$$\begin{aligned} \alpha &= \int q(\mathbf{z}) \frac{f(\mathbf{z})}{g(\mathbf{z})} g(\mathbf{z}) d\mathbf{z} && (1.21) \\ &= \mathbb{E}_g \left[q(\mathbf{z}) \frac{f(\mathbf{Z})}{g(\mathbf{Z})} \right] && (1.22) \end{aligned}$$

where $\mathbb{E}_g$ indicates the expectation is taken with respect to the density g.

If the $\mathbf{Z}_i$ are now independent samples from density g (rather than our original density f), the importance sampling estimator is given as

$$\hat{\alpha}_g = \frac{1}{n} \sum_{i=1}^{n} q(\mathbf{Z}_i) \frac{f(\mathbf{Z}_i)}{g(\mathbf{Z}_i)}. \tag{1.23}$$

The quantity or weight $\frac{f(\mathbf{Z}_i)}{g(\mathbf{Z}_i)}$ is the likelihood ratio evaluated at $\mathbf{Z}_i$.

One can show that $\hat{\alpha}_g$ is an unbiased estimator of α, thus in considering the success in variance reduction of importance sampling we can just compare second moments of the estimator with and without this technique. The second moment of the importance sampling estimator is

$$\mathbb{E}_g \left[\left(q(\mathbf{Z}) \frac{f(\mathbf{Z})}{g(\mathbf{Z})} \right)^2 \right] = \mathbb{E} \left[q(\mathbf{Z})^2 \frac{f(\mathbf{Z})}{g(\mathbf{Z})} \right] \tag{1.24}$$

which may be greater or smaller than the standard MC estimator's second moment $\mathbb{E}[q(\mathbf{Z}^2)]$. Indeed, this can even be infinitely larger or smaller with certain choices of g. Thus, the choice of the importance sampling density is crucial to the success of this variance reduction method and one must take great care when using this approach.

If q is taken as the indicator function for some set A, then $\alpha = \mathbb{P}(\mathbf{Z} \in A)$, and the (theoretical) zero-variance importance sampling density is $q(\mathbf{z})f(\mathbf{Z})/\alpha$, i.e., the

conditional density of $\mathbf{Z}$, given $\mathbf{Z} \in A$. Thus, when this method is applied in order to estimate a probability we should seek to find a density which is similar to the conditional density. Put more simply, we should choose our importance sampling density to make g more likely. This concept is commonly used in finance to reduce variance when the set A is rare under f, for example the event of a large number of obligors defaulting in credit risk management modelling.

Example 1.3. As a very simple example of importance sampling let us attempt to estimate $\alpha = P(Z > 4)$, where Z is a standard normal distribution. To estimate this naively by MC simulation, first 1,000,000 standard normal random variates are simulated and from this sample the standard MC estimator gives a 95% confidence interval for γ of $(-1.94 \times 10^{-3}, 1.98 \times 10^{-3})$. This corresponds to a standard error of 1×10^{-3}. Instead, one could generate normal random variates with mean 4 and variance 1, density g say, and employ the likelihood ratio between this and the standard normal density. This likelihood ratio is given by

$$\frac{f(z)}{g(z)} = \frac{e^{-z^2/2}}{e^{-(z-4)^2/2}} = e^8 e^{-4z} \tag{1.25}$$

evaluated at each of the sampled values from g. With this approach we obtain a 95% confidence interval for γ of $(3.156 \times 10^{-5}, 3.182 \times 10^{-5})$, or, an estimate of 3.169×10^{-5} with a standard error of 6.73×10^{-8}. This compares well to the value obtained from accurate statistical tables, i.e., $1 - \Phi(4) = 3.167 \times 10^{-5}$. Also, there is clearly a dramatic reduction in variance using the importance sampling estimator. Of course, this is a hypothetical example and more efficient methods for calculating γ exist. However, it demonstrates that if we can utilise some knowledge of the model or payoff being considered, the increase in accuracy of a MC estimator for certain problems can be vast.

Quasi-random sampling

Quasi-random (or low-discrepancy) simulation algorithms have a different philosophy in comparison to standard MC simulation. Pseudo-random sampling algorithms aim to mimic randomness by computing a long list of numbers from a deterministic sequence which appear very much as if they are uniformly sampled variates. With quasi-random sampling, however, the aim is not to mimic randomness, but to obtain

an increased accuracy in MC estimators "specifically by generating points which are too evenly distributed to be random" (Glasserman [Gla03]). These methods could potentially offer convergence in these estimators of up to $O(1/n)$, compared to the standard MC simulation convergence of $O(1/\sqrt{n})$. The variance reduction techniques which have been discussed in this section will only decrease the implicit constant in the $O(1/\sqrt{n})$ convergence. Thus, the quasi-random sampling approach can be very rewarding in increasing the efficiency of the MC simulation technique for certain financial problems.

To further illustrate the idea behind quasi-random sampling, consider again our general financial MC estimator $\hat{\alpha}$. This can be expressed as

$$\hat{\alpha} = \frac{1}{n}\sum_{i=1}^{n} p(S(\mathbf{Z}_i)) = \frac{1}{n}\sum_{i=1}^{n} p(S(\Theta^{-1}(\mathbf{U}_i))) = \frac{1}{n}\sum_{i=1}^{n} r(\mathbf{U}_i), \tag{1.26}$$

where we have expressed the d-dimensional random normal vector $\mathbf{Z}_i$ in terms of a d-dimensional vector of uniform random variates using the inverse distribution function, $\Theta^{-1}(z)$, of the random variable z. When being applied to a vector, we assume this distribution function will act on each element, returning another vector. We have set $r(x) = p(S(\Theta^{-1}(x)))$ to simplify the notation. A general quasi-random MC estimator for α is then given by

$$\hat{\alpha} = \frac{1}{n}\sum_{i=1}^{n} r(\mathbf{x}_i) \tag{1.27}$$

for systematically chosen points $\mathbf{x}_1, \ldots, \mathbf{x}_n$ in the unit hypercube $[0,1]^d$. The algorithm for choosing these points is designed to sample as uniformly as possible over $[0,1]^d$, where $d = \dim(\mathbf{U}_i)$ for each random number stream i.

Of course, pseudo-random sampling aims to generate these points such that they mimic a randomly generated sequence. With the quasi-random estimator this systematic sampling of points is given by a deterministic sequence which aims to fill the unit hypercube as uniformly as possible. A naive approach might be to sample these points in a regularly spaced grid, however this approach suffers from a couple of major shortcomings, which will be outlined later in Section 2.4. Instead, algorithms have been developed which generate sequences which aim to minimise the discrepancy of a set of points in the d-dimensional hypercube. Discrepancy is

a mathematical notion capturing the deviation from uniformity of a set of points – this concept is formalised mathematically in Glasserman [Gla03].

To give some insight into how these quasi-random (or low discrepancy) sequences are generated, let us investigate a quasi-random sequence in one-dimension over the unit interval. This deterministic algorithm generates what is known as a Van der Corput sequence. Let us set an integer $b \geq 2$ which we will call the base. Every positive integer has a unique representation as a linear combination of non-negative powers of b with coefficients, $a_j(k)$, in the set $0, 1, \ldots, b-1$. This can be expressed as

$$k = \sum_{j=0}^{\infty} a_j(k) b^j \tag{1.28}$$

for every $k \in \mathbb{N}$, with all apart from a finite number of the coefficients being equal to zero.

Introducing the radical inverse function ϕ_b as a mapping from $\mathbb{N}$ to $[0, 1)$. For each positive integer k it will flip the coefficients of k about the base-b "decimal" point to obtain a base-b fraction $.a_0 a_1 a_2 \ldots$ More formally, this function is defined by

$$\phi_b(k) = \sum_{j=0}^{\infty} \frac{a_j(k)}{b^{j+1}}. \tag{1.29}$$

The base-b Vand der Corput sequence is then given as $0 = \phi_b(0), \phi_b(1), \phi_b(2), \ldots$ An example of the first few terms of the base 2 Vand der Corput sequence is given in the Table 1.2 (which is reproduced from Glasserman [Gla03]).

A naive refinement might just add the new values in increasing order, for example if the sequence already consisted of 0, 1/4, 1/2 and 3/4, the next terms would be (in

k	K(binary)	$\phi_2(k)$(binary)	$\phi_2(k)$
0	0	0	0
1	1	0.1	1/2
2	10	0.01	1/4
3	11	0.11	3/4
4	100	0.001	1/8
5	101	0.101	5/8
6	110	0.011	3/8
7	111	0.111	7/8

Table 1.2: First few terms of base-2 Van der Corput sequence.

order) 1/8, 3/8, 5/8 and 7/8. The Van der Corput sequence adds these in a balanced way, appearing on alternating sides of 1/2, first of all, and then on alternative sides of 1/4 and 3/4. This property continues as the sequence extends and fills the unit interval with greater detail. The base, b, parameter has the property that the larger the value it takes, the greater the number of points which are required to achieve uniformity.

This simple Van der Corput sequence illustrates the basic idea behind a quasi-random number sequence. Quasi-random sequences for multi-dimensional problems extend this approach to generate points which have low-discrepancy over the multi-dimensional unit hypercube. In the analysis later in this thesis, we generate quasi-random points in multiple dimensions using the Sobol sequence. For a detailed introduction to the Sobol sequence see Section 5.2.3 of Glasserman [Gla03]. Glasserman also shows that the uniformity property of quasi-random sequences is much more apparent in the lower dimensions of a high-dimensional problem. Thus, it may be beneficial to use knowledge of the underlying stochastic model in conjunction with quasi-random sampling to improve accuracy. In other words, we should use the dimensions offering the greatest uniformity for the most important risk-drivers.

1.4.2 Summary of the MC technique in finance

In summary, this section has hopefully given a brief, but informative account of some of the fundamental principles of the MC simulation technique. This should provide enough of a background of the method to understand the research which will be developed throughout the thesis and how such innovations would be implemented in practice. There are many textbooks which give a more extensive introduction to the application of MC simulation in finance. Some references are Glasserman [Gla03], Higham [Hig04] or Huynh, Lai and Soumaré [Huy08]. We will now move on to introduce a novel extension of the MC technique which has been proposed as a solution to the challenge of calculating a solvency capital requirement for complex insurance liabilities. Investigating this approach is the purpose of Part I of the thesis.

Part I

LSMC method for insurance liability projection

Chapter 2

Introduction to LSMC

2.1 Idea behind the Least-Squares Monte Carlo (LSMC) method

In Part I of the thesis the least-squares Monte Carlo (LSMC) method for projecting insurance liabilities will be investigated. This approximation technique could prove very useful for practitioners in the insurance industry looking for an efficient approach for calculating a solvency capital requirement (SCR) under the upcoming Solvency II regulatory framework. To begin with, a general overview of the LSMC simulation technique for the calculation of an insurance SCR will be given. Under the impending Solvency II regulatory framework, insurance companies operating inside the European Union will be required to calculate a SCR as part of their risk management practice. As was discussed in Section 1.2 this can either be done using an industry-wide standard formula or by an insurer employing their own model which takes into account the specific risks they face. This thesis will concentrate on the more sophisticated of these approaches by employing an economic scenario generation methodology. This is much more aligned with the philosophy of Solvency II and is the method most of the large insurers are considering in their SCR calculations.

To aid this introduction, let us imagine we are an insurer wishing to calculate an accurate and reliable SCR over some future time horizon. This more sophisticated approach proceeds by accurately calculating the projected liability distribution over this future time horizon and then determining some percentile of this distribution corresponding to the confidence level at which the SCR is required. There are two stages to this calculation; firstly, we would look to project a number of realisations of the key economic variables, or risk-drivers, on which the future liability is likely to depend. This idea is often known as economic scenario generation. After these future states of the economy have been simulated, the liability the insurer faces at the projection date will be calculated, conditional on each of these simulated future states. With many insurance products the liabilities faced are complex and path-dependent in nature. As such, there is usually no simple, analytic closed-

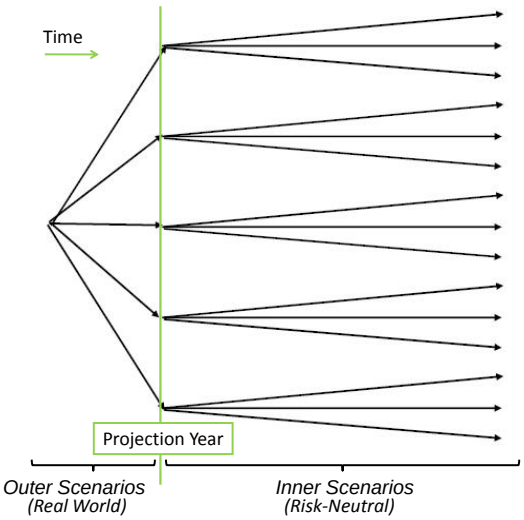

Figure 2.1: Nested simulation.

form valuation formula for these liabilities. Furthermore, these liabilities are often multi-dimensional in nature, meaning the natural approach to valuing the liabilities efficiently is to use the technique of Monte Carlo simulation, introduced in Section 1.4. Thus, this challenge of projecting insurer's liabilities has naturally led to 'nested' or 'stochastic-on-stochastic' simulation. This can be summarised as a simulation with a number of 'outer' economic scenarios each branching out to form a number of 'inner' valuation scenarios for calculating the (conditional) liability. A schematic of this nested simulation set-up is given in Figure 2.1.

One important point to note is the different approaches for calibrating the models from which we simulate these outer economic projections and inner liability valuation scenarios. For the outer scenarios, we wish to employ a model which will realistically capture the behaviour of the key risk-drivers from today out until the projection date. This chosen model will be calibrated to some historical data to attempt to best capture the likely observed path of these variables over the period of time until the projection date. This is often known as a real-world model calibration. On the other hand, the inner liability valuations are calibrated using a completely different philosophy. Here, the valuation model will not be calibrated using the observed behaviour of key economic variables, but instead calibrated in such a manner as to ensure that the expected value of all financial assets in our economy will grow at the same rate. This rate of growth will be the return one would achieve by investing in a risk-free asset, which we assume is available for

investment in this economy. Calibrating the model in such a manner avoids the possibility of arbitrage opportunities, where an investor could begin with zero wealth and expect to make a risk-free profit. This type of model calibration is often known as risk-neutral, or sometimes market-consistent. There has been a vast amount of literature over the last forty years, developing the risk-neutral approach to pricing financial instruments. Some textbooks which introduce this theoretical framework are Bingham and Kiesel [Bin04], Baxter and Rennie [Bax96] and Wilmott [Wil00]. In practice, market-consistent calibrations are typically obtained by fitting these models to quoted market prices, since any arbitrage opportunities would be quickly eliminated and priced out of a liquid market. With the terms of insurance products often extending many decades into the future, obtaining a market-consistent calibration would require quoted prices for options with a similar maturity. Such options are very illiquid and this leads to complications in attempting to obtain risk-neutral valuations of insurance liabilities. An excellent reference discussing such issues and proposing methodologies for making progress in the face of this challenge is the article by Pelsser [Pel11]. For the moment, however, let us assume that a satisfactory risk-neutral calibration can be achieved for the liability valuation model under consideration.

In this SCR calculation, the number of outer scenarios needed to give a reliable estimate will, of course, depend on the confidence level the insurer wishes to have in determining the amount of risk-based capital to hold. As mentioned in Section 1.2, typical SCR confidence levels are reasonably far into the upper-tail of the required capital distribution, often at the 99.5-th percentile. In order to obtain an accurate estimate of such a percentile of the distribution, a large number of outer economic scenarios must be simulated. This will ensure we have enough resolution in the upper tail to get an accurate estimate of this percentile through the calculation of simple estimators based on upper order statistics. Also, given that insurance liabilities are often long-term, complex, path-dependent and multi-dimensional in nature, obtaining an accurate liability valuation (conditional on each of the outer economic scenarios) will require a large number of inner valuation scenarios to be simulated. Thus, the total number of liability valuations required for an accurate SCR calculation under the full nested simulation framework is prohibitively large.

To illustrate the computational challenge of nested simulation let us specify some numbers: At the typical confidence level of 99.5% (over a one year time horizon) a minimum of 10,000 outer scenarios would be necessary to get reasonable resolution in the upper tail of the projected liability distribution. And within each of these outer scenarios, a minimum of around 10,000 inner scenarios are required to obtain an accurate estimate of the value of the liability. Thus in total around $10,000 \times 10,000 = 100,000,000$ valuation scenarios are needed. Furthermore, this simulation is only for a one-year time horizon. If an insurer wished to project n years into the future, a total of approximately $n \times 10,000 \times 10,000 = 100,000,000n$ scenarios would be required. An illustration of how this nested simulation extends to multiple projection dates is given in Figure 2.2. Bauer, Bergmann and Reuss [Bau11] analyse the optimal outer and inner scenario allocation for a SCR calculation in the nested simulation framework. They argue that reasonably accurate estimates can be obtained with a smaller number of inner scenarios than this. However, given the long-term and complex nature of typical insurance liabilities, the required number of valuation scenarios would still be infeasible given current levels of technology. Thus, the full nested simulation approach is not computationally practical for insurers.

In order to make this simulation framework computationally tractable, one approach would be to dramatically decrease the number of (conditional) inner valuation scenarios, given each of the outer economic projections. Let us imagine reducing this number of inner valuation scenarios from 10,000 to just a few, perhaps even just one. This will, of course, give a liability valuation of very poor accuracy. However, if we regress each of these poor single inner scenario liability estimates on some key risk-drivers which influence the liability at the projection date, then the accuracy of each of these estimates can be vastly improved. Essentially, we are using the cross-sectional information from across all these inaccurate single inner scenario valuations to correct the estimate for each of these in isolation. By employing a least-squares regression to improve the accuracy of the liability estimates at each projection date within this reduced simulation framework, far more accurate and computationally efficient estimates of the insurance SCR can be achieved. This technique will be referred to as the least-squares Monte Carlo (LSMC) method and investigating it further will form a large part of the content of this thesis. A schematic representation of the LSMC method is given in Figure 2.3.

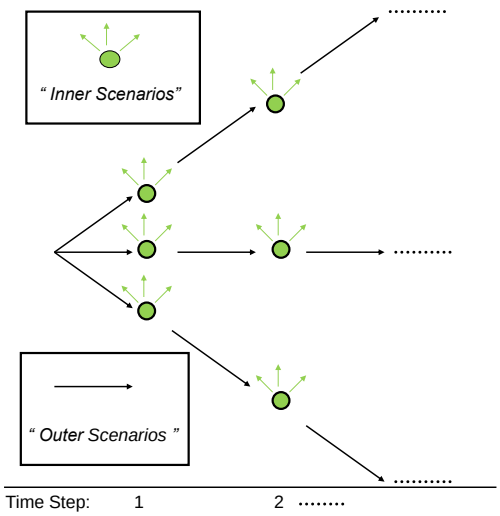

Figure 2.2: Nested simulation (a multi-year projection).

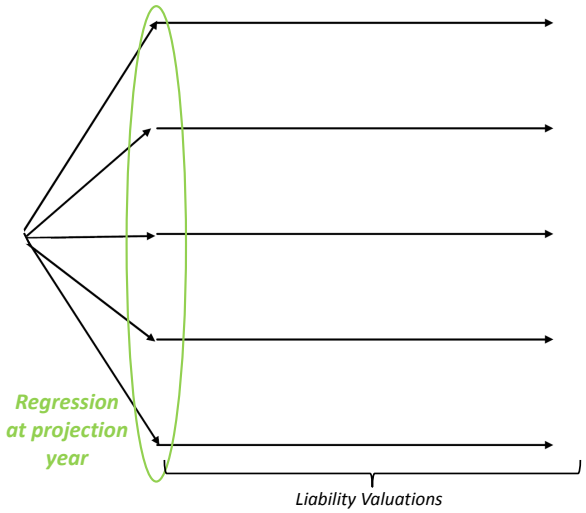

Figure 2.3: A schematic representation of LSMC simulation.

The LSMC was originally proposed in the context of American option pricing and developed by Carriere [Car96], Tsitsiklis and Roy [Tsi99] and Longstaff and Schwartz [Lon01]. It is worthwhile taking a short digression from the problem of projecting insurance liabilities to look at how LSMC is used to price American options. This will help further illustrate the approach in our context and also show the method being applied in another important area. Furthermore, many results given in the literature for the LSMC method in this alternative context may extend to our problem of projecting complex insurance liabilities.

2.2 LSMC for American option valuation

Let us attempt to value an American put option written on some stock with expiry in two years time by means of Monte Carlo simulation. The buyer of the option can, however, exercise this option at the end of the first year and at expiry only. This type of option is often denoted a Bermudan Option. In reality, in order to approximate the value of an American option, one would consider a Bermudan option with a large number of exercise opportunities before expiry. However, for illustrative purposes we shall just allow the one additional exercise opportunity at the end of the first year. The first stage in this simulation task is to generate a number of scenarios for the underlying stock price and calculate the resulting cashflows that arise on exercising the option. Firstly we need to project forward to the first available exercise opportunity (in our case at the end of the first year). However, at this stage the cashflow is dependent on whether or not the option is exercised. If the option is in the money the holder has the choice between exercising or holding the option until a later time, believing that the stock price will fall further. Of course the choice to continue holding the option has value. Therefore, in deciding whether or not to exercise the option when it is in the money at the end of the first year, the holder will compare the amount they will receive from exercising the option early with the (estimated) value from continuing to hold on to the option until expiry. The choice which provides the greatest value will dictate the action of the option holder.

In order to determine the continuation value of the option, we must generate a set of 'inner scenarios' branching from each outer scenario (taking us from initially to the end of the first year). This approach was originally proposed by Broadie and Glasserman who refer to it as a 'simulated tree' [Bro97]. Comparing this with the problem of projecting market-consistent balance sheets, we see both simulation challenges are very similar. In pricing an American option, both outer and inner scenarios are simulated under a risk-neutral model, though. Naturally, to price an American option under this approach, one would take a large number of exercise opportunities and working backward from maturity apply this method to determine whether the expected continuation value was greater than the early exercise value at each of these timesteps. The final value, corresponding to time zero, would then be the LSMC approximation for the price today of the American option.

2.3 LSMC framework/algorithm

Let us return to the challenge of projecting market-consistent insurance liabilities using the LSMC method. In this section, however, we will take the general concept of the LSMC method outlined in Section 2.1 and formalise it in a mathematical framework. This will give a clearer understanding of the details of the method and be crucial in later parts of the thesis where the technique will be thoroughly investigated and the basic idea of the method extended.

The following notation and framework will be based on that used to describe the LSMC method in Longstaff and Schwartz [Lon01], however here the algorithm will be presented in our insurance context, rather than in the pricing of American options. The market model we shall work under is a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$, where Ω is the set of all possible realisations of the random economy over a time period $[0, T]$, $\mathcal{F}$ is the sigma field of events occurring throughout $[0, T]$ and $\mathbb{P}$ is the 'real-world' probability measure defined on the elements of $\mathcal{F}$. We define $\mathcal{F}_t$ to be the filtration generated by the price processes for securities and underlying economic models in this economy until time t. We assume that $\mathcal{F}_T = \mathcal{F}$. Consistent with the notion of there being no arbitrage in this economy, we assume there exists an equivalent martingale measure $\mathbb{Q}$ for this economy (under which discounted stock prices are martingales).

Now, we want to construct an algorithm to give a pathwise approximation to the future liabilities associated with some insurance product at the projection time. Firstly let us set $t_1, t_2, \ldots$ as the points in time where the underlying fund level of the product is re-balanced due to the investments made with the funds since the last re-balancing date and/or when a withdrawal is made from this fund. Now define the cashflows from the insurer's reserves to the policyholder at time t as $C(\omega, t)$, which will be non-zero only if the guaranteed income due to the policyholder exceeds the underlying fund level at time t. The ω indicates these cashflows are different for each scenario. The future liabilities associated with the product at time t_k can then be expressed as:

$$L(\omega; t_k) = E_{\mathbb{Q}}\left[\sum_{j=k+1}^{K} \exp\left(-\int_{t_k}^{t_j} r(\omega, s)ds\right) C(\omega; t_j)\Big|\mathcal{F}_{t_k}\right], \tag{2.1}$$

where $r(\omega, t)$ is the riskless discount rate and the expectation is taken conditional on $\mathcal{F}_{t_k}$, or the information available about the product (e.g., fund and guaranteed income levels) and the underlying economy (e.g., interest rates) at time t_k. This information would be available in the outer simulation stage of the nested simulation set-up. The problem is how do we estimate this conditional expectation?

The LSMC method uses least-squares to obtain an approximation for this conditional expectation function at time t_k. At this time the assumption is made that the unknown functional form of $L(\omega; t_k)$ can be approximated by a linear combination of a countable set of $\mathcal{F}_{t_k}$-measurable basis functions. Longstaff and Schwartz give a formal justification of this assumption. Possible choices of basis functions include (weighted) Laguerre, Hermite, Legendre and Chebyshev orthogonal polynomials. Longstaff and Schwartz state that simply using powers of the state variables as basis functions also gives accurate results. A more in depth discussion of possible choices of basis functions in the LSMC method is given in Section 2.5. Some tests investigating the accuracy of the LSMC estimates under different choices of basis functions are given in Section 3.8.

By choosing an appropriate set of basis functions, $B_1(\mathbf{X}), B_2(\mathbf{X}), \ldots$, the liability value can then be expressed as

$$L(\omega; t_k) \approx \sum_{q=0}^{Q} a_q B_q(\mathbf{X}(t_k)) \tag{2.2}$$

where the a_j are constants and $\mathbf{X}$ is the vector of appropriate explanatory variables (which are found to influence the expected future liabilities of the product). If the LSMC method is used in multi-dimensional problems one may expect the number of basis functions (including cross-terms) to grow exponentially with an increasing number of state variables. Longstaff and Schwartz cite research which shows this number may not increase at such a rate, making the method robust to multi-dimensionality.

With this LSMC approach to the liability valuation, we can significantly reduce the number of inner scenarios required for each outer scenario projection, perhaps to even just a single inner scenario. This computational task is then easily manageable by a standard desktop computer. Taken alone each single inner scenario will provide a poor estimate of the true liability value. However, by regressing over a large

number of these single scenario estimates, a refined, more accurate liability valuation for each economic projection can be achieved.

Within this LSMC framework there is still a lot of flexibility in how the technique is implemented in practice. In Sections 2.4-2.6 some of the main choices one has in configuring this approach for the calculation of an insurance SCR will be discussed.

2.4 LSMC fitting scenario sampling

The first issue in the implementation of the LSMC method which we will discuss is the approach for sampling in the outer scenarios. Ultimately we will generate outer scenarios using a realistic real-world model and then feed these into our calculated regression function. This will generate liabilities which are consistent with how we believe the economic risk-drivers will behave until the projection date. In order to obtain an accurate fit for this regression function, however, we can sample or structure the outer scenarios in any way we like. These scenarios are used to calibrate the regression model. We will refer to these as outer fitting scenarios to distinguish them from the real-world outer scenarios.

For these outer fitting scenarios we can simply choose where to place the points to fill the space spanned by the risk-drivers which drive the projected value/liability, or sample them from any statistical distribution we choose. Therefore, how best to distribute these outer fitting scenario points in space is a key issue in the configuration of the LSMC method.

To explore this idea further, let us set-up an interval such that it is extremely unlikely that the risk-driver will take a value at the projection time outside of these given values. For clarity, let us assume there are only two factors which drive the future liabilities measured at projection time (at end of year one); these are the equity level at year one and the volatility level at year one, say. Now, we make the assumption that it is extremely unlikely that by the end of year one the equity return will not lie within the interval [0.3, 1.8] and the volatility within the interval [0, 50%]. This notion could be generalised by setting this interval as the range of values the risk-drivers will take with some extremely high level of confidence, say 99.9%. These intervals then define the limits over which we wish to determine a regression surface in the two risk-drivers.

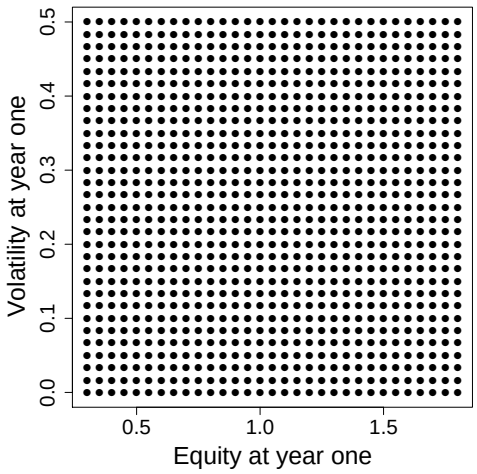

Figure 2.4: Full (discrete) grid sampling (961 fitting points shown).

2.4.1 Full (discrete) grid sampling

Probably the most natural method under which to sample points from this finite two-dimensional space is to create a regularly-spaced discrete grid. If we have 961 (= 31x31) outer fitting points, they will fill our two-dimensional space, as is shown in Figure 2.4. The main problem with this approach is how the number of outer fitting points grows as the number of risk-drivers increases. This is often referred to as the curse of dimensionality. Even with just four risk-drivers, choosing 31 points in each risk-driver would yield $31^4 = 923,521$ outer fitting points. Furthermore, as the number of risk-drivers or discretisation points is increased further, the number of fitting points increases rapidly.

2.4.2 Latin hypercube sampling

One approach which does not suffer from this problem with many risk-drivers is Latin hypercube sampling (LHS). Under LHS each risk-driver interval is partitioned into strata and the outer points are sampled in such a way as to ensure that exactly one outer point will feature in each stratum, for each of the given risk-drivers. More technically, the stratification only takes place in the one-dimensional marginals of a multi-dimensional joint distribution. This sampling approach does not pose problems in high dimensional problems, as the total number of points (and hence strata used) is independent of the number of risk-drivers. For clarity, we show how

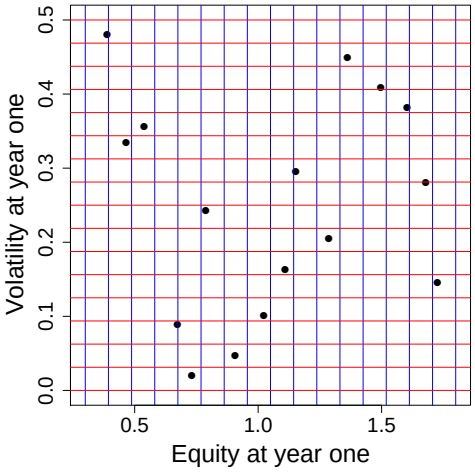

Figure 2.5: Latin Hypercube sampling (16 fitting points shown).

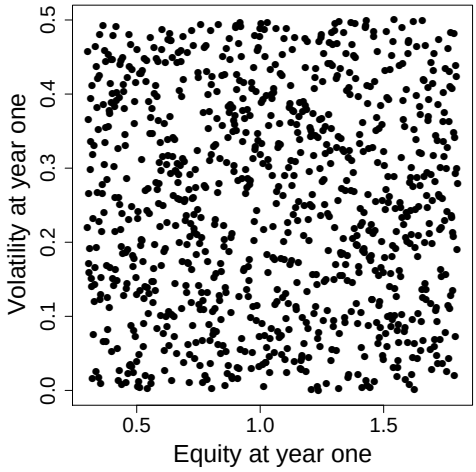

Figure 2.6: Latin Hypercube sampling (961 fitting points shown).

16 outer fitting points would fill our finite two-dimensional space in Figure 2.5. The strata are shown alongside the points in this plot. The results of generating 961 points by LHS in our two-dimensional risk-driver space is then shown in Figure 2.6.

2.4.3 Quasi-random sampling

Another outer fitting scenario sampling approach which is robust to a large number of risk-drivers is to sample using quasi-random (or low-discrepancy) numbers. These were mentioned briefly in Section 1.4. The idea in sampling with this method is not to mimic randomness, as is the case in generating the pseudo-random numbers used in standard Monte Carlo simulation. Rather an increased level of accuracy is sought, by generating samples which are too evenly distributed to be random. For a more detailed introduction, see Glasserman [Gla03]. An example of filling the space spanned by the year one equity and volatility intervals with 961 two-dimensional Sobol numbers is shown in Figure 2.7.

2.4.4 Uniform (pseudo-random) sampling

A pseudo-random number generator could also generate 961 two-dimensional points from the space spanned by the chosen risk-driver intervals. This is just the standard uniform distribution sampling scheme over this space and gives another method by which we could sample points in the outer fitting scenarios. Naturally, this will

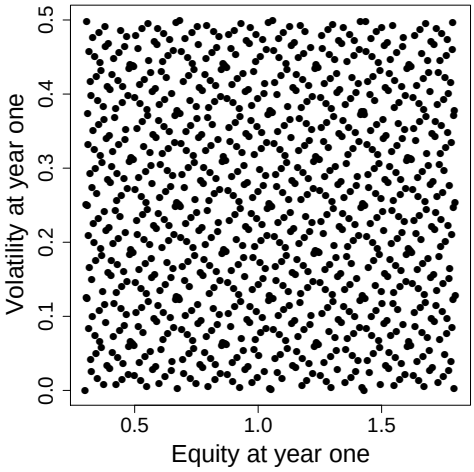

Figure 2.7: Quasi-random sampling (with 961 fitting points shown).

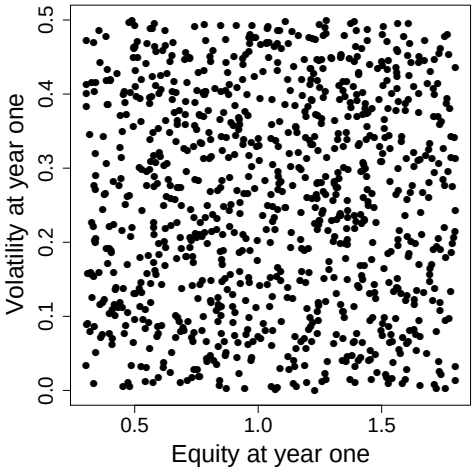

Figure 2.8: Uniform sampling (with 961 fitting points shown).

fill the space with a greater discrepancy than generating the points using a Sobol sequence. However, it is not clear that generating outer fitting points with this low discrepancy property will necessarily improve the fit of the resultant regression function. A graph of 961 standard uniform sampled year one equity and volatility pairs is given in Figure 2.8.

In this section we have shown a few different approaches for sampling the outer fitting scenarios in the LSMC method. In Sections 3.7 and 3.8, these different approaches will be investigated and some tests determining the accuracy of the LSMC estimates under these different sampling methods will be performed. For the moment, the key point to note is that the scenarios which are used to calibrate the regression model need not necessarily be drawn from the real-world probability model which we believe to realistically describe the evolution of the risk-drivers out to the projection date. This idea is not discussed in the American option pricing LSMC literature and is an original innovation in this thesis. In the context of pricing options, one is interested in estimating the mean or expected value of the continuation value distribution and using this as part of some approximation scheme to ultimately determine a fair value. In Section 3.7 we will see that the alternative sampling approaches for the fitting scenarios will only improve the success of the LSMC estimates when the aim is to estimate some upper percentile of a simulated distribution. Thus, these alternative sampling approaches will not add any benefit to the LSMC technique in the context of American option pricing. Of course, if one were interested in estimat-

ing a value-at-risk, or similar risk-measure, from positions held in American-style options, these alternative sampling approaches would improve the LSMC estimates of these quantities.

2.5 Basis functions in the LSMC method

Another issue one has to contemplate when implementing the LSMC method is the form the basis functions should take in the regression model. Recall, that the technique assumes the unknown functional form of the liability at the projection time, t_k, can be given as a linear combination of a countable set of $\mathcal{F}_{t_k}$-measurable basis functions. However, there are many different possible forms for these basis functions. One approach would be to take simple powers of the $\mathcal{F}_{t_k}$-measurable explanatory variables which are found to influence the liability at the projection date. Another approach often proposed is to construct the regression model from orthogonal polynomials of these $\mathcal{F}_{t_k}$-measurable explanatory variables. In this section we will give an introduction to the theoretical benefits of using orthogonal polynomials in a regression model. Some simple analysis of the design matrices under different types of basis functions will then be performed to investigate whether there seems to be evidence suggesting that using orthogonal polynomials could improve the success of the LSMC estimates.

Let us imagine we have some data on which we wish to perform a least-squares regression. For ease of illustration, it is assumed there is just one risk-driver influencing the regression, x say. Our data consists of n values of x and n values of the response variable L (which in the LSMC set-up would be the estimated liability values from each of the single inner scenarios). The regression model will then take the form

$$L_i = a_1 B_1(x_i) + a_2 B_2(x_i) + \cdots + a_Q B_Q(x_i) + \epsilon_i, \tag{2.3}$$

for $i = 1, \ldots, n$, where the a_j are constants, the $B_j(x)$ is the j-th basis function in the variable x and the ϵ_i is the error term, or noise, of the response L_i. The errors ϵ_i will be assumed to be normally distributed with mean zero and variance σ_ϵ^2 and be uncorrelated with the regressors, i.e., $\mathbb{E}[x_i \epsilon_i] = 0$.

This regression can be expressed more succinctly in the form of a matrix as

$$\mathbf{L} = \mathbf{aX} + \boldsymbol{\epsilon}, \tag{2.4}$$

where

$$\mathbf{L} = \begin{pmatrix} L_1 \\ \vdots \\ L_n \end{pmatrix}, \quad \mathbf{X} = \begin{pmatrix} B_1(x_1) & \cdots & B_Q(x_1) \\ \vdots & \ddots & \vdots \\ B_1(x_n) & \cdots & B_Q(x_n) \end{pmatrix}, \quad \mathbf{a} = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}, \quad \boldsymbol{\epsilon} = \begin{pmatrix} \epsilon_1 \\ \vdots \\ \epsilon_n \end{pmatrix}. \tag{2.5}$$

Under ordinary least-squares regression the parameters of the regression model are found using

$$\hat{\mathbf{a}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{L} \tag{2.6}$$

and the correlation matrix of the parameter estimates is given by

$$\text{corr}(\hat{\mathbf{a}}) = (\mathbf{X}'\mathbf{X})^{-1}\sigma_\epsilon^2. \tag{2.7}$$

It can be seen from these results that if the matrix $(\mathbf{X}'\mathbf{X})^{-1}$ is diagonal then this will result in uncorrelated parameter estimates. This would be beneficial in the regression because the larger the correlation between the parameter estimates the greater the resultant variance associated with each of the individual parameter estimates.

By setting the basis functions $B_0(x), \ldots, B_Q(x)$ as the first $Q+1$ polynomials from some family of orthogonal polynomials, the matrix $(\mathbf{X}'\mathbf{X})^{-1}$ will be diagonal in the limit $n \to \infty$ if $x_1, \ldots, x_n$ are evenly distributed across the domain on which the family of polynomials are orthogonal. In practice, we will have a finite number of x_i, but it is possible through the sampling schemes discussed in Section 2.4 to make these points as evenly-spaced across the domain as possible.

With this in mind, a few tests will now be performed to investigate whether the matrix $(\mathbf{X}'\mathbf{X})^{-1}$ is 'closer to diagonal' when using a family of orthogonal polynomials as the basis functions in the regression, compared with just using simple powers of the explanatory variable. For these tests, we will assume a large number of realisations of the explanatory variable x have been sampled from the real-world distribution and x_{upp} and x_{low} represents the maximum and minimum of this population. The range over which we wish to sample the fitting points used to determine the best-fitting

parameter estimates in the regression model is then given as $x \in [x_{low}, x_{upp}]$. The two different sampling approaches which will be used in these tests will be uniform and quasi-random simulation. For each of these two approaches we will consider both the simple polynomial and orthogonal polynomial basis functions.

Assume we have generated n fitting points, $x_i, i = 1, \ldots, n$, by either uniform or quasi-random sampling over $[x_{low}, x_{upp}]$. For the case of simple polynomial basis functions, we simply set the basis functions as

$$B_0(x_i) = 1, \quad B_1(x_i) = x_i, \quad B_2(x_i) = x_i^2, \quad \ldots, \quad B_Q(x_i) = x_i^Q. \tag{2.8}$$

For the orthogonal polynomial basis functions in this analysis, we will consider the class of Legendre orthogonal polynomials. The Legendre polynomials are orthogonal with respect to the L^2-inner product over the interval $[-1, 1]$:

$$\int_{-1}^{1} B_m(x) B_n(x) dx = \frac{2}{2n+1} \delta_{mn}, \tag{2.9}$$

where δ_{mn} is equal to 1 if $m = n$ and zero otherwise. As these polynomials are orthogonal over the interval $[-1, 1]$, we must scale and shift our fitting points so that they lie within this domain. This would be achieved by simply setting

$$\tilde{x}_i = \frac{2(x_i - x_{low})}{x_{upp} - x_{low}} - 1. \tag{2.10}$$

Taking the basis functions as the first few Legendre polynomials gives:

$$\begin{aligned} B_0(\tilde{x}_i) &= 1 && (2.11) \\ B_1(\tilde{x}_i) &= \tilde{x}_i && (2.12) \\ B_2(\tilde{x}_i) &= \frac{1}{2}(3\tilde{x}_i^2 - 1) && (2.13) \\ B_3(\tilde{x}_i) &= \frac{1}{2}(5\tilde{x}_i^3 - 3\tilde{x}_i). && (2.14) \end{aligned}$$

Further Laguerre polynomials can be determined using Bonnet's recursion, which gives the $(n+1)$-th order polynomial in terms of the n-th and $(n-1)$-th as

$$(n+1) B_{n+1}(\tilde{x}_i) = (2n+1) \tilde{x}_i B_n(\tilde{x}_i) - n B_{n-1}(\tilde{x}_i). \tag{2.15}$$

Other classes of orthogonal polynomials also have simple recursions to generate them. They may be orthogonal over different domains, however this can be accounted for by adapting Equation 2.10 as appropriate. If we use uniform sampling to generate 10,000 points over the interval $[-1, 1]$ and then use the first 4 Laguerre polynomials to give the columns of the design matrix $\mathbf{X}$, this will result in this matrix having 10,000 rows and 4 columns. Then, the 4×4 matrix $(\mathbf{X}'\mathbf{X})^{-1}$ will be given as is shown just below. The equivalent matrices for Sobol quasi-random sampling over $[-1, 1]$ together with Legendre orthogonal polynomial basis functions and each of the two sampling approaches with a simple third-order polynomial regression are also given below:

Uniform sampling with Legendre polynomials:

$$(\mathbf{X}'\mathbf{X})^{-1} = \begin{pmatrix} 1.0 \times 10^{-4} & -9.9 \times 10^{-7} & -7.6 \times 10^{-8} & 1.8 \times 10^{-6} \\ -9.9 \times 10^{-7} & 3.0 \times 10^{-4} & 3.9 \times 10^{-7} & -3.9 \times 10^{-6} \\ -7.6 \times 10^{-8} & 3.9 \times 10^{-7} & 5.0 \times 10^{-4} & -2.5 \times 10^{-6} \\ 1.8 \times 10^{-6} & -3.9 \times 10^{-6} & -2.5 \times 10^{-6} & 6.9 \times 10^{-4} \end{pmatrix} \tag{2.16}$$

Average of absolute value of off-diagonal elements: 1.67×10^{-5}.

Sobol quasi-random sampling with Legendre polynomials:

$$(\mathbf{X}'\mathbf{X})^{-1} = \begin{pmatrix} 1.0 \times 10^{-4} & 3.8 \times 10^{-8} & 2.5 \times 10^{-8} & 3.8 \times 10^{-8} \\ 3.8 \times 10^{-8} & 3.0 \times 10^{-4} & 1.2 \times 10^{-7} & 1.8 \times 10^{-7} \\ 2.5 \times 10^{-8} & 1.2 \times 10^{-7} & 5.0 \times 10^{-4} & 1.7 \times 10^{-7} \\ 3.8 \times 10^{-8} & 1.8 \times 10^{-7} & 1.7 \times 10^{-7} & 7.0 \times 10^{-4} \end{pmatrix} \tag{2.17}$$

Average of absolute value of off-diagonal elements: 7.14×10^{-8}.

Uniform sampling with simple polynomial:

$$(\mathbf{X}'\mathbf{X})^{-1} = \begin{pmatrix} 2.3 \times 10^{-4} & -1.2 \times 10^{-5} & -3.7 \times 10^{-4} & 1.7 \times 10^{-5} \\ -1.2 \times 10^{-5} & 1.9 \times 10^{-3} & 1.8 \times 10^{-5} & -2.6 \times 10^{-3} \\ -3.7 \times 10^{-4} & 1.8 \times 10^{-5} & 1.1 \times 10^{-3} & -2.4 \times 10^{-5} \\ 1.7 \times 10^{-5} & -2.6 \times 10^{-3} & -2.4 \times 10^{-5} & 4.4 \times 10^{-3} \end{pmatrix} \tag{2.18}$$

Average of absolute value of off-diagonal elements: 3.80×10^{-4}.

Sobol quasi-random sampling with simple polynomial:

$$(\mathbf{X}'\mathbf{X})^{-1} = \begin{pmatrix} 2.3 \times 10^{-4} & 4.6 \times 10^{-8} & -3.8 \times 10^{-4} & -1.2 \times 10^{-7} \\ 4.6 \times 10^{-8} & 1.9 \times 10^{-3} & -1.9 \times 10^{-7} & -2.6 \times 10^{-3} \\ -3.8 \times 10^{-4} & -1.9 \times 10^{-7} & 1.1 \times 10^{-3} & 6.3 \times 10^{-7} \\ -1.2 \times 10^{-7} & -2.6 \times 10^{-3} & 6.3 \times 10^{-7} & 4.4 \times 10^{-3} \end{pmatrix} \tag{2.19}$$

Average of absolute value of off-diagonal elements: 3.73×10^{-4}.

These matrices show that the off-diagonal elements of the matrix $(\mathbf{X}'\mathbf{X})^{-1}$ are on average closer to zero when Legendre polynomial basis functions are used, in comparison to employing a simple polynomial regression. Furthermore, by employing Sobol quasi-random sampling of the fitting points with orthogonal polynomial basis functions, these off-diagonal elements are closer to zero again, than when using simple uniform random sampling over the interval $[-1, 1]$. This is not a great surprise, as the theory of orthogonal polynomials suggests that the matrix $(\mathbf{X}'\mathbf{X})^{-1}$ should be diagonal under an even distribution of points and infinitely many of these orthogonal polynomials. Thus, by employing quasi-random sampling, we would expect that the matrix using four of these polynomials would be more 'diagonal-like', than by just employing a standard polynomial regression.

Of course, if the matrix $(\mathbf{X}'\mathbf{X})^{-1}$ is close to diagonal, this will give parameter estimates which have very little correlation with one another. Therefore, this analysis offers some evidence that employing orthogonal polynomial basis functions (particularly when in combination with quasi-random sampling) might give a more accurate regression in the LSMC method. Of course, all these matrices have fairly small off-diagonal elements, so any extra accuracy gained in the LSMC method may be small. Nonetheless, this should be investigated further and some tests of different orthogonal polynomials (with various fitting point sampling approaches) within the LSMC method are performed in Section 3.8.

2.6 LSMC outer and inner scenario allocation

The final issue remaining in the optimal configuration of the LSMC method is how to allocate the available computational budget between the simulation of the outer fitting scenarios and the inner valuation scenarios. By increasing the number of

inner valuation scenarios (per outer fitting scenario), the accuracy of the liability valuation for each fitting scenario will be improved. This means that each of the liability values on which we want to regress in terms of the key explanatory variables will have less associated sampling error, thus we would expect that each of the regression-corrected estimates will also be of greater accuracy. Therefore, by employing a greater number of inner scenarios (per fitting scenario), a more accurate LSMC estimated liability value distribution can be determined. On the other hand, if the number of outer fitting scenarios is increased (whilst keeping the number of inner valuation scenarios per outer scenario fixed), the accuracy of the LSMC estimated projected value distribution should also improve. The error of the conditional projected liability value for each fitting scenario will not depend on how many outer scenarios are sampled. However, with a greater number of fitting points through which to perform a regression, we would expect to get a more accurate estimated projected liability value function in terms of the explanatory variables. Hence, by sampling a greater number of outer fitting scenarios, the accuracy of the LSMC method will also increase.

Therefore, given a fixed computational budget, there is obviously a trade-off that must be made between the number of outer fitting scenarios and the number of inner valuation scenarios in the LSMC set-up. Would generating a large number of outer fitting scenarios, with a relatively small number of inner scenarios produce more accurate estimates than doing the converse? In Section 3.9 a test of the accuracy of the LSMC estimates under different scenario budgets will be performed. This will provide us with guidance how to allocate the computational budget in calculating an insurance SCR by this technique.

2.7 Alternative approaches to LSMC

To conclude this chapter, a couple of alternative approaches to the LSMC method which are sometimes used by practitioners will be outlined. Some of the similarities between these other two techniques and LSMC will also be examined in this section. This will provide a brief overview of some of the other possible techniques insurance companies could use to calculate their SCR and how these different approaches are related to one another.

2.7.1 The curve fitting approach

The first alternative technique that practitioners could use to estimate the projected liability distribution is known as curve fitting. This approach is similar in practice to LSMC and operates as follows: Under the full nested simulation approach for estimating the projected insurance liability distribution, outlined in Section 2.1, the resultant total number of valuation scenarios was too large to be computationally practical. Under the LSMC method, the idea was to reduce the number of inner valuation scenarios per outer scenario to a much smaller number, then correct the liability estimates by performing a regression. With the curve fitting approach, however, the idea is to reduce the number of fitting scenarios to a far smaller number instead. This small number of outer scenarios will usually be chosen to be representative of risk-driver values which will yield some key percentiles of the projected liability distribution, when possible.

After this a curve would then be chosen to exactly fit through this small number of points, which will of course be fairly accurately valued, given a full 10,000 or more inner scenarios are being used. If we assume these valuations have no error, then if we have p of these points, a polynomial of degree $p-1$ could be fit exactly to these liability percentiles. This is shown schematically in Figure 2.9. For complex insurance liabilities, however, there will often be some appreciable error, even if 10,000 valuation scenarios have been used. Thus, some approximation technique to fit the curve through these points must be used. One such approach could be to minimise the sum of the squared-errors. But then, of course, we are back in the regime of the LSMC technique and the issue at hand is just the scenario budget allocation issue between outer and inner simulations discussed back in Section 2.6.

Different scenario budget allocations within the LSMC method will be tested in Section 3.9 and some conclusions will be drawn here as to how practitioners should distribute their available computational resources to make the LSMC technique most efficient. These tests will also be relevant in the comparison of the LSMC and curve fitting techniques since typical insurance liability estimates will still yield some significant amount of error even when a large number of valuation scenarios are used.

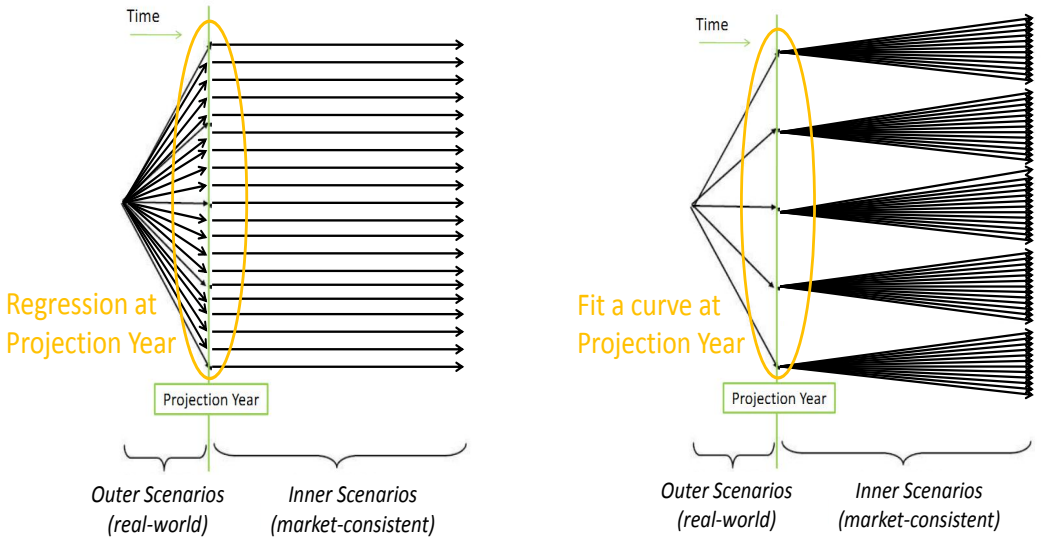

Figure 2.9: Schematic representation of the LSMC approach (left) with a large number of outer fitting scenarios and a small number of inner valuation scenarios (per outer scenario). The converse, known as curve fitting, is shown to the right.

2.7.2 The replicating portfolio approach

The second alternative technique which an insurer could use to calculate a SCR is known as the replicating portfolio approach. Under this method the insurance liability cashflows are approximated by the cashflows from a set of standard financial instruments. Theoretically, if the value of this portfolio of assets is equal to the value of the liability today, then it should be equal to the value of the liability given our real-world economic projection, by the law of one price. Valuing this portfolio at the projection date provides a faster method for generating market-consistent projections of the liability, compared to valuing all the liability cashflows for the complex insurance product over a term of many decades. For more information on the replicating portfolio approach, a couple of excellent introductory articles are Koursaris [Kou11] and Oechslin et al. [Oec07]. For a more detailed discussion, see the textbook of Wüthrich, Bühlmann and Furrer [Wüt10].

Although this technique is conceptually appealing, there are a number of technical challenges in applying this method in practice, including choice of appropriate candidate assets and economic scenarios. In addition, the long-term nature of insurance business and its exposure to numerous (often non-traded) risk-drivers means that it can be challenging to replicate even the simplest liabilities in the capital markets.

Chapter 3

Optimising the LSMC Algorithm

In Sections 2.4-2.6 some of the choices one can make in implementing the LSMC method were explored. In this chapter we will try to investigate some of these issues with a view to finding the optimal set-up of the LSMC algorithm. With so many factors affecting the success of the LSMC approximation, it is crucial that we can obtain as much information as possible concerning the algorithm for estimating the liabilities. Bearing this in mind it is sensible to introduce a case study involving a product which admits analytical values in the inner scenario simulation component of the algorithm. Then we can obtain an exact liability value for each real world outer scenario projection. The LSMC method gives a regression-adjusted simulation estimate for each of these real world outer scenarios. Thus, by using a product with a corresponding analytical value in the inner scenario stage, the LSMC estimates can be easily compared to these true conditional liability values. Furthermore, this comparison can be made either across the whole distribution of real world outer scenario projections, or perhaps just concentrating on a particular part of this real world distribution in isolation, for example an upper percentile. This is particularly appealing in an insurance risk-management perspective, where the the success of the approximation in the upper tail is of greater importance than other parts of the distribution.

3.1 Projected value of a European put option

The simplest such product which admits an analytical value is a European option under the Black-Scholes model for equity asset returns. Under the Black-Scholes model, the risk-free rate of interest and the volatility are assumed to be constant in time, thus in the valuation stage of our test of the LSMC algorithm this must also be the case. Obviously, this is not a realistic assumption for actual market behaviour, but the purpose of this set-up is simply to test the mechanics of the LSMC method. Once a better understanding has been obtained of the effect of the parameters of LSMC algorithm on the accuracy of the approximation, then these parameters can

be used for a more realistic product and model in the inner scenario valuation stage (where simple analytical valuation formulae will no longer be available to us).

For the analysis which follows, we will consider a European put option under the Black-Scholes model. The reason for choosing a put option is that the structure of the payoff is quite similar to that of the liabilities written on an insurance product. Under an insurance guarantee, the insurer typically must make some payment, P, to a policyholder from an underlying fund, F_t. This fund will be invested in some financial asset or portfolio from annuitisation until the point in time when the payment needs to be made, hence F_t is given by some stochastic process. At time T when the payment is due to be made to the policyholder, if $F_T > P$ the insurer will not face any liability. However, if $F_T < P$ then the insurer must pay the difference out of its own reserves—this is what we call a liability to the insurer. Thus, the liability L an insurer faces from this simple general insurance guarantee is given by

$$L = \max(P - F_T, 0). \tag{3.1}$$

One can see immediately that this is very similar in form to the payoff of a European put option. This hopefully provides some justification that considering a simple put option under the Black-Scholes model will at least capture some of the structure of typical liabilities which arise from insurance guarantees. Then, hopefully any results or stylised rules regarding the most efficient LSMC set-up will extend from this constructed example to the realistic setting of calculating an insurance solvency capital requirement (SCR).

3.2 LSMC Analysis Set-Up

For this equity put valuation set-up, we will consider the term to maturity to be T years and attempt to determine the projected value of this option in t year's time. The risk-free rate of interest will be assumed to be constant over the T year term of the option, but the underlying equity and volatility will be projected to year t using a stochastic real world model. Thus, the value of the option at year t will depend on two risk-drivers – the equity level at year t and the volatility level at year t. We will look to approximate the projected value of the option at year t using the LSMC algorithm. The first stage of this algorithm is to simulate (or set)

M fitting scenarios, which will be used to find the regression proxy function, which approximates the value of the option at year t in terms of our risk-drivers. Different possible sampling methods for these outer fitting scenarios were discussed in Section 2.4. For each of these fitting points, $(\widetilde{S}_t^m, \widetilde{\sigma}_t^m)$, we then simulate N inner valuation scenarios under our valuation (risk-neutral Black-Scholes) model. This will give N simulated equity prices at year T of $\widetilde{S}_T^{m,n}$, $n = 1 \ldots, N$, for each fitting scenario m. The corresponding option payoff for each of these is just given by

$$\widetilde{V}_T^{m,n} = \max(K - \widetilde{S}_T^{m,n}, 0). \tag{3.2}$$

The estimated option value at year t for fitting scenario m, V_t^m, is then given by taking the average of the subsequent N option payoffs and discounting:

$$\widetilde{V}_t^m = e^{-r(T-t)} \frac{1}{N} \sum_{n=1}^{N} \widetilde{V}_T^{m,n}. \tag{3.3}$$

Thus, we now have an estimate for the value of the option at year t for each of the fitting scenarios, $m = 1, \ldots, M$. As explained in Section 2.1, when the financial product is complex and path-dependent the required computational run-time for each inner valuation scenario can be relatively large. As such, the value of N must be chosen to be relatively small, otherwise the required run-times for projecting liabilities will become prohibitively large. The idea of the LSMC approach is that the large standard errors from using a small number of valuation scenarios can be overcome if we perform a regression on these estimated option values on the two risk-drivers over all the fitting scenarios generated. Under a standard Ordinary Least-Squares (OLS) regression

$$\widetilde{V}_t^m = f(\widetilde{S}_t^m, \widetilde{\sigma}_t^m) + \epsilon_m, \tag{3.4}$$

where the ϵ_m are the regression residuals, assumed to be independent and identically distributed normal random variables. Generally, these residuals will be independent as Monte Carlo errors, however in practice they are unlikely to be identically distributed. Indeed, the variance of the error ϵ_m will usually depend on the values of the corresponding explanatory variables $\widetilde{S}_t^m$ and $\widetilde{\sigma}_t^m$. Fortunately, the OLS estimator is still consistent even if the residuals in the regression are not identically distributed.

This fact is shown formally by White [Whi00]. Furthermore, there are studies which show that variable selection criteria can still perform reasonably well in the presence of heteroscedasticity. See for example Baek, Karaman and Ahn [Bae05].

Let us express the function $f(\widetilde{S}_t^m, \widetilde{\sigma}_t^m)$ as being made up of $B+1$ functions of the risk-drivers $\widetilde{S}_t^m$ and $\widetilde{\sigma}_t^m$:

$$f(\widetilde{S}_t^m, \widetilde{\sigma}_t^m) = \sum_{q=0}^{Q} a_q B_q(\widetilde{S}_t^m, \widetilde{\sigma}_t^m), \tag{3.5}$$

where the $a_q \in \mathbb{R}$ are the regression coefficients and $B_0(\widetilde{S}_t^m, \widetilde{\sigma}_t^m) = 1$ meaning a_0 is just a constant in the regression model. The $Q+1$ regression coefficients can then be determined by minimising the sum of the squared errors from the regression. That is, we choose $a_q \in \mathbb{R}$, $q = 0, \ldots, Q$, to minimise

$$\sum_{m=0}^{M} \left(\widetilde{V}_t^m - f(\widetilde{S}_t^m, \widetilde{\sigma}_t^m)\right)^2. \tag{3.6}$$

Let us denote the function which minimises the sum of squared errors in the regression, or the function with optimal choice of coefficients a_q, as $\hat{f}(\widetilde{S}_t^m, \widetilde{\sigma}_t^m)$. We then call $\hat{f}$ the (fitted) regression proxy function of the two risk-drivers.

Having found our regression proxy function using the fitting scenarios, we are now in a position to estimate the projected option values for the real world economic scenarios. These are simulated separately under a realistic model for the behaviour of the underlying equity and volatility. Sampling R scenarios from a real world stochastic volatility model yields R equity prices and volatilities at year t. Let us denote these values S_t^k and σ_t^k, for $k = 1, \ldots, R$. The fitted regression can now be used to give the LSMC estimate for real world scenario k as

$$\widehat{V}_t^k = \hat{f}(S_t^k, \sigma_t^k). \tag{3.7}$$

Of course, under a Black-Scholes model for the inner valuation scenarios an analytical option price exists for each of these R real world scenarios. This allows us to find the true value of the option for each fitting scenario and compare these to the LSMC estimates. The true value of the European put option for real world scenario

k is then given by

$$V_t^k = p^{\text{BS}}(\tilde{S}_t^k, K, T-t, r, \tilde{\sigma}_t^k), \tag{3.8}$$

where $p^{\text{BS}}(S, K, \tau, r, \sigma)$ is the analytical formula for the price of a European put option under the Black-Scholes model, with initial equity price S, strike price K, risk-free interest rate r, volatility σ and time τ remaining until maturity.

The values of $\widehat{V}_t^k$ and V_t^k can be used to determine a metric for 'success' in the LSMC method in various ways. We will outline two possible ways of measuring the accuracy of the LSMC estimates now. Firstly, the difference between the LSMC estimates and analytic values across the whole real world distribution can be considered. This first measure of error in the LSMC estimate, $\mathcal{M}_1$, can be expressed mathematically as

$$\mathcal{M}_1 = \frac{1}{R}\sum_{k=0}^{R} |\widehat{V}_t^k - V_t^k|. \tag{3.9}$$

The absolute value is required as otherwise there would be a cancelling out effect in the summation from a value of k where $\widehat{V}_t^k > V_t^k$ and a value of k where $\widehat{V}_t^k < V_t^k$. Furthermore, for now we are not concerned whether $\widehat{V}_t^k$ is larger than V_t^k for each scenario k - just with how close the two values are to one another.

The second metric for studying the 'success' of the LSMC estimate, $\mathcal{M}_2$, will be very different in nature to $\mathcal{M}_1$. Instead of attempting to capture how well the LSMC approximation technique performs across the whole real world distribution for S_1 and σ_1, this metric will concentrate on the upper percentile of the projected option value distribution at year t. This metric will be determined as follows; firstly the R analytic values of the option value at year t are sorted into increasing order. Let us denote these values as $V_t^{[j]}$, where $V_t^{[1]} < V_t^{[2]} < \cdots < V_t^{[R]}$. The 99.5-th percentile of the projected option value distribution at year t is then given by $V_t^{[j']}$ where $j' = \lfloor 0.995 \times R \rfloor$. We should note that this is actually an approximation to the percentile, based on our R real world simulations, however by choosing a sufficiently large value of R this estimate estimate will be very close to the true percentile of the distribution. We then take the real world scenario which subsequently gave $V_t^{[j']}$, let's say indexed by $k = k(j')$, and take the difference between the LSMC estimate of real world scenario k and the true 99.5-th percentile of the projected

value distribution. Mathematically, this metric $\mathcal{M}_2$ can then be expressed as

$$\mathcal{M}_2 = |\hat{f}(S_t^k, \sigma_t^k) - V_t^{[j']}|. \tag{3.10}$$

Again the absolute value of the difference has been taken, as we are not concerned with whether this value is positive or negative. Other possible metrics for studying the success of the LSMC algorithm could be constructed in a similar manner. The two which we have defined here, however, are quite representative of the major requirements practitioners have in approximating projected liability distributions: firstly, that some upper percentile is approximated fairly accurately, as ultimately this sort of method is most likely to be used to calculate a solvency capital requirement or value-at-risk. Secondly, as well as being a good approximation in some upper percentile of the distribution, some practitioners also feel that the approximation should also be fairly accurate across the rest of the distribution too. We note here that the metric $\mathcal{M}_2$ suffers from a major shortcoming which will be outlined in Section 3.4. An adjusted version of this metric which overcomes this problem will be presented there.

In the analysis which follows in subsequent sections we will perform tests on the LSMC set-up using the projected put option valuation problem under the Heston model, which will be introduced in Section 5.1. In these tests the projection time is $t = 1$ and the maturity of the option is $T = 10$ (years). Also, we set the initial equity price as $S_0 = 1$, the strike price to be $K = 1.3$ and the risk-free rate to be $r = 5\%$. For the Heston stochastic volatility model, the initial volatility is given by $\sigma_0 = 10.1\%$, the mean reversion level $\kappa = 1$, the mean reversion level $\theta = 20\%$ and the volatility of the variance process $\sigma_v = 0.15$. In these tests $1,000$ outer scenarios are sampled uniformly and 50 inner valuation scenarios are generated given each outer scenario.

3.3 Building up the LSMC regression model

The choice of the basis functions is clearly one of the key decisions one has to make when implementing the LSMC algorithm. In this section we will look at some different choices of regression model for the problem of estimating the projected value of the European put. As this issue is investigated some commentary will be

given as to how this analysis and any findings extend to the more general context of projecting insurance liabilities. This is, of course, the area where the LSMC algorithm introduced in this thesis will be useful in practice.

With the inner valuation scenarios simulated under the Black-Scholes model, the analytic value of the put option at time t can be easily calculated, given the projected equity price and volatility at year t. In this way, we can construct a true projected value distribution surface plot in the two risk-drivers, S_1 and σ_1. Two different views of this true projected value distribution surface are shown in Figure 3.1.

An alternative way of looking at this surface is via a contour plot. This plot, given in Figure 3.2, is a top-down view of the surface showing lines of constant projected option value, V_1. Where the lines of this plot are close together, the valuation surface is steepest. Studying Figure 3.2 in comparison with Figure 3.1, we can see the sharp fall in projected option value for S_1 increasing from 0.3 until around 0.8, when σ_1 close to zero. For $S_1 > 0.8$ the projected value of the option is close to zero, and the contour plot shows this region is very flat. The contour plot gives a nice two-dimensional interpretation of the three-dimensional projected value surface and is a useful visual tool in analysing these valuation surface plots.

The challenge we face is to obtain a fairly accurate representation of this true surface in terms of some proxy function of the two risk-drivers S_1 and σ_1. This proxy function is often described in terms of some function of basis functions, which are given by simple polynomials or from some family of orthogonal polynomials. The choice of basis functions in the LSMC method was discussed in Section 2.5. In Section 3.8 some tests will be performed investigating the accuracy of the LSMC estimates under different choices of basis functions. However, for the moment we shall just take powers of the explanatory variables as the basis functions in the implementation of the LSMC method.

Given that the basis functions are just powers of the two explanatory variables, S_1 and σ_1, the only issue remaining with the regression component of the LSMC method is to determine what form the proxy function should take. That is, what function of $S_1, \sigma_1, S_1^2, \sigma_1^2, S_1\sigma_1, S_1^3, \sigma_1^3, S_1^2\sigma_1, S_1\sigma_1^2, S_1^4, \sigma_1^4, \ldots$, will give the 'best' proxy function for estimating the projected option value at year one. It is important to be careful that the proxy function does not fit 'too closely' to the option values from

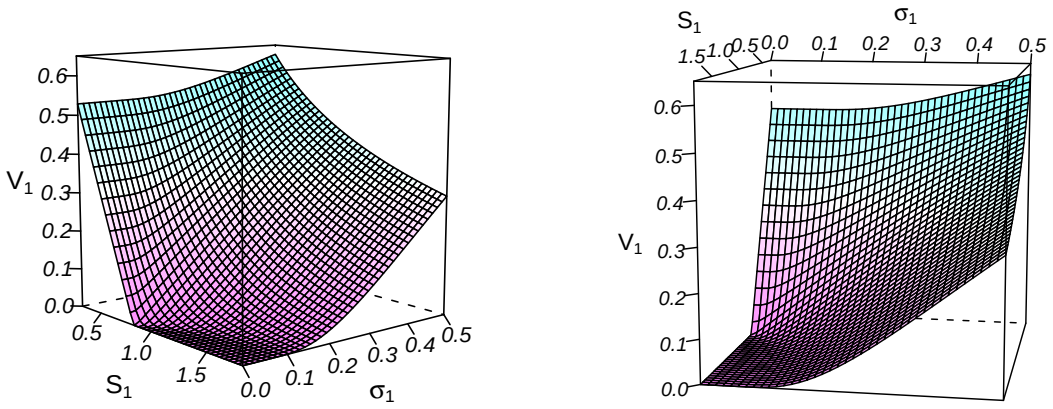

Figure 3.1: Two views of the analytic surface of V_1 varying with S_1 and σ_1.

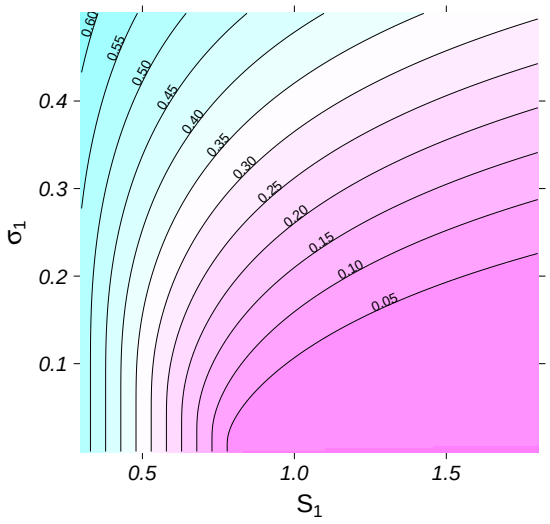

Figure 3.2: Analytic surface contour plot.

the outer fitting points, as the estimated value at these points has some inherent uncertainty. This situation is known as statistical over-fitting and will be discussed in Section 3.5. As we are testing the regression models with real world data which is independent of the outer fitting scenarios, the problem of statistical over-fitting should be accounted for with the regression error metrics, $\mathcal{M}_1$ and $\mathcal{M}_2$.

To begin with, let us consider a few simple proxy functions and study them with a view to understanding how each of the basis functions seem to be affecting the behaviour of the regression surface. Perhaps the simplest proxy function is just the

linear sum of the basis functions. That is, we fit the regression model using M fitting points by choosing a_0, a_1 and a_2 such that $\sum_{m=1}^{M} \epsilon_m^2$ is minimised, where

$$\epsilon_m = \widetilde{V}_t^m - f_1(\widetilde{S}_t^m, \widetilde{\sigma}_t^m) = \widetilde{V}_t^m - \left(a_0 + a_1 \widetilde{S}_t^m + a_2 \widetilde{\sigma}_t^m\right). \tag{3.11}$$

This fitted proxy function can then be plotted as a surface for different values of S_1 and σ_1 in the same manner as the true analytical surface plot was earlier. This surface is plotted in Figure 3.3, showing two views from the same angles that we displayed for the true projected value surface in Figure 3.1. Also given in Figure 3.3 is the contour plot corresponding to this regression surface and a contour plot of the surface of $\hat{V}_1 - V_1$. This is the difference surface between the regression proxy function and true surface of projected option values for the different values of S_1 and σ_1 considered.

These plots of the regression surface show that, although some of the characteristics of the true projected option value surface have been captured by this simple form of proxy function, f_1, there are still clear differences in their overall structures. By comparing the contour plots, the proxy function significantly underestimates the projected option value when both S_1 and σ_1 are relatively low. It is clear from Figure 3.2 that the true projected value function behaves quite differently with S_1 for different fixed values of σ_1. In order to capture this effect in the regression, one needs to include a cross-term, or a term with some power of S_1 multiplying some power of σ_1.

To study how adding such a term to our proxy function improves our regression, let us consider adding an extra $S_1\sigma_1$ term, as well as S_1^2 and σ_1^2 terms. The proxy function is now

$$f_2(S_1, \sigma_1) = a_0 + a_1 S_1 + a_2 \sigma_1 + a_3 S_1^2 + a_4 \sigma_1^2 + a_5 S_1 \sigma_1. \tag{3.12}$$

In Figure 3.4 a plot of the fitted regression surface is given from the two different viewing angles, together with a contour plot of this surface. Comparing these plots to Figure 3.1, show that including a cross-term and also the square of the explanatory variables in the proxy function gives a regression surface which is much more consistent with the true projected value surface. Studying the surface contour plot shows that the addition of the cross-term captures the bivariate dependence of the

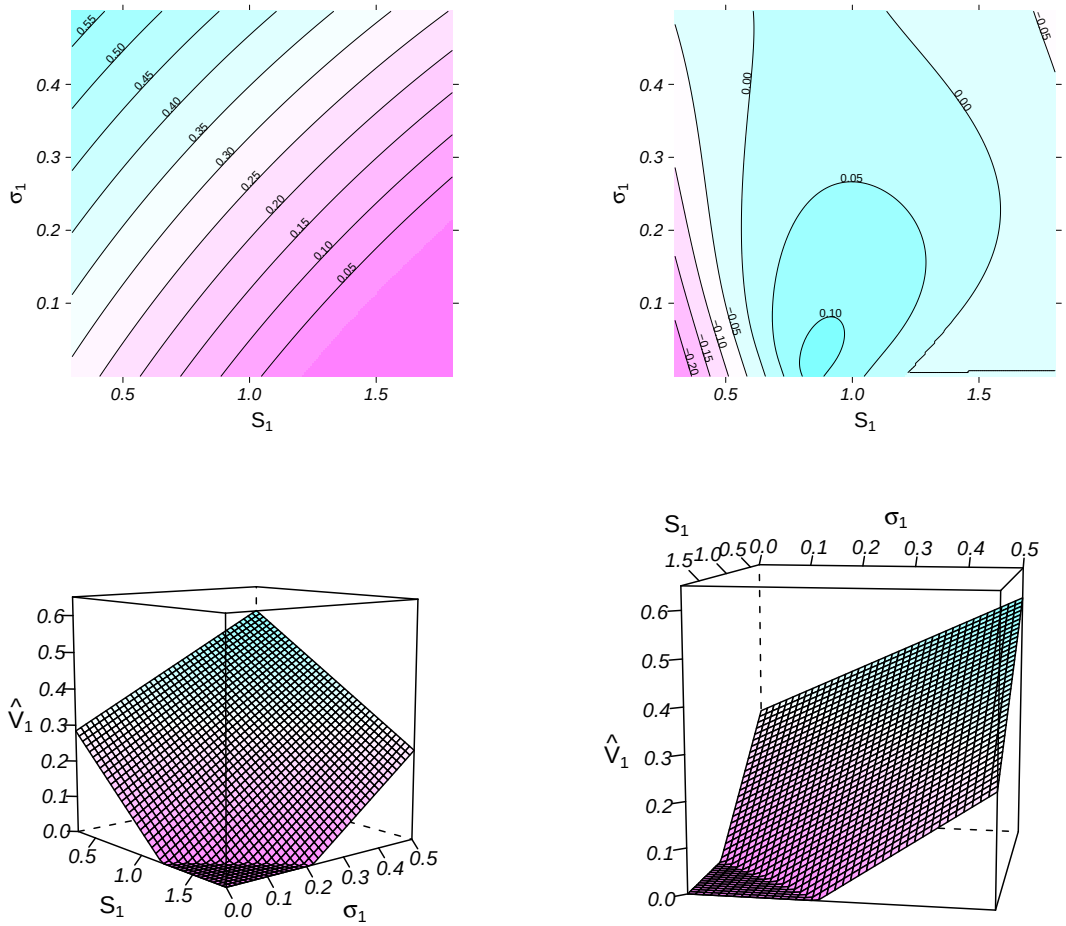

Figure 3.3: Bottom Row: Two views of regression surface of $\hat{V}_1$ varying with S_1 and σ_1, with $f_1(S_1, \sigma_1) = c_0 + c_1 S_1 + c_2 \sigma_1$. Top Left: Surface contour plot. Top Right: Difference surface $(\hat{V}_1 - V_1)$ contour plot.

projected option value much more accurately than proxy function f_1. The difference surface contour plot gives further clarity on the improved fit to the true surface from proxy function f_2.

Given that the regression surface appears more like the true projected value surface with the addition of the second-order terms and the second-order cross term, perhaps it would be sensible to try adding higher powers of the explanatory variables and higher order cross-terms to the proxy function. To experiment with the inclusion of some higher power basis functions, we can look at t-distribution hypothesis tests on the significance of each basis function within the proxy function and take any out which are not significant.

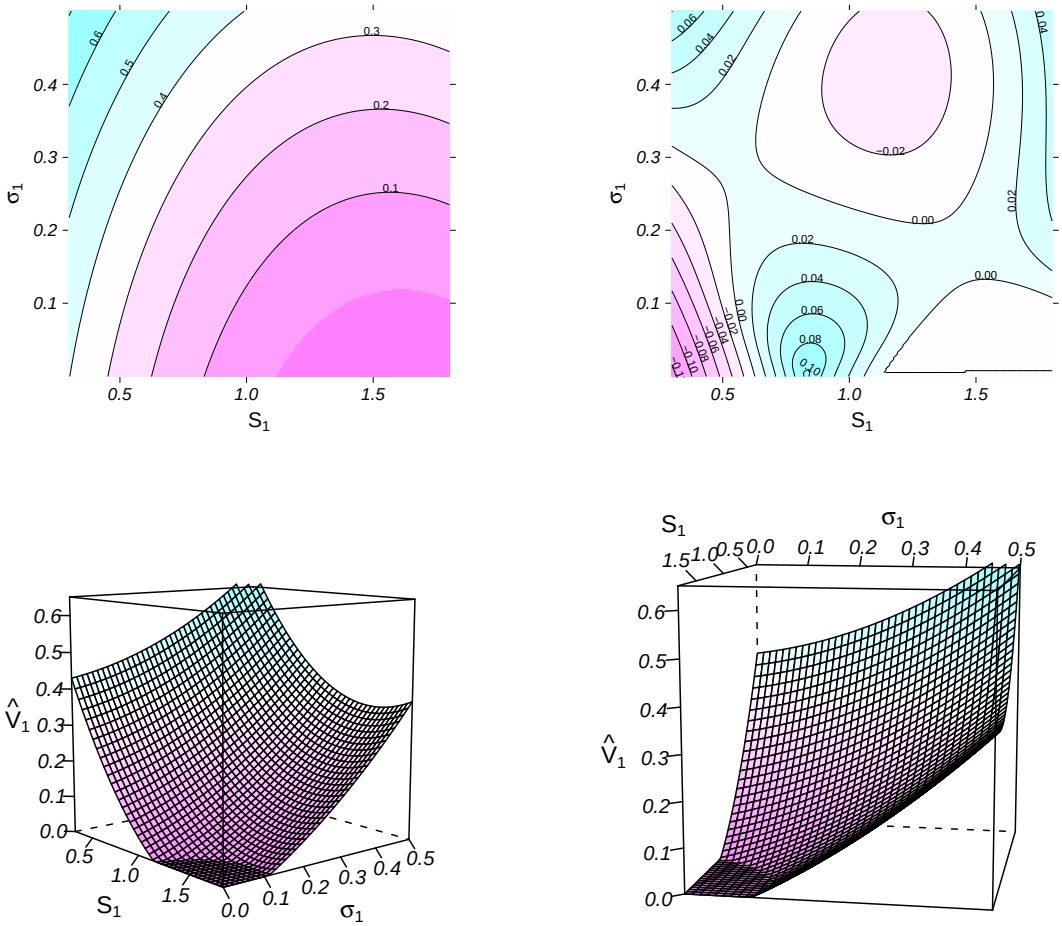

Figure 3.4: Bottom Row: Two views of regression surface of $\hat{V}_1$ varying with S_1 and σ_1, with $f_2(S_1, \sigma_1) = c_0 + c_1 S_1 + c_2 \sigma_1 + c_3 S_1^2 + c_4 \sigma_1^2 + c_5 S_1 \sigma_1$. Top Left: Surface contour plot. Top Right: Difference surface $(\hat{V}_1 - V_1)$ contour plot.

Proceeding in this manner results in the proxy function f_3 being chosen as a statistically successful form of model, where

$$f_3(S_1, \sigma_1) = c_0 + c_1 S_1 + c_2 \sigma_1 + c_3 S_1^2 + c_4 \sigma_1^2 + c_5 S_1 \sigma_1 + c_6 S_1^3 + c_7 S_1^2 \sigma_1 + c_8 S_1 \sigma_1^2 + c_9 S_1^3 \sigma_1. \tag{3.13}$$

A regression surface for this proxy function is shown from two different views in Figure 3.5 alongside the corresponding surface contour plot. If we compare Figure 3.5 with Figures 3.1 and 3.2, it is clear that proxy function f_3 gives a far superior fit to the true projected value distribution, than either of the two simple proxy functions. In fact at first glance, the regression surface is not very different from the analytic surface at all. Studying these surface plots closely, one can see the profile of

the projected option value for $\sigma \approx 0$ is slightly too smooth for the regression surface. This is no real surprise, as it is always going to be difficult to capture this non-smooth behaviour of the valuation surface in this region perfectly by considering just smooth basis functions. A second difference between these plots is the slight underestimation of projected option value in the regression model for low values of S_1 and high values of σ_1, and slight overestimation for low values of S_1 and reasonably low values of σ_1. This is perhaps more noticeable in the difference contour plot. These slight differences may not seem very important, however in practice the LSMC method will most likely be used to estimate some upper percentile of a liability distribution. Therefore, even around these extremes of the valuation surface we still want to obtain an accurate representation of the value from the regression proxy function.

In Figure 3.7 we give another plot of the regression surface for the proxy function f_3. However, this plot shows the contour lines projected onto the upper surface of the imaginary cube enclosing the surface. This helps in visualising the connection between the surface plot and the two-dimensional contour plot representation. (Note that this surface has not been floored at zero, so the surface and contours are slightly different in the corner immediately facing the reader in Figure 3.7).

Of course, manually adding and subtracting basis functions based on statistical significance tests is both time consuming and not very scientific. Furthermore, in the case of complex insurance liabilities in multiple dimensions one is likely to have less intuition about the valuation surface. Therefore, we will now introduce an approach which can systematically pick out the best form of proxy function in the regression for the chosen set of fitting points with which we are working.

3.3.1 Stepwise AIC regression approach

The stepwise AIC regression approach is an algorithm which can be used to find the best regression model for some fitting points (See Venables and Ripley [Ven02] for a reference to this algorithm). The method is initialised by choosing some simple regression proxy function. The proxy function f_2 from before would be a suitable choice here. Next we choose a maximal proxy function for the search. For this we take the explanatory variables and all cross-terms up to order 6. We assume the regression proxy function will not involve terms out-with this set of basis functions.

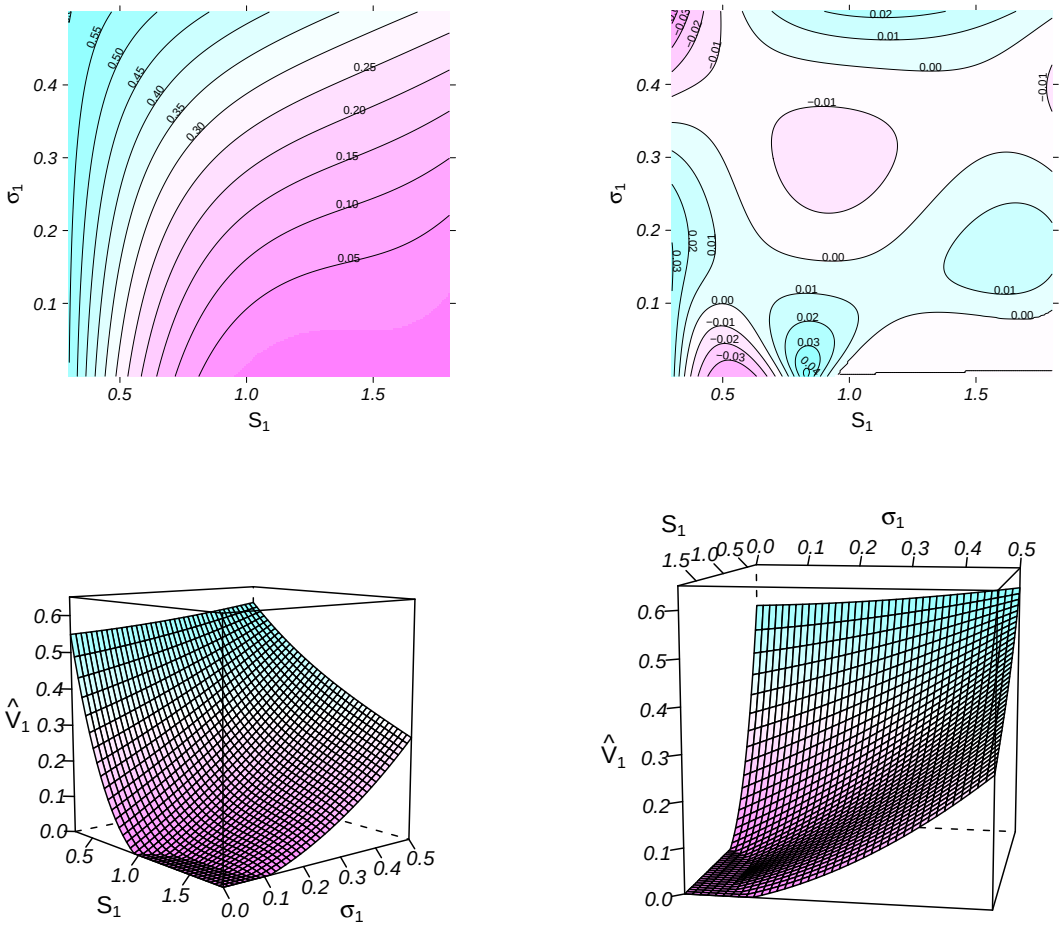

Figure 3.5: Bottom Row: Two views of regression surface of $\hat{V}_1$, with $f_3(S_1, \sigma_1) = c_0 + c_1 S_1 + c_2 \sigma_1 + c_3 S_1^2 + c_4 \sigma_1^2 + c_5 S_1 \sigma_1 + c_6 S_1^3 + c_7 S_1^2 \sigma_1 + c_8 S_1 \sigma_1^2 + c_9 S_1^3 \sigma_1$. Top Left: Surface contour plot. Top Right: Difference surface $(\hat{V}_1 - V_1)$ contour plot.

With these inputs defined, the model selection algorithm can now run.

The first regression model the stepwise AIC method considers is $c_0 + c_1 S_1 + c_2 \sigma_1 + c_3 S_1^2 + c_4 \sigma_1^2$ (the model which we chose to initialise the algorithm). Firstly, the algorithm calculates the Akaike information criterion (AIC) for this regression model. Next, the method will attempt to add each of the basis functions not in this proxy function (but which are present in the maximal model) in isolation and calculate the new AIC for each of these models. As 1, S_1, σ_1, S_1^2 and σ_1^2 are in the current proxy function, the algorithm will take each of these out of the model in isolation and calculate the AIC for each of these new possible models too. It will then compare the AIC for all these new possible models and the one which gives the biggest decrease, compared to the current model being considered, will be chosen. This model will

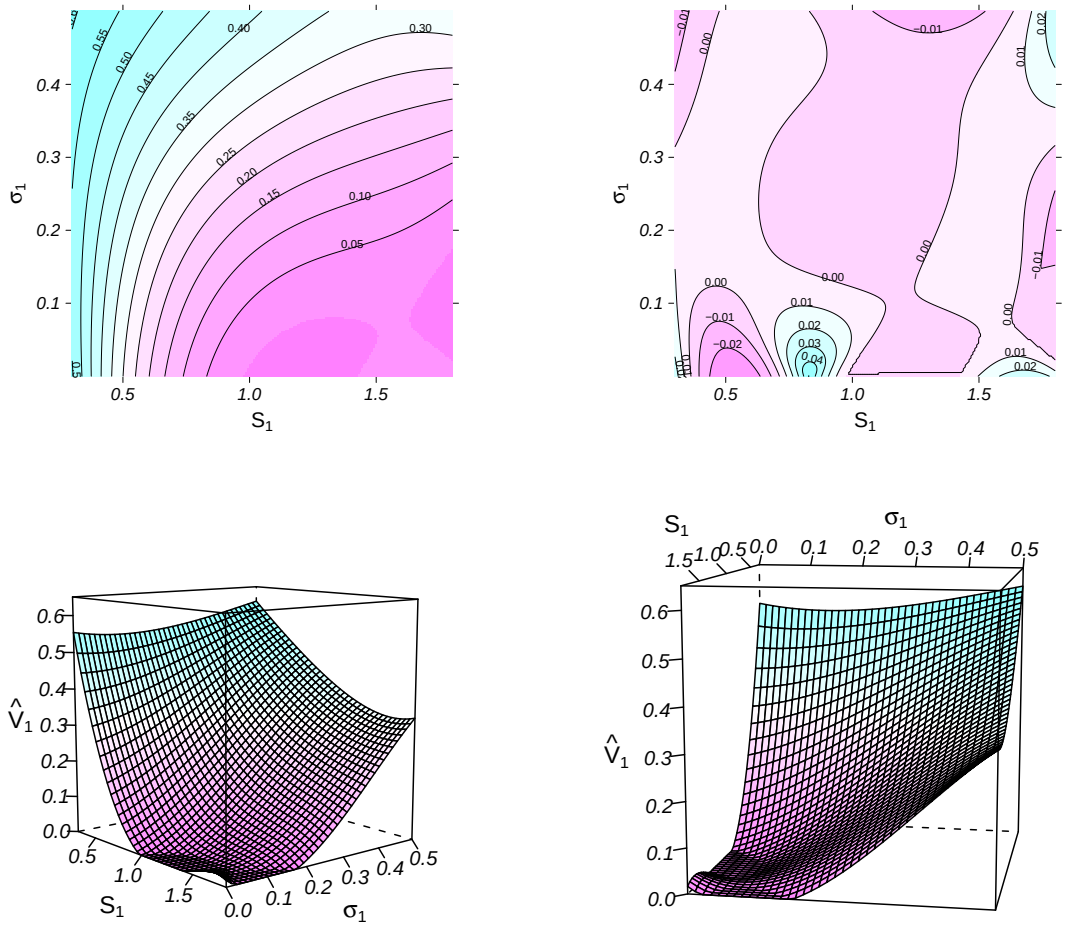

Figure 3.6: Bottom Row: Two views of regression surface of $\hat{V}_1$, with $f_4(S_1, \sigma_1)$, which was found using the stepwise AIC algorithm. Top Left: Surface contour plot. Top Right: Difference surface $(\hat{V}_1 - V_1)$ contour plot.

then become the new model which will be tested. Again, the AIC values will be calculated for all models which can be obtained by adding or subtracting a single basis function to or from this second model being tested. The model which yields the greatest decrease in AIC will be chosen as the third model to be tested.

This process carries on until the addition or subtraction of any single basis function will only increase the AIC. The algorithm will then terminate and the model which is being tested when this occurs is chosen as the optimal regression model.

Running the stepwise AIC algorithm for the fitting point data from Section 3.3 results in the proxy function f_4 being selected as the best regression model for the

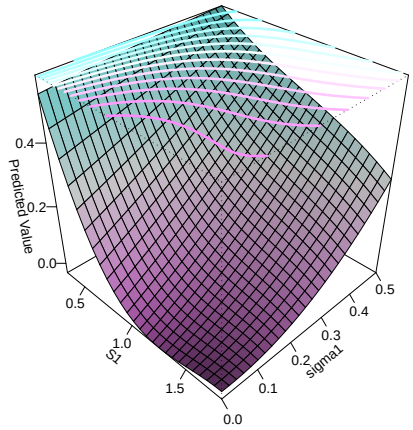

Figure 3.7: Contour lines projected above surface plot for proxy function f_3.

projected put option value at year 1, where

$$f_4(S_1, \sigma_1) = c_0 + c_1 S_1 + c_2 \sigma_1 + c_3 S_1^2 + c_4 \sigma_1^2 + c_5 S_1 \sigma_1 + c_6 \sigma_1^3 + c_7 S_1^2 \sigma_1 + c_8 S_1 \sigma_1^3 + c_9 S_1^2 \sigma_1^2 + c_{10} S_1^5 + c_{11} S_1^5 \sigma_1. \tag{3.14}$$

The regression surface for this proxy function is shown from two different views in Figure 3.6 alongside the corresponding surface contour plot. The two views of the surface show this regression model behaves quite similarly to proxy function f_3. Both surfaces are generally quite similar to the true projected valuation surface. It is difficult to infer from these plots which regression model fits best out of of f_3 and f_4. It is perhaps easier to compare how these two surfaces correspond with the true projected value distribution in the centre of the distribution if we study the surface contour plots. Looking at the contour plots in Figures 3.5 and 3.6 compared with Figure 3.2, we can see that both proxy functions f_3 and f_4 approximate the true valuation surface fairly closely over the range of values considered. The difference contour plot, again, shows this reasonably accurate fit. By eye it is difficult to distinguish which of proxy functions f_3 or f_4 gives a better fit to the true projected value distribution from these contour plots. However, looking at the values of the regression error metrics for these proxy functions may give us an approach to choose between these two models.

3.4 Performance of regression error metrics

Having compared some different choices of proxy function visually using surface and contour plots, we will now investigate how our two regression error metrics, introduced in Section 3.1, perform in determining how successful each form of proxy function is in the LSMC method. Shortly, we will see that the metric $\mathcal{M}_2$ is a poor indicator of the success of fit of a regression model, however an adjusted version of $\mathcal{M}_2$ can address the problems of this metric.

To test these risk metrics, we use $R = 50,000$ real world economic projections out to year one using the Heston stochastic volatility model. This gives us $50,000$ different values of S_1 and σ_1, which are, of course, the two risk-drivers on which the projected option value, V_1, depends. These tests use 1,000 uniformly sampled outer scenarios and 10 inner valuation scenarios per outer scenario. An average of these metrics is taken over 10,000 of these tests to factor out any dependence on a particular realisation of the uniform sampled fitting scenarios. Note, that f_4 will take a different form given each different fitting scenario tested, depending on what is output from the stepwise AIC algorithm.

Using this data let us calculate the risk metric $\mathcal{M}_1$ associated with each of the regression proxy functions. We recall this metric is given by:

$$\mathcal{M}_1 = \frac{1}{R}\sum_{k=0}^{R} |\widetilde{V}_t^k - V_t^k|, \tag{3.15}$$

where $\widetilde{V}_t^k$ is the regression estimate for real world projection k and V_t^k is the analytic Black-Scholes value given real-world projection k. The values of the metric $\mathcal{M}_1$ for the four different proxy functions which were considered in the last section are given in Table 3.1.

Regression error metric $\mathcal{M}_1$:	
Proxy Funct.	$\mathcal{M}_1$
f_1	0.0630
f_2	0.0181
f_3	0.0037
f_4	0.0026

Table 3.1: Regression error metric $\mathcal{M}_1$ for different proxy functions.

The values of the metric $\mathcal{M}_1$ given in this table are consistent with our discussion of the plots of the regression surfaces and contour plots for f_1 to f_4, compared with the corresponding plots for the true projected value distribution. The smaller the value of $\mathcal{M}_1$, the more accurately the regression proxy function has captured the true projected value distribution on average across its whole range. In moving from regression proxy function f_1 to function f_2, the value of $\mathcal{M}_1$ fell by around 70%. When we considered Figures 3.3 and 3.4 in comparison with the true distribution surface, we argued that introducing the higher power basis functions and the cross-term had significantly improved the accuracy the proxy surface modeled the true surface at essentially all the values of S_1 and σ_1 considered. Thus, this fall in $\mathcal{M}_1$ from proxy function f_1 to f_2 is definitely consistent with what we see visually in the diagnostic plots given earlier. There is a further fall of nearly 80% in the value of this error metric when we move from proxy function f_2 to proxy function f_3. This, again, fits with our discussion of the regression surface and the true distribution surface plots given earlier. Comparing both Figure 3.5 and Figure 3.4 with the true projected value distribution surface, we saw that using proxy function f_3, found by considering the significance of each of the terms in the regression model, seemed to give a much closer fit to the true distribution surface.

By considering the surface and contour plots, it was impossible to definitively conclude which of functions f_3 and f_4 gave the best fit across the whole distribution. Both seemed to fit well in the centre of the projected value distribution. The regression error metric $\mathcal{M}_1$ shows a relatively small fall in moving from proxy function f_3 to f_4. This can probably be explained by the fact that f_4 selected the best fitting regression model for each specific set of fitting data considered, whereas f_3 was fixed upfront. This is reassuring as function f_3 was chosen for one specific set of fitting data, yet it performed almost as well as f_4 in terms of the metric $\mathcal{M}_1$.

Overall, it would appear that if we measure success in the LSMC method as a fairly accurate approximation across the whole projected value distribution, then $\mathcal{M}_1$ appears to be a good statistic to consider. Of course, we do need to know the option values corresponding to the inner scenario model analytically, which means we do not need an LSMC approximation in the first place. Thus, this metric can only be used in this test product set-up. The findings on how to construct the regression model should extend to a more general insurance framework, however.

Suppose we are less concerned with how successful the LSMC approximation is across the whole distribution, and more concerned how successful it is in estimating large outcomes of projected option value? The second LSMC error metric, introduced in Section 3.1, would seem more appropriate when this is the case. Recall, this metric was given by

$$\mathcal{M}_2 = |f^\star(S_t^k, \sigma_t^k) - V_t^{[j']}|, \tag{3.16}$$

where $V_t^{[j']}$ is the (estimated) true 99.5-th percentile, obtained by taking the corresponding order statistic of the R analytic projected option values and $f^\star(S_t^k, \sigma_t^k)$ is the regression estimate of the real world projection which subsequently gave the 99.5-th percentile of the projected value distribution. The estimated value of the 99.5-th percentile of the true projected equity put value distribution, $V_t^{[j']}$, was found to be 0.4467. The values of $\mathcal{M}_2$ for each of the proxy functions considered earlier are given in Table 3.2.

Regression error metric $\mathcal{M}_2$:	
Proxy Funct.	$\mathcal{M}_2$
f_1	0.0242
f_2	0.0021
f_3	0.0033
f_4	0.0029

Table 3.2: Regression error metric $\mathcal{M}_2$ for different proxy functions.

Comparing the proxy function f_2 with f_1 it is not that surprising that the corresponding value of $\mathcal{M}_2$ is significantly smaller. The plots of the regression surfaces show that f_2 is a much better approximation to the true projected value surface, including in the upper tail. On the other hand when we move from the proxy function f_2 to f_3 or f_4 the value of $\mathcal{M}_2$ increases. This is surprising, as the regression surface for f_3 certainly appears to be a much closer fit to the true projected value surface than the regression surface of f_2 for the vast majority of values of S_1 and σ_1 considered here. There is, however, a fairly simple explanation of why this is, which will now be discussed.

When we talk about the 99.5-th percentile of the true projected value distribution this is simply a certain projected option value – in our example a fairly accurate estimate of this is 0.4467. There are various combinations of the two risk-drivers, S_1 and σ_1, which will give this projected option value analytically. Indeed, in Figure 3.8

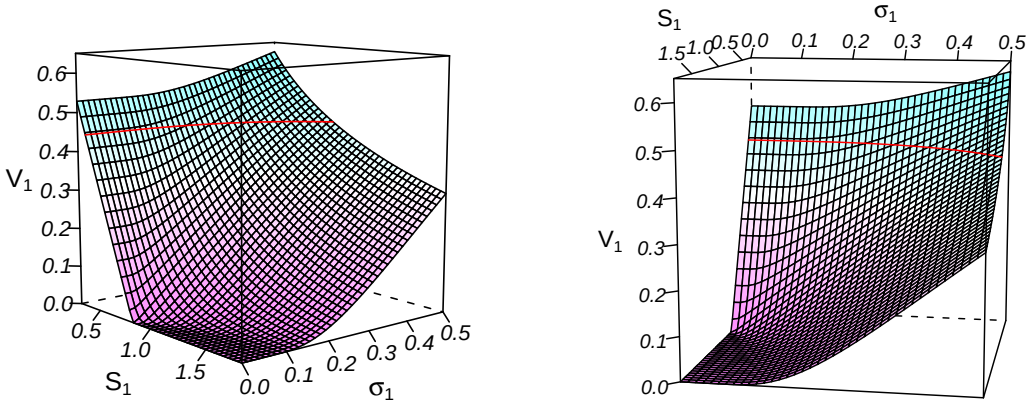

Figure 3.8: Two views of the analytic surface of V_1. The red line shows the contour at the level of the 99.5-th percentile of the true projected value distribution.

we plot the analytic projected value surface again (with the usual two views), but this time we superimpose a contour line on the surface showing where this surface takes the value $\hat{V}_1 = 0.4467$. It is clear that this line spans many different values of S_1 and σ_1. A contour plot of this surface, given in Figure 3.9, perhaps shows this more clearly. We can see that the 99.5-th percentile of the true projected value distribution could be a result of a value of σ_1 from 0 to 0.5 (the whole range of volatilities we consider). Also, there are certain values of σ_1 with which any value of S_1 from just below 0.4 to around 0.8 will give the 99.5-th percentile of the projected option value.

Thus, a realisation of the 99.5-th percentile of the true projected value distribution can be a result of many different possible pairs (S_1, σ_1). Of course, the difference between the regression estimate and the true percentile projected value will be different for all these different pairs. In fact Figure 3.10 shows the contour plot of the surface of the differences between the regression proxy function f_3 and the true projected value distribution, again, but this time with a line showing the S_1 and σ_1 pairs which generate the upper percentile superimposed. We see that along this line there are points where the regression surface underestimates the percentile and points where it overestimates it (as well as three points where it estimates it perfectly).

Bearing this in mind, it is not hard to appreciate that choosing the one pair (S_1, σ_1)

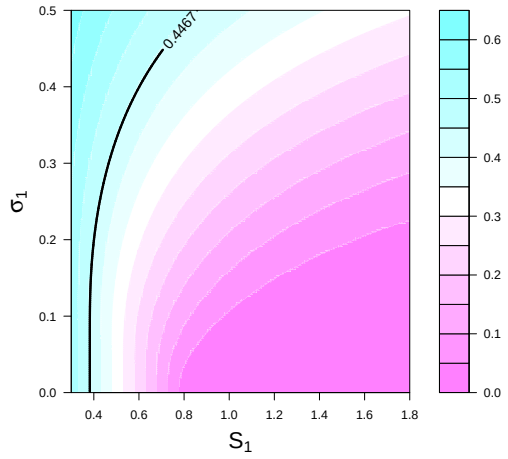

Figure 3.9: Analytic surface contour plot. 99.5-th percentile contour of true distribution shown in black.

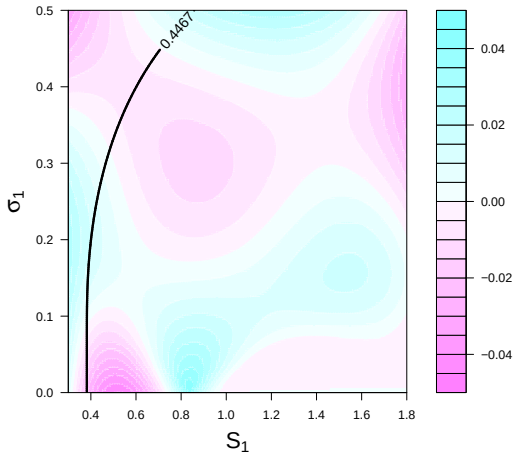

Figure 3.10: Contour plot of difference surface $\hat{V}_1 - V_1$ with proxy f_3. Black line, same as in Figure 3.9.

as being 'the' projection which generated the upper percentile of the value V_1 could give a quite different estimate of $\mathcal{M}_2$, than if we chose a different pair (S_1, σ_1) lying on this upper percentile contour line. Naturally, when comparing two different proxy functions, one of them could happen to be a pair where the corresponding regression surface and analytic surface are relatively close to one another. With the other proxy function, we could choose a pair where the two surfaces happen to have diverged away from each other slightly. Despite the fact that along this upper percentile contour line the two surfaces could be closer together than for the other proxy function on average, the metric $\mathcal{M}_2$ will indicate that the first proxy function is a preferable regression model. Therefore, this metric is not a good statistic with which to measure the fit in the upper tail of the projected value distribution.

We will now attempt to modify the definition of the metric $\mathcal{M}_2$ so that it still gives an indication of the fit of regression proxy functions around the upper 99.5-th percentile of the true projected value distribution, but is much more robust than it is in its current form. The obvious problem is that the metric is based upon a single pair (S_1, σ_1) which lies on the upper percentile contour line. The metric would be much more robust if we could base it on a larger sample of the pairs (S_1, σ_1) on which this contour line lies. As we use a reasonably large sample of $R = 50,000$ real world economic projections in calculating the metric $\mathcal{M}_2$, rather than just choosing the single pair (S_1, σ_1) which corresponds to the $49,750$-th largest analytic projected

value, let us now select a set of risk-driver pairs around this percentile. This will involve selecting the pairs corresponding to the $49,650$-th through to the $49,850$-th of the ordered analytic projected values. With such a large number of real world scenarios this will still give a reasonably close approximation to the specific upper percentile that we wish to consider, however now we will hopefully get a selection of many different risk-driver pairs (S_1, σ_1). We then take the difference between these analytic values and the regression estimates from feeding these real world pairs into the proxy function being tested. The absolute value of each of these 201 differences between the analytic surface and regression will then be averaged to give us an adjusted version of the regression error metric $\mathcal{M}_2$, which we denote $\mathcal{M}_3$:

$$\mathcal{M}_3 = \frac{1}{201} \sum_{i=-100}^{100} |\widetilde{V}_t^{k(i)} - V_t^{[0.995R+i]}|. \tag{3.17}$$

The terms $\widetilde{V}_t^{k(i)}$ and $V_t^{[j]}$ are defined as in the discussion given before Equation 3.10. In Table 3.3 we consider how this metric rates each of our four proxy functions:

Regression error metric $\mathcal{M}_3$:	
Proxy Funct.	$\mathcal{M}_3$
f_1	0.0137
f_2	0.0058
f_3	0.0030
f_4	0.0028

Table 3.3: Regression error metric $\mathcal{M}_3$ for different proxy functions.

These values for $\mathcal{M}_3$ are more consistent with what was seen when we studied the plots of the regression surfaces in comparison with the true projected value surface. Thus, using this adjusted metric, we can say the proxy functions f_3 and f_4 are preferable to the proxy functions f_1 and f_2 when our only interest in the regression is to obtain a reasonably good approximation around the vicinity of the upper 99.5-th percentile of the true projected value distribution. Overall, the metrics $\mathcal{M}_1$ and $\mathcal{M}_3$ suggest that choosing a fixed form of regression model either by considering the significance of each term in the model or using a stepwise AIC algorithm is a sensible way to implement the LSMC method. Bauer, Bergmann and Reuss [Bau10] suggest using Mallows' complexity parameter, C_p, as an alternative statistic for model ranking in the proxy function selection algorithm. See Section 4.6 for more discussion of this idea.

3.5 Issue of statistical over-fitting

In studying different possible proxy functions for the regression in the LSMC method in Sections 3.3 and 3.4, it became clear that the order of the basis functions required for a good fit to the projected put option value function needs to be greater than just one or two. Indeed, when an Akaike information criterion-based stepwise regression selection algorithm was employed a sixth-order term was included in the 'best fitting' regression function.

Theoretically if we were to keep adding higher order basis functions to the chosen proxy function, the fit from the regression to the outer fitting points should keep improving. This is not the same thing as the fit of the regression model to the true underlying distribution of projected option values improving, though, since the estimated value of the option at each fitting point has some associated sampling error. Thus, by choosing too complex a regression model, one can end up fitting to the sampling error as well as the underlying option value at each of the fitting points. This is the problem of statistical over-fitting mentioned earlier. Using an AIC-based approach to select the best regression model should help overcome this problem, as it ranks regression models by their value of log-likelihood with a penalty applied for each term included in the proxy function. Thus, under this approach for selecting the proxy function, extra terms will only be added to the model if they can be justified by increasing the accuracy of the fit by enough.

To study whether the issue of over-fitting applies to the LSMC algorithm and at what point this phenomenon seems to occur, let us simplify our test case somewhat. Rather than considering estimating the projected value of an equity put option with two risk-drivers by LSMC, we will now only consider this to be driven by a single risk-driver. To do this, we assume the underlying asset volatility is constant throughout the whole term of the option. Thus, the underlying equity level at the projection date, S_1, is now the sole explanatory variable in the LSMC method and the outer fitting points will only be one-dimensional.

Naturally, the real world outer scenarios will no longer need to be given by a stochastic volatility equity model. Instead now we will just assume the Black-Scholes model with annualised rate of growth μ as our real world distribution for equity returns. In full, the projection time is in one year's time, that is $t = 1$, and the maturity

of the option is in $T = 10$ year's time. As before we set $S_0 = 1$, $K = 1.3$ and $r = 5\%$. The real world growth in the Black-Scholes model, μ, will be taken as 3%. Of course, when the Black-Scholes model is simulated in the (risk-neutral) inner valuation scenarios, the growth will be equal to the risk-free rate $r = 5\%$. We will use 1,000 outer fitting scenarios and 20 inner valuation scenarios (per outer fitting scenario). To test the regression fits, we will generate 3,000,000 values of S_1 under the (real world) Black-Scholes model.

The outer fitting scenario points are sampled at regularly spaced intervals over $[0.3, 3.0]$. Three different choices of regression proxy function will be considered. The first, $g_1(S_1)$, will be found by using the stepwise AIC regression model selection function in R. The starting model will just be a constant plus S_1, and the maximal model for the scope of the search will be all terms in S_1 up to ninth order. When this is implemented, the model which was returned as the best fitting was $g_1(S_1) = c_0 + c_1 S_1 + c_2 S_1^2 + c_3 S_1^3 + c_4 S_1^4$. The second regression proxy which we will consider is all terms in S_1 up to ninth order, or $g_2(S_1) = c_0 + c_1 S_1 + \cdots + c_{10} S_1^{10}$. The third proxy function which will be studied will include all terms up to order 25, that is $g_3(S_1) = c_0 + c_1 S_1 + \cdots + c_{25} S_1^{25}$.

The fitted versions of these proxy functions give the LSMC estimate for the projected equity value at year one, $\hat{V}_1$, as a function of the risk-driver S_1. The function $\hat{V}_1$ is plotted for the $3,000,000$ values of S_1 obtained by simulating from the real world Black-Scholes model in the left-hand column plots of Figure 3.11. Also shown, as a red line, in these plots is the analytic value of the projected option, V_1, plotted as a function of S_1. The top plot in this column corresponds to regression proxy function g_1, the middle plot shows regression proxy function g_2 and the bottom plot displays proxy function g_3.

The plot of $\hat{V}_1$ for proxy function g_1 compared with the true distribution line of V_1, shows that this regression model fits the analytic projected value function very well. The plot of $\hat{V}_1$ for proxy function g_2 also appears to fit the true projected value distribution well.

The plots to the immediate right of these show the corresponding difference plots between the proxy function estimate and the true distribution, $\hat{V}_1 - V_1$, as a function of the risk-driver S_1. This shows the difference between the proxy function and the

true distribution is smaller for slightly more values of S_1 with g_1 compared to g_2. However, the difference between these fits is fairly small.

What if we were to use as high-order a proxy function as can be handled computationally. Would the resultant fit be better than with g_1 and g_2? If we look at the plot of $\hat{V}_1$ for proxy function g_3, the proxy function up to order 25, compared with the true distribution, there is evidence that this is not so. At first glance this appears similar to the fit of g_2, however if we look very closely at the black regression estimates they are almost oscillatory in nature for certain parts of the $\hat{V}_1$ distribution. This is certainly not the case for the distribution of V_1 which we are trying to approximate. This effect can be explained by the phenomenon of over-fitting. By considering too complex a model for $\hat{V}_1$, the proxy function is trying to fit to the outer fitting scenario points too closely. If the estimated value at any of those points lie somewhat above the true distribution, the proxy function will now locally try to move toward that point (even though it is lying away from the true distribution line simply because of sampling error). Then if a point lies slightly below the true distribution, the proxy function will try to move down to fit that point closely. This results in some regions of the distribution of $\hat{V}_1$ appearing almost oscillatory and is obviously not representative of the behaviour of the true projected value distribution.

The corresponding difference plot for proxy function g_3 seems to have much more extreme rises from a trough to a peak, and corresponding falls, than the corresponding plot for g_2. Perhaps this is an indication of the local oscillatory-like behaviour mentioned above and a further exhibition of over-fitting in the regression model g_3.

Let us now consider how our regression error metrics, $\mathcal{M}_1$ and $\mathcal{M}_3$, perform in the presence of possible over-fitting. The values of these two metrics for each of the proxy functions g_1, g_2 and g_3 are given in Table 3.4.

Regression error metrics:		
Proxy Funct.	$\mathcal{M}_1$	$\mathcal{M}_3$
g_1	0.00241	0.00042
g_2	0.00396	0.00197
g_3	0.00583	0.00128

Table 3.4: Regression error metrics $\mathcal{M}_1$ and $\mathcal{M}_3$ for proxy functions g_1, g_2 and g_3.

Clearly, $\mathcal{M}_1$ also seems to indicate over-fitting in proxy function g_3. There is a significant increase in the average absolute difference between the regression estimates and the analytic projected values across the whole real world S_1 distribution in going from proxy function g_2 to g_3. This is likely caused by the fact that g_3 is now trying to fit too closely to the outer fitting points, that it is fitting not just to the underlying true distribution, but also to the inherent statistical error in each of the fitting points simulated projected values. On the other hand, $\mathcal{M}_3$ does not seem to identify over-fitting in the complex regression model. Although the value of $\mathcal{M}_3$ is smaller for proxy function g_1 than g_2, the value of $\mathcal{M}_3$ for g_3 is also smaller than that for proxy g_2. This is perhaps not surprising, as this metric only considers a very small part of the $\hat{V}_1$ distribution, hence it is not likely to be a good indicator of statistical over-fitting in our regression model. The table below shows the values of these metrics for all orders of proxy function from 2 to 25. It shows over-fitting seems to occur with orders greater than about 5. Notice that the best model by AIC does not exactly correspond to the best model in terms of $\mathcal{M}_1$. However, these two models have similar values of $\mathcal{M}_1$ and AIC, indicating that the stepwise regression approach does seem to give a method of selecting the form of the regression proxy function which is robust to over-fitting.

Regression error metrics for different orders of proxy function:							
Order	$\mathcal{M}_1$	$\mathcal{M}_3$	AIC	Order	$\mathcal{M}_1$	$\mathcal{M}_3$	AIC
2	0.0295	0.0084	-22553.4	14	0.0054	0.0007	-23421.0
3	0.0058	0.0083	-23343.2	15	0.0052	0.0017	-23419.7
4	0.0024	0.0004	-23424.8	16	0.0053	0.0011	-23418.7
5	0.0022	0.0005	-23423.3	17	0.0054	0.0014	-23418.5
6	0.0035	0.0017	-23424.8	18	0.0054	0.0010	-23418.7
7	0.0039	0.0023	-23424.2	19	0.0055	0.0014	-23416.8
8	0.0039	0.0023	-23422.2	20	0.0056	0.0012	-23417.8
9	0.0041	0.0026	-23420.3	21	0.0056	0.0005	-23418.1
10	0.0040	0.0020	-23418.5	22	0.0056	0.0003	-23417.1
11	0.0042	0.0001	-23420.4	23	0.0056	0.0007	-23416.9
12	0.0054	0.0002	-23424.5	24	0.0060	0.0023	-23420.1
13	0.0055	0.0003	-23422.9	25	0.0058	0.0013	-23419.2

Table 3.5: The $\mathcal{M}_1$ and $\mathcal{M}_3$ errors for different orders of polynomial regression in the variable S_1 with **1,000 outer and 20 inner scenarios** being simulated.

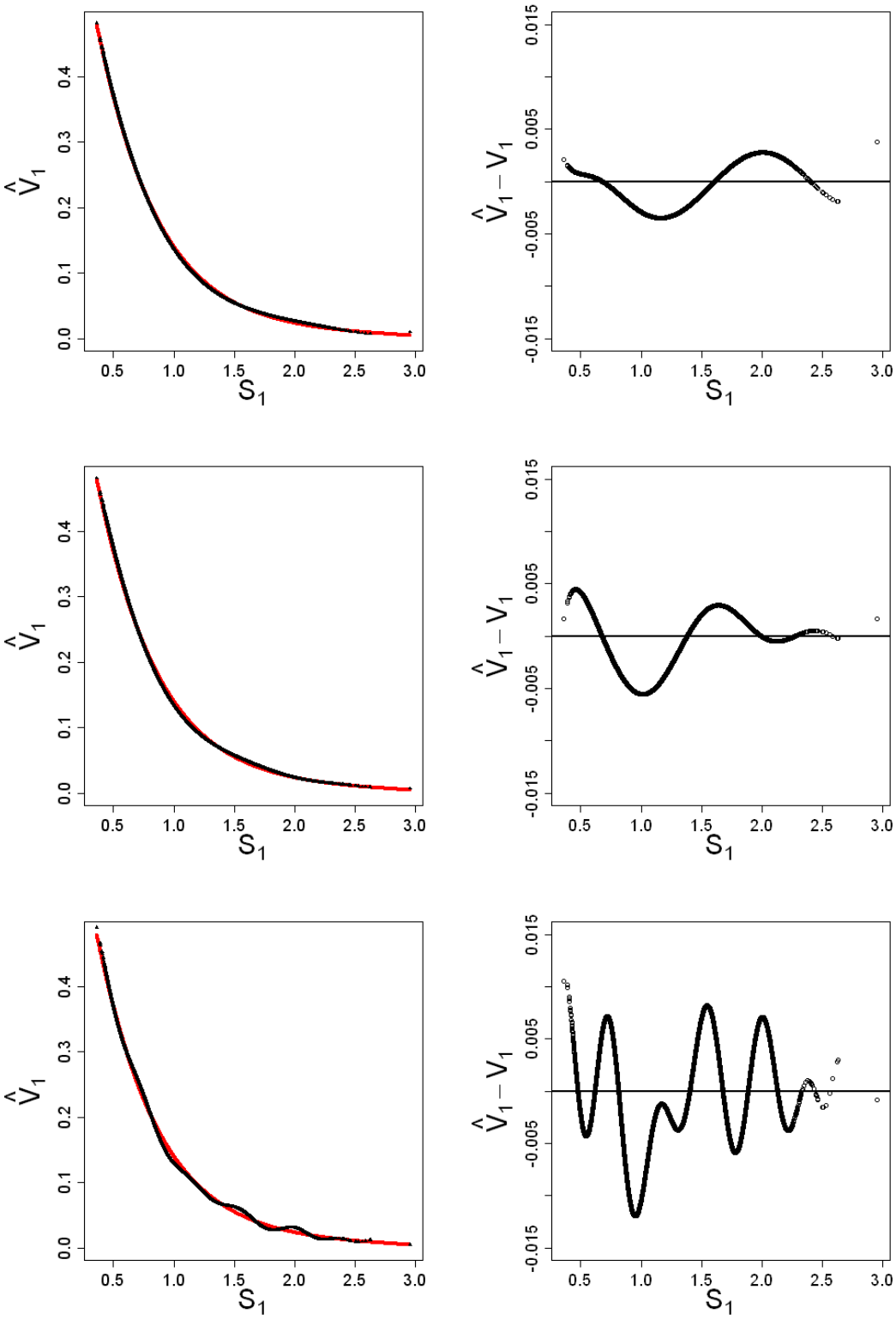

Figure 3.11: Left-hand column plots: Regression est.'s for 3,000,000 real world outer scenarios shown as black triangles, with analytic value of the option, V_1, shown by the red line. Right-hand column: Differences between the regression estimates and analytic values for each of the 3,000,000 real world outer scenarios. Regressions: In the first row $g_1 = c_0 + c_1 S_1 + c_2 S_1^2 + c_3 S_1^3 + c_4 S_1^4$. In the second row $g_2 = c_0 + c_1 S_1 + \cdots + c_{10} S_1^{10}$. In the third row $g_3 = c_0 + c_1 S_1 + \cdots + c_{25} S_1^{25}$.

3.6 Over-fitting and the number of outer/inner scenarios

In the analysis just given, we fixed the number of outer and inner scenarios and then increased the complexity of the regression model used in the LSMC method to investigate the issue of over-fitting. This approach, however, does not tell us anything about the relationship between the number of outer or inner scenarios being generated and the level of complexity of the regression model at which over-fitting occurs.

In this section some analysis will be performed to determine whether the number of inner scenarios being sampled in the LSMC method has an influence over the level of complexity in the regression model beyond which there is evidence of over-fitting. Also, we will look to test whether the number of outer scenarios being sampled has an influence over the same issue.

Recall, statistical over-fitting occurs when the regression model includes a great enough degree of complexity it begins to fit to the sampling error associated with each of the estimates, as well as to the true underlying projected value distribution. If we increase the number of these inner valuation scenarios being performed, the sampling error associated with each of these projected option value estimates will decrease. As such, it seems reasonable that with a smaller variance associated with each of these estimated values the regression proxy function will be less likely to fit to this. Thus, intuitively, if a greater number of inner scenarios are performed, then a more complex regression model will be necessary before over-fitting can occur.

In Table 3.6 the $\mathcal{M}_1$ and $\mathcal{M}_3$ errors are, again, given for different orders of polynomial regression model in the variable S_1, but now the number of inner scenarios has been increased from 20 to 100, in comparison to Table 3.5. The number of outer scenarios has been kept at 1,000. This table shows that the metric $\mathcal{M}_1$, which we argued in the previous section was a good indicator of over-fitting, suggests that orders of polynomial beyond 10 or 11 are too complex and are fitting to the sampling error of the estimates. Recall, that for 1,000 inners and 20 inners, the order of polynomial beyond which this was the case was 5. Therefore, this simple test gives evidence to justify our hypothesis that increasing the number of inner valuation scenarios allows a more complex regression model to be employed before the phenomena of statistical over-fitting becomes a problem.

Regression error metrics for different orders of proxy function:							
Order	$\mathcal{M}_1$	$\mathcal{M}_3$	AIC	Order	$\mathcal{M}_1$	$\mathcal{M}_3$	AIC
2	0.0304	0.0082	-110249	14	0.0009	0.0025	-114006
3	0.0066	0.0082	-113714	15	0.0008	0.0036	-114008
4	0.0011	0.0014	-114007	16	0.0009	0.0036	-114006
5	0.0009	0.0009	-114005	17	0.0010	0.0036	-114004
6	0.0008	0.0020	-114008	18	0.0010	0.0036	-114004
7	0.0007	0.0025	-114010	19	0.0011	0.0038	-114002
8	0.0005	0.0022	-114010	20	0.0011	0.0037	-114002
9	0.0004	0.0018	-114009	21	0.0011	0.0038	-114000
10	0.0002	0.0013	-114007	22	0.0011	0.0037	-114000
11	0.0002	0.0011	-114006	23	0.0011	0.0037	-114000
12	0.0006	0.0010	-114006	24	0.0011	0.0039	-113998
13	0.0008	0.0014	-114005	25	0.0011	0.0038	-113998

Table 3.6: The $\mathcal{M}_1$ and $\mathcal{M}_3$ errors for different orders of polynomial regression in the variable S_1 with **1,000 outer and 100 inner scenarios** being simulated.

In Table 3.7 the values of $\mathcal{M}_1$ and $\mathcal{M}_3$ are given for different orders of polynomial regression for the case of 10,000 outer scenarios and 20 inner scenarios. The results can be compared to the results of 1,000 outers and 20 inners, given in Table 3.5, to investigate how increasing the number of outer fitting scenarios affected the point at which over-fitting seemed to occur. The $\mathcal{M}_1$ metric suggests that for 10,000 outer fitting scenarios and 20 inner scenarios over-fitting seems to occur for polynomial regression models beyond order 10. This is of much greater complexity than order 5, which was found to be the corresponding level for 1,000 outer and 20 inner scenarios. Thus, this test gives evidence that increasing the number of outer fitting scenarios allows a more complex regression model to be employed before over-fitting becomes an issue.

Similar tests reducing the number of outer and inner scenarios showed that the level of complexity at which over-fitting occurred was lower than was the case for 1,000 outer and 20 inner scenarios. This is, again, in correspondence with the hypotheses given at the beginning of this section.

3.7 Fitting point sampling in LSMC

When implementing the LSMC method, the natural approach for the sampling of the fitting scenarios might appear to be to use the real-world distribution which we believe to give a realistic description of the behaviour of the risk-drivers from today

Regression error metrics for different orders of proxy function:							
Order	$\mathcal{M}_1$	$\mathcal{M}_3$	AIC	Order	$\mathcal{M}_1$	$\mathcal{M}_3$	AIC
2	0.0298	0.0082	-221379	14	0.0009	0.0014	-228932
3	0.0065	0.0083	-228328	15	0.0009	0.0012	-228930
4	0.0012	0.0016	-228919	16	0.0008	0.0014	-228929
5	0.0006	0.0001	-228931	17	0.0008	0.0014	-228927
6	0.0006	0.0010	-228935	18	0.0008	0.0014	-228927
7	0.0005	0.0014	-228938	19	0.0008	0.0017	-228926
8	0.0004	0.0013	-228937	20	0.0008	0.0015	-228925
9	0.0004	0.0011	-228935	21	0.0008	0.0019	-228925
10	0.0003	0.0010	-228933	22	0.0008	0.0018	-228925
11	0.0004	0.0007	-228932	23	0.0008	0.0017	-228924
12	0.0007	0.0007	-228932	24	0.0008	0.0019	-228923
13	0.0009	0.0011	-228933	25	0.0008	0.0018	-228923

Table 3.7: The $\mathcal{M}_1$ and $\mathcal{M}_3$ errors for different orders of polynomial regression in the variable S_1 with **10,000 outer and 20 inner scenarios** being simulated.

out until the projection date. After all, the regression model which is found will be used to infer a liability based on this real-world distribution, so why fit the regression using any other distribution? In Section 2.4 it was argued that sampling from other statistical distributions may be beneficial in the application of the LSMC technique in a risk-management setting. In this section, we will perform some simple tests using the projected put value framework to show that it is, indeed, preferable to sample the fitting points from other distributions in this context.

For this analysis we will work with the two-dimensional projected put value problem which was studied earlier in this chapter. The real-world distribution for the behaviour of the two risk-drivers will be the Heston stochastic volatility model - which will give us simulated values of the equity price, S_1, and the equity volatility, σ_1, at the projection time at year one.

Histograms of 50,000 real-world scenarios sampled from the Heston model are shown in Figure 3.12 for each of the two risk-drivers. In the top-left of this Figure is a histogram of the year one equity levels and the top-right Figure displays a histogram of the year one volatilities. These plots show the Heston model has a reasonably thick lower tail for the year one equity returns and a reasonably thick upper tail for the year one volatility levels. From the surface plots shown earlier in this chapter, relatively low equity returns together with relatively large volatilities is what yields the largest projected option values. Thus, the projected option value distribution will also have a reasonably heavy upper tail.

This projected option valuation task is being used to guide us in how successful the LSMC might be in terms of calculating insurance solvency capital requirements and how the method should be optimised. Naturally, if a good fit can be achieved across the whole liability distribution, then this would be the most successful outcome in the LSMC method. However, when this is not possible then achieving a good fit in the upper tail takes a far greater priority over the fit in the lower tail or centre of the distribution. The plot in the bottom-left of Figure 3.12 shows 50,000 year one equity and volatility pairs sampled from the Heston stochastic volatility model. Superimposed on top of these points is the contour line of the 99.5-th percentile of the projected option value distribution, exactly as was plotted in Figures 3.9 and 3.10. The plot in the bottom-right of this Figure shows the same, except with just 1,000 sampled pairs from the Heston model. These plots show that when sampling from the real-world model used to project the two risk-drivers to year one, there will not be a large amount of points which will yield projected option values around this 99.5-th percentile. Of course, by definition there will only be a small number of real-world scenarios which yield values in the upper tail of this projected option value distribution. Therefore, if the upper tail is of key concern to practitioners, perhaps it is not wise to use this real-world distribution to fit the regression proxy function in the LSMC method. Using some other distribution to sample these fitting points could generate more scenarios which will yield valuations around this upper 99.5-th percentile, compared with using the real-world model.

To demonstrate this fact, 1,000 equity and volatility pairs were sampled from a uniform distribution, which was scaled and shifted to span the region of this two-dimensional risk-driver space which the real-world model will generate realisations within, to some very high level of confidence. These 1,000 risk-driver pairs are plotted alongside the 99.5-th projected value percentile contour line in Figure 3.13 (the left-hand side plot of this Figure). Comparing this plot with the bottom-right plot of Figure 3.12, shows that using simple uniform sampling of these risk-drivers will give more scenarios which yield option valuations around the 99.5-th percentile, than if we use the real-world Heston model for this. In Figure 3.13 (right-hand side plot), 1,000 fitting scenario pairs generated from a two-dimensional Sobol quasi-random number scheme are displayed alongside this upper percentile contour line. Again, sampling these fitting points using quasi-random numbers over this subset

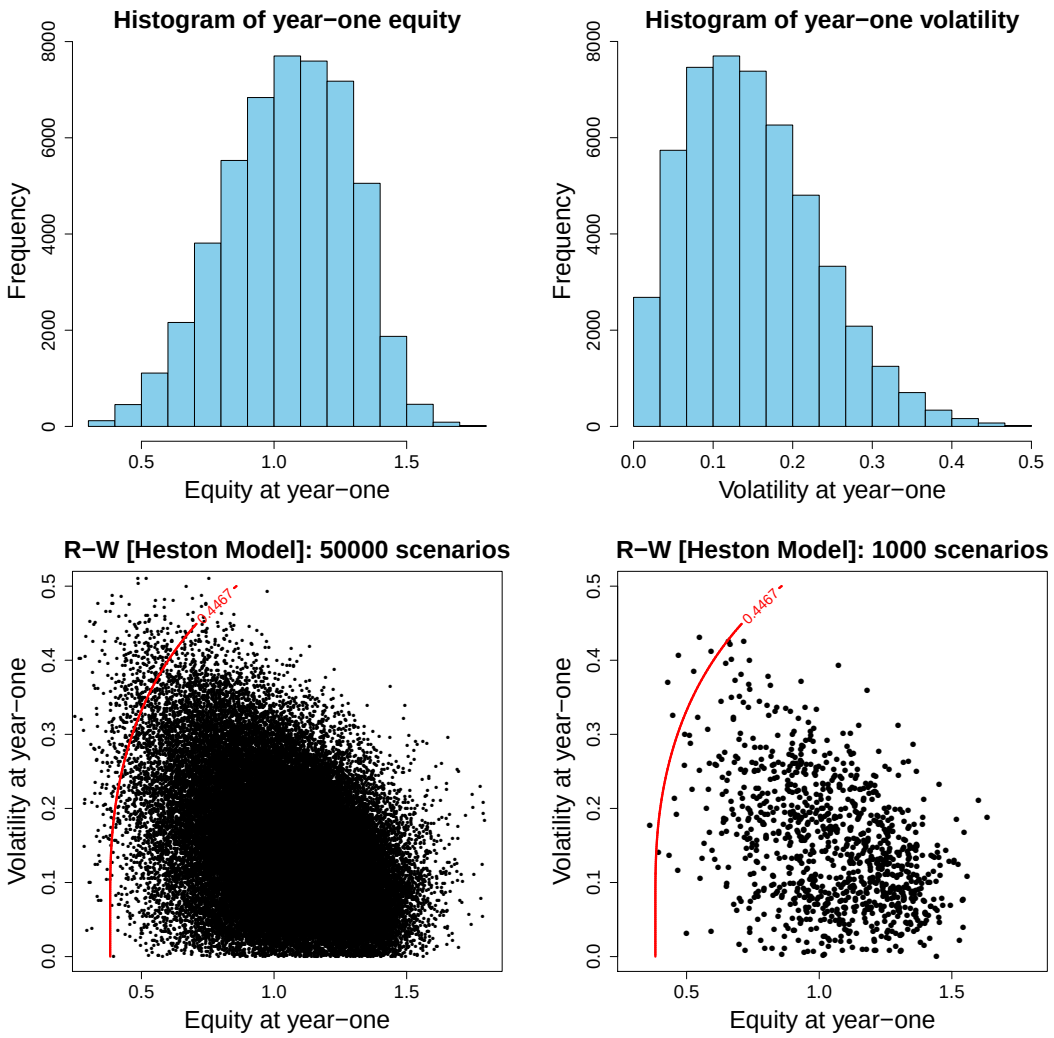

Figure 3.12: Top row plots: Histograms of the year one equity values (left-side) and year one volatilities (right-side) sampled from the Heston stochastic volatility model. Bottom row plots: Left-side plots 50,000 year one equity and volatility pairs sampled from the Heston model, with the analytic projected option value 99.5-th percentile from Section 3.4 superimposed (red line). Right-side plot shows the same, but with only 1,000 equity and volatility pairs.

of risk-driver space results in a much greater proportion of fitting scenarios in the vicinity of the upper tail, in comparison with using the real-world distribution for this purpose.

The discussion up until now seems to indicate that generating the fitting scenarios used to calibrate the regression proxy function in the LSMC method via a uniform distribution or by quasi-random sampling may give a better fit in the upper tail of the projected value distribution, compared with using the real-world model. However, this is just a hypothesis and needs to be investigated. In order to test different sampling methods for these regression fitting scenarios, we generate 961 pairs of year

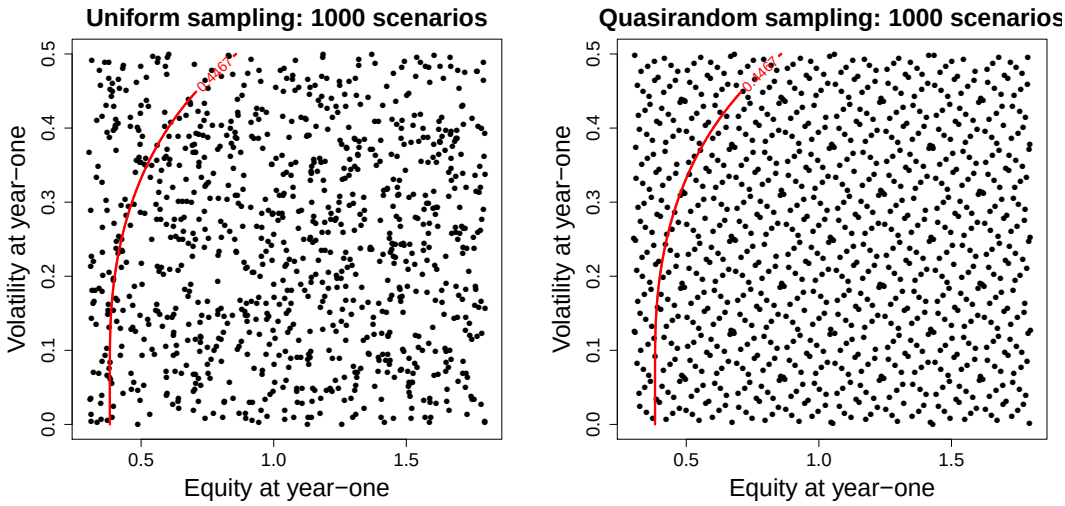

Figure 3.13: Left-side plots 1,000 year one equity and volatility pairs sampled uniformly, with the analytic projected option value 99.5-th percentile from Section 3.4 superimposed (red line). Right-side plot shows the same, but with 1,000 equity and volatility pairs generated using Sobol quasi-random numbers.

one equity and volatility using the Heston stochastic volatility model; (independent) uniform distributions; Sobol quasi-random numbers; Latin Hypercube sampling; and placing the fitting points in a fixed grid of even spacing across the subset of risk-driver space in which the real-world distribution typically generates values. In these tests we will use the same option parameters given at the end of Section 3.1 and employ our metrics, $\mathcal{M}_1$ and $\mathcal{M}_3$, for determining the success of the LSMC method.

In Figure 3.14 the LSMC $\mathcal{M}_1$ error is plotted for each of the five different methods of fitting point sampling. Recall that the $\mathcal{M}_1$ error metric took the average of the absolute differences between the analytical projected option value and LSMC estimates for 50,000 real-world (Heston model) test scenarios. This plot shows that if one wants a good fit across the whole projected option value distribution, it is slightly preferable to sample the fitting points using the real-world (Heston) model. Since we are also generating the 50,000 test scenarios from the Heston model this result makes sense. Choosing the fitting scenarios from the same distribution that we sample the test scenarios is going to give the best fitting regression proxy function (and hence most accurate LSMC estimates) when accuracy is measured across the whole distribution.

On the other hand, if we consider the fit of the regression proxy function only in the upper tail, then the same result need not necessarily hold. Indeed, in Section

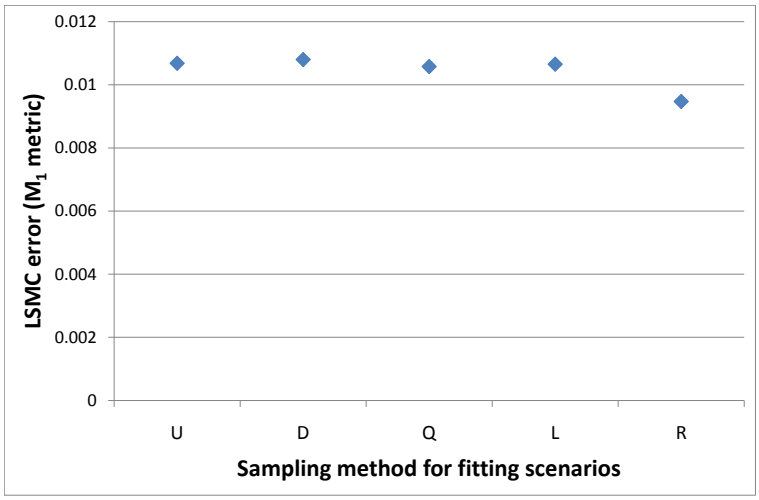

Figure 3.14: Plot of error in LSMC estimates across the whole projected option value distribution (metric $\mathcal{M}_1$) for different methods of sampling fitting scenarios. U: uniform; D: discrete grid; Q: quasi-random; L: Latin hypercube; R: real-world.

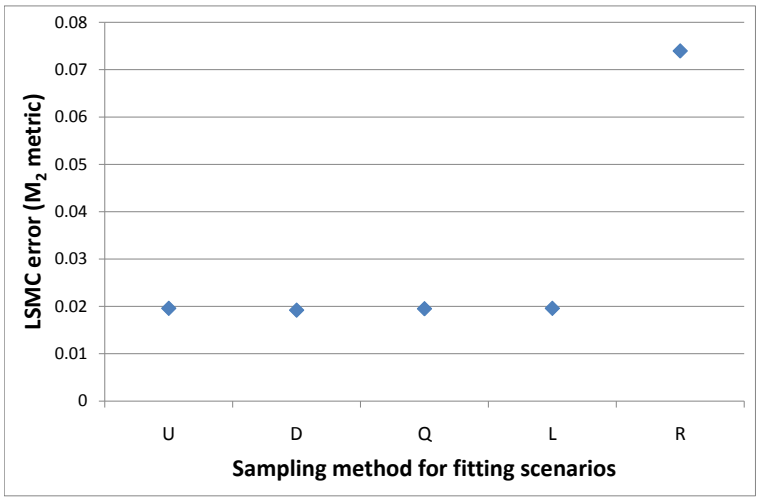

Figure 3.15: Plot of error in LSMC est's around the 99.5-th percentile of projected option value dist'n (metric $\mathcal{M}_2$) for different methods of sampling fitting scen's. U: uniform; D: discrete grid; Q: quasi-random; L: Latin hypercube; R: real-world.

2.4 we argued that it would be reasonable to expect a better fit in the upper tail of the projected value distribution when sampling the fitting scenarios uniformly or using quasi-random sampling, in comparison with sampling from the real-world (Heston) model. In Figure 3.15 the LSMC $\mathcal{M}_3$ error is plotted for each of the five different methods of fitting point sampling. Recall that the $\mathcal{M}_3$ error metric took the average of the absolute differences between the analytical projected option value and LSMC estimates for scenarios which yield projected values in the close vicinity of the 99.5-th percentile. Figure 3.15 shows that using the real-world Heston model to sample the fitting scenarios for the regression proxy function does, indeed, give poorer LSMC estimates in the upper tail of the projected option value distribution, compared with the other sampling approaches. Thus, the above hypothesis seems to be correct.

In Figure 3.14 we saw that using the real-world model to sample the fitting points gave (by a fairly slight margin) the best fitting regression proxy function across the whole projected value distribution. Yet, in Figure 3.15 we saw that using the real-world distribution for this purpose, gave much poorer LSMC estimates in the upper tail of the projected value distribution, compared with other sampling methods. So, which method for sampling the fitting points should be preferred? Well, the fit across the whole projected value distribution is only marginally better using the real-world distribution to sample the fitting points, while the fit in the upper tail is far poorer using this sampling approach. Furthermore, the fit in the upper tail is of key concern to insurance practitioners when calculating a solvency capital requirement. Therefore, this all suggests that when using the LSMC method in calculating a solvency capital requirement, one would be better off sampling the fitting scenarios from one of the alternative schemes tested, other than the real world (Heston) model.

But which one of these other distributions should be chosen for sampling the fitting scenarios? In Figure 3.16 an enlarged plot of the $\mathcal{M}_1$ errors in the LSMC estimates for sampling types uniform, discrete grid, quasi-random and Latin hypercube is given. Although there is not a great difference between these errors, this Figure suggests that employing quasi-random sampling will give the best fitting regression function across the whole projected value distribution, followed closely by Latin hypercube and uniform sampling. Discrete grid sampling is somewhat poorer in the subsequent LSMC fit across the whole distribution. The reason for this will be dis-

cussed in a moment. In Figure 3.17, a similar plot is given for the $\mathcal{M}_3$ errors with these four sampling types for the fitting scenarios. This plot suggests that if one is only interested in the accuracy of the fitted proxy function in the upper tail of the projected value distribution, the best approach for generating the fitting scenarios is to place them in a discrete grid in risk-driver space. There is no significant difference between the other sampling approaches in terms of the success of the LSMC estimates in the upper tail. The discrete grid sampling approach places a relatively large proportion of the total outer scenarios in the corner of the risk-driver hypercube over which we sample. For the projected put valuation problem, the upper tail of the distribution comes from the scenarios in one of these corners, and hence this sampling approach provides the best fit in the upper percentile. Conversely, because of this disproportionate number of points in the corner of this hypercube, there is relatively fewer of them in the middle of the distribution. This explains the poorer fit around the centre of the projected value distribution from using discrete grid sampling. This increased accuracy in the upper quantiles and decreased accuracy around the median in comparison to real-world sampling is also seen for uniform, quasi-random and Latin hypercube fitting point sampling and is explained similarly. This is analogous to importance sampling, where simulating more scenarios from the tail of the distribution will increase the accuracy of estimators in this region, but at the expense of the accuracy estimators based on the centre of the distribution.

In more complex insurance problems, it might not be so easy to determine which part of risk-driver space will yield the largest liability values. Also, the risk-drivers which yield the largest liabilities might not even be concentrated in a localised part of risk-driver space, as is the case for the simple put option. Furthermore, placing the points in a discrete grid is only sensible for a low number of dimensions, perhaps up to three. As the dimensionality of the problem grows, the total number of scenarios grows exponentially. This total number will become computationally infeasible for situations when the dimensionality of the problem is not low. Thus, when implementing the LSMC method in a typical insurance setting, employing discrete grid sampling of the fitting scenarios is likely to be impractical. In such cases any of the alternative sampling methods should offer a similar degree of accuracy in fit in the upper tail, assuming the analysis from the projected put option valuation problem can be extended to the insurance setting.

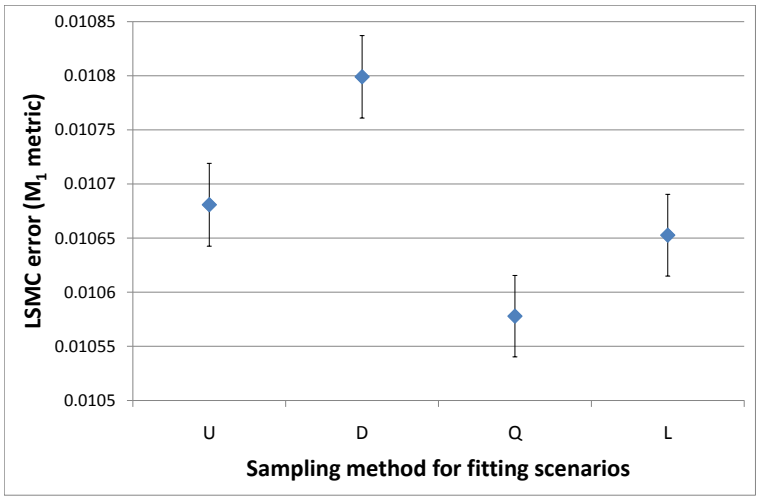

Figure 3.16: Enlarged plot of error in LSMC estimates across the whole projected option value dist'n (metric $\mathcal{M}_1$) for different methods of sampling fitting scenarios. U: uniform; D: discrete grid; Q: quasi-random; L: Latin hypercube.

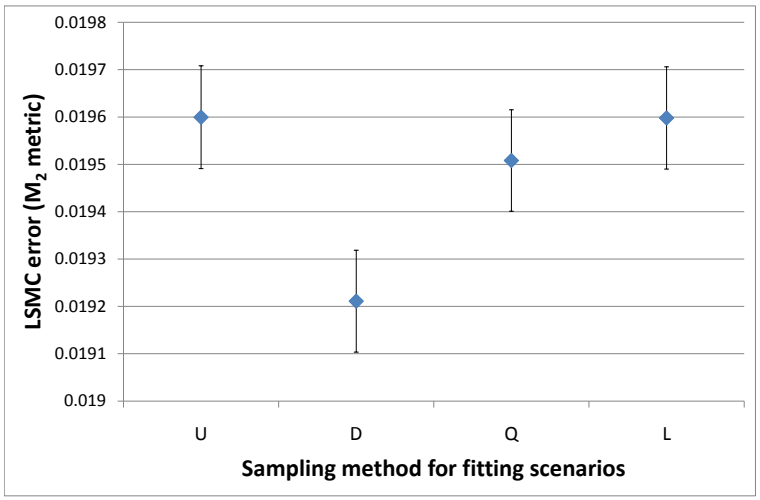

Figure 3.17: Plot of error in LSMC estimates around the 99.5-th percentile of the projected option value dist'n (metric $\mathcal{M}_2$) for different methods of sampling fitting scenarios. U: uniform; D: discrete grid; Q: quasi-random; L: Latin hypercube.

3.8 Form of basis functions in LSMC

In the previous section, the choice of the sampling method for the fitting scenarios in the LSMC method was investigated. In that analysis, the regression proxy function was composed of standard polynomial functions. That is, for the two risk-drivers in this projected option valuation problem, S_1 and σ_1, the regression model consisted of simple powers and cross-terms of these variables. Of course, the terms which make up the regression proxy function, which we refer to as the basis functions, need not just be simple powers and cross-terms. Indeed, there is a wide class of functions which would make a sensible choice of form for these basis functions which are known as orthogonal polynomials. An introduction to the use of orthogonal polynomials and why they may improve the accuracy of a regression was given in Section 2.5. The purpose of this section is to perform some simple tests to ascertain whether the use of orthogonal polynomials improves the performance of the LSMC method. Furthermore, a few different families of orthogonal polynomials will be tested to see whether this choice has an effect.

Under a standard polynomial regression, the value of the option is described in terms of a simple polynomial function as

$$f(S_1, \sigma_1) = a_{0,0} + \sum_{i=0}^{I} a_{i,0} S_1^i + \sum_{j=0}^{J} a_{0,j} \sigma_1^j + \sum_{p=1}^{P} \sum_{q=1}^{Q} a_{p,q} S_1^p \sigma_1^q, \tag{3.18}$$

where the $a_{m,n}$ are constants (to be determined). The parameter I is the maximum power of S_1 and J is the maximum power of σ_1 in the regression model. The parameters P and Q are the maximum powers that S_1 and σ_1, respectively, can take in a cross-term of the regression model.

This set-up can be generalised to allow for the basis functions to be constructed from the set of some orthogonal polynomial, as

$$f(S_1, \sigma_1) = a_{0,0} + \sum_{i=1}^{I} a_{i,0} B_i(S_1) + \sum_{j=1}^{J} a_{0,j} B_j(\sigma_1) + \sum_{p=1}^{P} \sum_{q=1}^{Q} a_{p,q} B_p(S_1) B_q(\sigma_1), \tag{3.19}$$

where $B_m(x)$ is the m-th-order orthogonal polynomial in the variable x. Examples of orthogonal polynomials which could be used in the regression model include Laguerre, Legendre, Chebyshev and Hermite polynomials.

The accuracy of the LSMC estimates will now be considered for the different combinations of fitting scenario sampling method and choice of orthogonal polynomial in the regression proxy function. The four different fitting scenario sampling approaches considered in Figures 3.16 and 3.17 are combined with standard polynomial, Chebyshev, Laguerre and Legendre basis functions to give 16 different possible LSMC set-ups. For each set-up, 10,000 independent repititions were performed and the $\mathcal{M}_1$ and $\mathcal{M}_3$ values will be calculated for each of these repetitions. This allows us to construct approximate 95% confidence intervals for the values of $\mathcal{M}_1$ and $\mathcal{M}_3$ under each of the 16 LSMC set-ups. These confidence intervals are shown in Figures 3.18 and 3.19 for the $\mathcal{M}_1$ and $\mathcal{M}_3$ metrics, respectively.

The plot in Figure 3.18 gives evidence that that using Legendre polynomials, particularly in combination with quasi-random sampling, can improve the accuracy of the LSMC method on average across the whole projected value distribution. The improvement in accuracy is relatively small however. Figure 3.19 shows similar findings for the accuracy of the LSMC method in the upper tail region of this distribution. An area of interesting further research is to investigate if these results extend to more complex products.

3.9 Optimal scenario budget allocation

One of the other fundamental questions in implementing the LSMC method is how the total computational budget should be allocated between generating the outer fitting scenarios and generating the inner valuation scenarios. This issue was introduced in section 2.6. It was argued there that although increasing either the number of outer or inner scenarios will increase the accuracy of the LSMC estimates, it is not immediately clear where the balance between the number of these scenarios should be drawn to make the method most efficient. We then assume that the general findings in this example will still apply when we consider a typical insurance product or book of liabilities and assume a more realistic real-world distribution. Bearing this in mind, it seems more reasonable to consider the computation time of generating the outer fitting scenarios and the inner valuation scenarios in the context of projecting complex insurance liabilities under realistic real-world and market-consistent distributions.

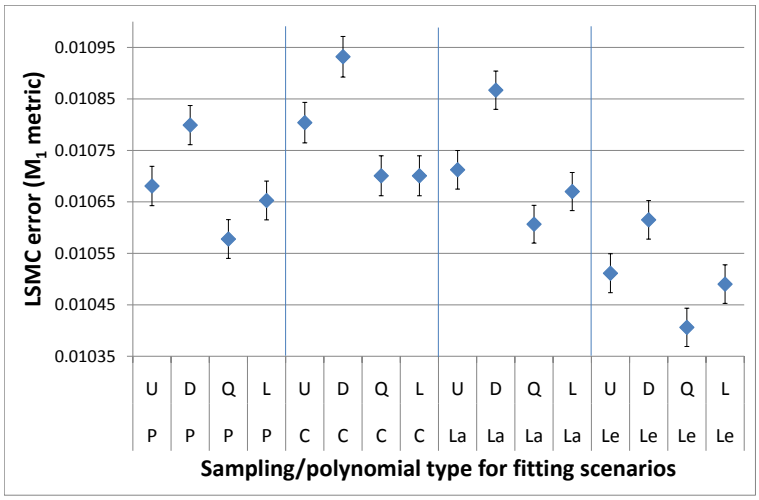

Figure 3.18: Enlarged plot of error in LSMC estimates across the whole projected option value dist'n (metric $\mathcal{M}_1$) for different methods of sampling fitting scenarios, U: uniform; D: discrete grid; Q: quasi-random; L: Latin hypercube and different choices of orthogonal polynomial, P: standard polynomial; C: Chebyschev; La: Laguerre; Le: Legendre.

In order to investigate this optimal scenario budget allocation problem we require some information about the relative time it takes to sample these different types of scenarios. We could calculate how long an outer and inner scenario takes for this projected put option set–up and then run different outer–inner allocations which have the same run-time. However, this would be specific to the form of the payoff of the European put option and our assumption of the Black-Scholes model for the valuation scenarios. Of course, the only reason we study a European put option under the Black-Scholes model is because it admits a simple analytical value for each outer scenario.

For typical insurance products the liability cashflows which need to be calculated in the inner valuation stage will often be complex, path-dependent and multi-dimensional in nature. Furthermore, these cashflows will often extend 30 or 40 years into the future. In short, calculating the liability for such insurance products is a difficult, path-dependent task which is very computationally demanding. On the other hand, sampling the outer fitting scenarios is a relatively efficient task. One may just have to generate a set of fitting points uniformly over some hypercube in risk-driver space. Even if these fitting scenarios were drawn from some realistic

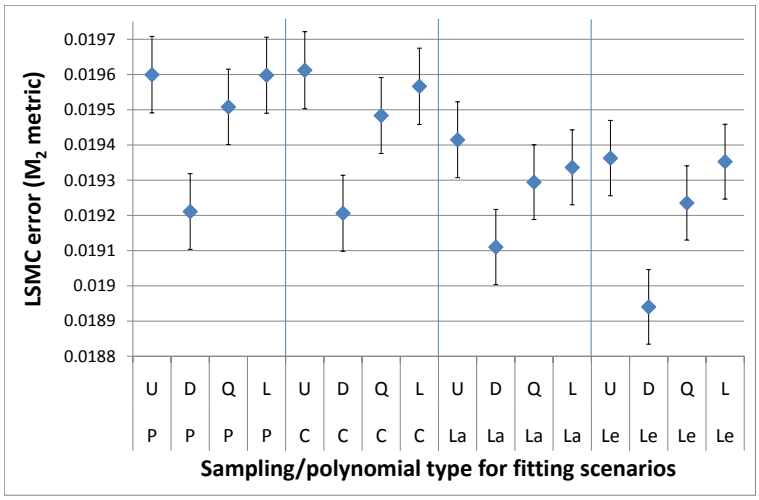

Figure 3.19: Plot of error in LSMC estimates around the 99.5-th percentile of the projected option value dist'n (metric $\mathcal{M}_2$) for different methods of sampling fitting scenarios, U: uniform; D: discrete grid; Q: quasi-random; L: Latin hypercube and different choices of orthogonal polynomial, P: standard polynomial; C: Chebyschev; La: Laguerre; Le: Legendre.

real-world model, such as the Heston stochastic volatility model, these scenarios will typically only be sampled out one-year into the future. By employing a fairly efficient sampling scheme, this would not be too onerous a task computationally. Thus, it seems perfectly reasonable to assume that generating a single fitting scenario will take far less time, than calculating the liability on some complex insurance product by Monte Carlo valuation. In fact, for typical insurance products on the market today, we can assume the time it takes to sample an outer fitting scenario is negligible compared with a liability valuation scenario. Given this assumption, the total computation time for the LSMC method will be given solely by the simulation of the inner scenarios. There are N of these per outer scenario and there are M outer scenarios altogether. Thus, we want to consider pairs of M and N, such that $M \cdot N = C$, where C is some fixed total number of inner valuation scenarios which expends our given computational budget. Thus, the question is what ratio of M to N will yield the most accurate LSMC estimated projected liability distribution?

To answer this we consider the projected put value problem, as this yields an analytical projected value for each outer real-world test scenario, and hence allows us to test the accuracy of the LSMC method for different possible budget allocations. The

budget allocations will, however, be consistent with the previous assumption made in the context of insurance liabilities. For this analysis, we assume the total computational budget will be equivalent to a total budget of $C = 10,000$ inner valuation scenarios. Then we want to test the error associated with LSMC estimated projected value distribution for different scenario allocations, (M, N) such that $M \cdot N = C$. The metrics by which we measure the success of the LSMC estimated projected value distributions will again be $\mathcal{M}_1$ and $\mathcal{M}_3$ introduced in Section 3.4. For each of the different budget allocations the optimal regression model is chosen using the stepwise AIC algorithm, introduced in Section 3.3. Also, 10,000 repeats were performed using different uniformly sampled fitting points in order to construct 95% confidence intervals for the value of these error metrics for each of the different budget allocations we consider. The error metrics will be calculated both for the LSMC with antithetic variates being employed in the inner valuation stage and without. Antithetic variates were introduced in Section 1.4.

In Figure 3.20 the $\mathcal{M}_1$ error for the estimated projected value distribution is plotted for a number of scenario budget allocations. The 95% confidence intervals are too small to be seen on this scale. This plot shows that in implementing the LSMC method, the smallest error is achieved by choosing as small a number of inner scenarios per outer fitting scenario, N, as possible. Thus, the optimal number of inner scenarios (per outer scenario) to use is just one antithetic pair. The level of this error metric gradually increases as the number of inner scenarios, N, is increased. Figure 3.21 shows similar findings for the error metric $\mathcal{M}_2$, again suggesting setting N as a single antithetic pair. Both of these plots also show that employing antithetic variates within the LSMC algorithm can offer increased accuracy in the valuations obtained. As antithetic variates are fairly easy to employ within the method, these tests suggest that one should use this variance reduction technique when calculating an insurance SCR using this approach.

3.10 Conclusion

In this chapter, some of the choices in the way the LSMC method is implemented have been examined. We summarise these briefly: the stepwise AIC algorithm seems to give a reasonably good approach for selecting the regression model and which is robust against over-fitting; sampling the fitting scenarios from the real-world dis-

tribution gives a good fit in the centre of the projected value distribution, but a poor fit in the upper tail. Sampling the points from an alternative scheme, such as Sobol sampling, gives a much better fit in the upper tail with only a slightly inferior fit in the centre of the distribution, compared with real-world sampling; There was no overwhelming evidence to suggest greatly improved accuracy from using orthogonal polynomials in the regression function, although in combination with Sobol sampling a small increase in accuracy was perhaps noticeable; the algorithm should only use one antithetic pair of valuation scenarios and the remaining computational budget should be expended generating outer fitting scenarios for the most efficient implementation of the algorithm.

Obviously these inferences have been made in the context of the fairly simple problem of calculating the projected value of a European option. However, the payoff function in this example is not too dissimilar to the form of payoff functions encountered in valuing projected insurance liabilities. Thus, the findings made throughout this chapter should provide reasonable guidance as to how to set-up the LSMC algorithm in an insurance liability context. With this in mind, we shall use some of the findings made in this chapter as we move on to considering applying the LSMC method to a stylised life insurance product in Chapter 4.

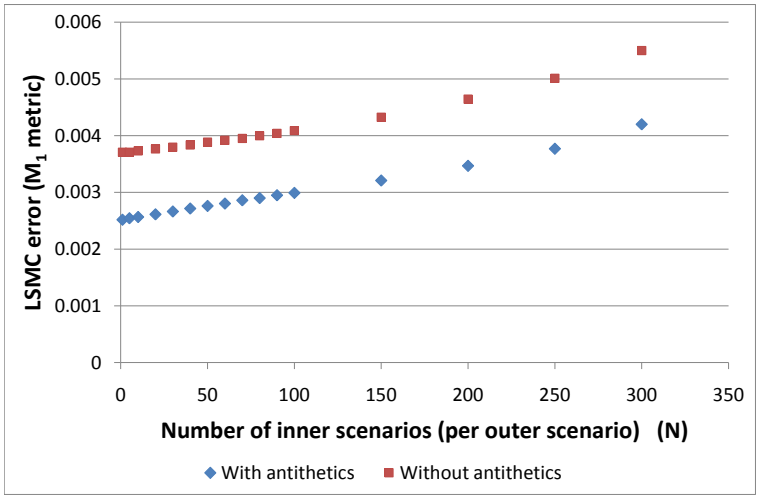

Figure 3.20: Distribution-wide average ($\mathcal{M}_1$) error of LSMC estimates for different pairs of outer-inner scenario budget allocations (indicated with reference to the number of inner scenarios per outer scenario, N). For all points, the total valuation scenario budget is C=10,000. Results with/without antithetics given.

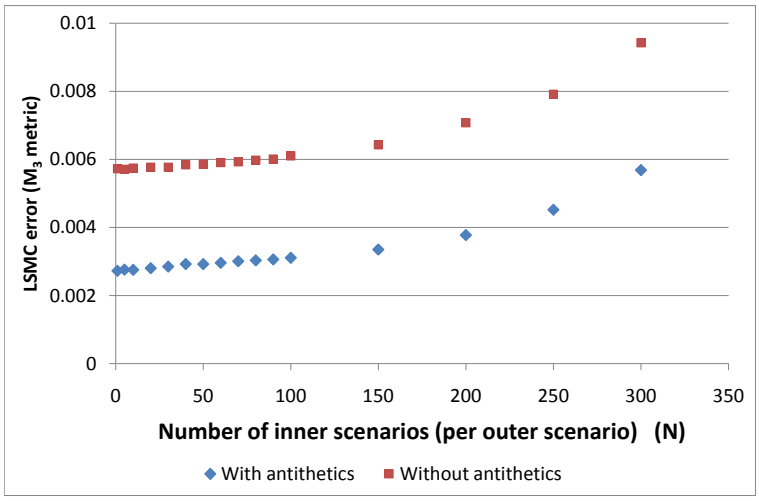

Figure 3.21: Upper-tail region local ($\mathcal{M}_3$) error of LSMC estimates for different pairs of outer-inner scenario budget allocations (indicated with reference to the number of inner scenarios per outer scenario, N). For all points, the total valuation scenario budget is C=10,000. Results with/without antithetics given.

Chapter 4

LSMC insurance case study

Having given an introduction to the application of the LSMC method in the context of the projection of insurance liabilities and calculating an SCR, we will now turn our attention to considering an illustrative example which tests the technique for a realistic insurance product. The LSMC approach will be of use to practitioners for insurance products featuring complex, path dependent and multi-dimensional liabilities. In such instances, no simple analytic formulae for calculating the liabilities are available and even a nested simulation approach with an optimal scenario allocation can be quite inefficient. The LSMC technique offers a method which is both efficient and can offer reliable estimates of the projected liability distribution. This technique was discussed in more detail in Chapters 2 and 3. One class of unit-linked insurance product for which the liabilities are often fairly complex, path dependent and multi-dimensional is the variable annuity, which was introduced in Section 1.3. In this chapter the LSMC technique will be applied to approximate percentiles of the projected liability distribution for a stylised product from this class.

4.1 Variable Annuity (VA) stylised product

The analysis of the LSMC method in an insurance context given later in this chapter will be based on the following variable annuity type product which is introduced in a Faculty of Actuaries Variable Annuity Working Party report [Led10]. The guarantee outlined in Figure 4.1 on the next page would sit as an option on a pension contract with income draw-down. This figure summarises the main features of the variable annuity guarantee and can be understood in relation to the discussion of Section 1.3. A brief overview of these features will now be given to ensure full clarity of this product. Firstly, the contract owner is a 65 year old male and from the onset of the contract has the 'guarantee rider' activated, i.e., this option on the insurance contract initially sits on the pensions contract from the policyholder's 65th birthday. He will pay an extra 1% of the fund value in management fees as a result of this option being activated. If the policyholder wishes to turn off this option, he would

be entitled to do so and this would cancel this extra guarantee charge. We shall refer to this as a customer 'lapse'. The modelling of the rate of policyholder lapse is by no means trivial and is an important issue in the calculation of the product's liabilities. Some possible lapsation models will be discussed in Section 4.2.

This product comes under the guaranteed minimum withdrawal benefit (GMWB) type of the VA product class, as described in Section 1.3. The policyholder is guaranteed to receive income at the level of 5% of the 'guarantee base' each year after his 65th birthday, until he turns off the guarantee 'rider' or dies. This 'guarantee base' is initially set at the amount of the policyholder premium, but can increase in value with an increasing VA fund level for the first ten years after annuitisation. After this window passes, the 'guarantee base' will remain at the same level for the remainder of the product's lifetime. This will be discussed in more detail shortly.

The underlying VA fund, which is initially funded by the policyholder premium, is invested each year as follows: 60% will be invested in an equity index tracker, while the remaining 40% will be invested in a ten year zero coupon swaps fund.

The option is a Guaranteed Minimum Withdrawal Benefit (GMWB), which guarantees the customer that they can receive the minimum level of income for the rest of their life.

- Sample Customer: 65, male
- GMWB level: 5% of Guaranteed Base
- Guarantee Base (GB): Initially 100% of premium.
  Each anniversary before age 75 steps up to the fund value (FV) if this is higher, subject to a maximum increase at any time of 15%
- Asset mix of fund: Equity: 60% (FTSE All Share Total Return tracker)
  Fixed Interest: 40% (Zero coupon swaps fund, duration 10 years)
  Rebalanced once a quarter (without discretion)
- Charges: AMC - 0.75% of fund value,
  Commission - 0.50% of fund value
  Guarantee - 1.00% of Guarantee Base
- Flexibility: Guarantee rider can be "switched off" and the guarantee charge will cease to apply. The customer can then decide to buy an open-market annuity or to remain invested in income drawdown, subject to regulatory restrictions.

Figure 4.1: An outline of the features of the stylised VA guarantee. [VWP07]

4.2 Calculating the stylised product liabilities

In this section a precise description of how the liabilities associated with this stylised product can be calculated will be given. We will then highlight the features of the product which make determining these liabilities a path-dependent calculation. This will hopefully give the reader a clear idea of the structure of the liability cashflows of this product and the variables on which they depend.

In order to ultimately deduce the liability faced by an insurer offering this product on the market, there are a few intermediate results which must be determined. Therefore, for clarity of exposition we shall treat each of these intermediate calculations separately, though still demonstrating how they interlink. It is important to recall that we assume the customer for this product is a 65 year old male, whose prior payments have accumulated to £100,000 and has initiated the guaranteed minimum withdrawal benefit rider feature (as stated in the product specification in the table in Figure 4.1) starting immediately from their 65th birthday.

Policyholder Life Expectancy

Before looking to formalise the features of the product with the intention of calculating its associated liability, one piece of information we shall certainly require is the life expectancy of the annuitant at each stage in their life from the age of 65 years. Naturally, as the product pays a guaranteed income level for the remainder of the policyholder's lifetime, the future liability to the insurer from this contract will depend on the future life expectancy of the policyholder. Determining the future life expectancy of the policyholder involves the use of actuarial tables and performing some calculations using this data. The exact nature of these calculations is not essential in the following analysis which leads to the VA product's liabilities, however, and thus shall not be discussed further. We should note a couple of assumptions. Firstly, a policyholder aged 120 is assumed to have a future life expectancy of zero years, as the process requires backward iteration from some future time. Therefore, in the inner scenario component of the Monte Carlo simulations, we must simulate at least 55 years into the future from the policyholder's 65th birthday. Secondly, mortality rates are assumed to be deterministic. More realistically, one should employ a stochastic mortality model to capture the additional risk of uncertainty within the mortality rates. This is an area for further research.

VA Fund Value

We now must calculate three key values related to the VA product for each of the 55 (possible) future years from annuitisation, which are all interlinked and stochastic. These are the 'fund value', 'guarantee base' and 'income level'.

The fund value is the amount of money in the policyholder's fund, which starts at £100,000, but can grow or diminish from one year to the next depending on the performance of the equity index and the ten year zero-coupon swaps fund (or bond) investment returns. This is information which is obtained through simulation. The fund value level, F_i, is determined, in terms of the fund return, R_i, in year i as

$$F_i = \max\Big(\Big(F_{i-1} - I_{i-1}\Big) \cdot \Big(1 + R_i\Big) \ , \ 0\Big), \tag{4.1}$$

with

$$R_i = \alpha E_i + \beta B_i - \nu - \lambda, \tag{4.2}$$

where i is the number of years after annuitisation, α and β are the percentage of the fund being invested in equity and bond respectively and E_i and B_i give the return from equity and bond in each year. The term I_{i-1} gives the VA income level from the previous year, $i-1$, which is determined as will be discussed shortly. The parameter ν in the fund return calculation represents the general management charge on the VA product, while the λ represents the additional charge payable whilst the GMWB rider is activated. The maximum function simply encapsulates the fact that the fund value can never take values which are less than zero.

VA Guarantee Base

The next property of the VA related product which shall be discussed is the guarantee base. This arises because the income level which the policyholder draws from the fund each year is not simply dependent on the VA fund value. Instead it depends on this guarantee base which is linked to the fund value, but captures the 'ratchet feature' of this product. Initially at annuitisation (on the policyholder's 65th birthday) the guarantee base will be equal to the fund value of £100,000. If the number of years since annuitisation is less than the 'ratchet term' of the product (which for this product is given as ten years), then the guarantee base can 'ratchet up' along with the fund value. Thus, if the fund value increases from £100,000 to £105,000 by the end of the first year, the guarantee base will also rise to £105,000. If however, a similar rise in the fund value occurred 15 years after annuitisation, then the guarantee base would remain the same value as it was in the previous year (as this is outwith the ratchet term of the product). The product also specifies that

the maximum increase in guarantee base from one year to the next is 15%. Thus, if the fund value increases by 20% from one year to the next, the guarantee base will only do so by 15%. The guarantee base also must stay at the same level or increase from that of the previous year. This is consistent with the product having a minimum withdrawal level of 5% of the initial fund value. Mathematically, the guarantee base can be expressed as follows:

If we denote the ratchet term as γ, then the guarantee base i years after annuitisation is given as :

$$G_i = I(i \leq \gamma) \cdot \min\Big(1.15 \times G_{i-1}, \max\Big(G_{i-1}, F_i\Big)\Big) + I(i > \gamma) \cdot G_{i-1}. \tag{4.3}$$

VA Income Level

The VA income level can now be simply deduced for each future year after annuitisation from the guarantee base value in that year. Indeed, this value is just the set withdrawal percentage level (stated in the product to be 5%) multiplied by the guarantee base value at the year considered. Under a typical insurance contract the policyholder will be able to choose how much to withdraw at each valuation date, up to some maximum amount. We assume the policyholder always withdraws the full amount of 5%. However, Milevsky and Salisbury [Mil06] show that, under certain conditions, withdrawing the full amount will be optimal for the policyholder in the sense that it maximises the value of the embedded option. Therefore, in the context of insurance solvency testing, such an assumption regarding policyholder behaviour seems reasonable.

Policyholder Lapse Rate

The lapse rate gives the rate of the policyholder lapse (as a percentage) at each future time. Recall that a policyholder lapse is the term given to the situation where the policyholder decides to switch off the guarantee withdrawal feature. They can then decide to buy an open-market annuity or remain invested in income draw-down. This lapse rate can be modelled under either a static or dynamic approach. The static approach simply sets the policyholder lapse rate at a fixed level. However, this is not an entirely satisfactory approach. Obviously if the fund value is growing

rapidly, but we are outwith the ratchet term, the guarantee base will remain fixed. This situation should make a policyholder more likely to 'turn off' the rider feature of the product ('lapsing' on their policy), as the growth in the underlying fund is not being realised in the income the policyholder receives. With this in mind, it is more appropriate to model the policyholder lapses 'dynamically'. This means this lapse rate will depend on the levels of fund value and guarantee base (which are of course dependent on the simulated underlying economy) and also on whether or not the ratchet term has passed. Furthermore, it will also depend on the present cash roll-up in the underlying economy, i.e., the future value of £1 in one years time. Encapsulating these features should give a much more realistic appreciation of actual policyholder lapse behaviour.

It should be noted that although policyholders will rarely make mathematically optimal decisions when it comes to their policy, they will often realise when the market situation is not advantageous to keeping their VA policy rider feature active, perhaps through financial advice. Thus, a dynamical model for lapse rate seems more realistic than a simple fixed rate static approach. Indeed, a survey carried out by the Society of Actuaries showed that in 2005, 15 out of 18 insurers assumed dynamic lapse behaviour in their liability valuations [Led10]. An area for further research could be to employ a stochastic risk factor to the policyholder lapse model.

The dynamic lapse rate condition that shall be employed in the VA product liability calculation component of our model shall operate as follows:

- A 4% Base lapse rate amongst policyholders shall be assumed to hold unless one of the following three conditions hold:
- The lapse rate will be 1% if the present value of income payments is greater than the current fund value.
- The lapse rate will, naturally, be 0% if the fund value has diminished to zero.
- The lapse rate will reach a maximum of 7% if the current fund value is greater than the current guarantee base *and* we are outwith the ratchet term of the product.

Rate of Policies in Force

The next property of the VA product we shall consider is the rate of policies (still) in force at each future projected year. This basically gives the proportion of policies still in force (in a large completely diversified portfolio of policies) given the current simulated state of the economy. Thus, at annuitisation this takes value 1, but with time this value will diminish as there is the possibility of the policyholder lapsing, as described by the lapse rate model, or policyholder death. If the number of policies in force is 0.75 at some future given time, then, given the simulated economy until this time, this means there is a probability of 0.75 that the policy would still be in force, i.e., the policyholder will not have lapsed or died since annuitisation.

The calculation of this property follows fairly easily from the policyholder lapse rate at any given future year. Naturally, the initial rate of policies in force at annuitisation is 1, as there has not been any opportunity for policyholder lapse. Then, i years after annuitisation, the proportion of policies still in force, P_i, is given by the recursion

$$P_i = P_{i-1} \cdot \left(1 - q^x_{i-1} - X_{i-1}\right), \tag{4.4}$$

where q^x_i gives the probability of policyholder death within the next year, given they are at an age of $65 + i$ and X_i gives the lapse rate in terms of the year, i, after annuitisation. This formula is fairly intuitive, as the guarantee will, of course, remain active on the policy unless the policyholder lapses or dies from one year to the next.

(Discounted) Cost of Guarantees

Having introduced the key values with regard to the policyholder and the stylised VA product (some of which of course depend on the state of the economy, which will be simulated), we are now in a position to calculate the first cashflow associated with the product. This cashflow is the 'cost of the guarantees' to the insurer associated with the issue of this insurance contract. It essentially gives the cost the insurer will have to pay at each of the future years after the year of projection as a result of the embedded minimum withdrawal guarantee in the contract. Naturally, these values will be different for each simulation path.

Of course, the insurer will only face costs in guaranteeing this minimum withdrawal to the policyholder if the amount cannot be met from the fund value at the time due. Also, there is the chance that the policy may have already lapsed, in which case the subsequent information and cashflows of this VA product would be irrelevant from then on. Thus, we must factor in the proportion of policies which remain in force (which is essentially the probability that the guarantee feature still remains active). As a result, the cost of guarantees, V_i, that the insurer faces i years after annuitisation of the VA contract is given by

$$V_i = -P_i \cdot \min(F_i - I_i, 0). \tag{4.5}$$

The factor of -1 is needed as any shortfall in the payment of the required income level from the fund value is a positive cost to the insurer as counter-party to the guarantee.

Naturally, we must discount this value of cost of guarantee by the discount factor at the time the cost is considered, as is always the case for risk-neutral valuation. This means we divide V_i by the amount £1 at annuitisation has grown to i years later.

Present Value of the Cost of (all future) Guarantees

This brings us to the key product cashflow of this chapter – the present value of the cost of all the future guarantees. Before going on to discuss the financial importance of this value, the procedure for its calculation shall be given.

At the time 55 years after annuitisation, the present value of the cost to the insurer of all future years guarantees (PV CoG) is just the cost of guarantee associated with the product for that year, i.e., V_{55} in the notation used above. For all the years up until 55 years after annuitisation, the PV CoG is given as the sum of all the subsequent years cost of guarantees discounted back to the year at which this PV CoG is being calculated. Thus, for example, the PV CoG 53 years after annuitisation would be given by the sum of the cost of guarantee (CoG) for year 55 discounted back to time 53 years after annuitisation plus the CoG for the year 54 years after annuitisation discounted back one year. Recall, the discounting is with reference to the discount factor of the underlying (simulated) economy, and thus will be at different rates for each simulation path.

Mathematically, the PV CoG i years after annuitisation, L_i, is given in terms of the cost of guarantees, V, as:

$$L_i = \sum_{k=i}^{55} V_k \frac{C_i}{C_k}, \tag{4.6}$$

where C_j gives the value at year j of C_0 at annuitisation, due to interest. This will generally be stochastic, unless a deterministic model of interest-rates is assumed.

Economically, this value gives the insurer a measure of the total liability associated with this VA contract at each year i after annuitisation. Thus, if we consider this value one year after annuitisation, this will give the insurer an idea as to the level of capital required to match all the liabilities related to this product from one years time onward. This will be beneficial in allowing the insurer to plan their finances and strategy for the coming year, safer in the knowledge that they have a good measure of required capital levels to remain solvent. For the remainder of this chapter this is what we will denote a 'liability' associated with this stylised product. Thus, L_1, will be referred to as the liability of the product at year one.

An alternative definition for the liability arising from this product would be to consider the 'net costs'. This is simply the difference between the PV CoG and the present value of the sum of all the future annual guarantee charges the insurer will receive from the policyholder. For a SCR calculation, it perhaps makes more sense to define the liability in this way, since the liabilities incurred each year can be 'offset' against the charges the insurer takes in. However, for the purposes of this chapter, we will just consider the 'liability side' of the cashflows and use our definition of liability, L, given above.

(Present Value of the) Guarantee Charges and VA Net Costs

For completeness, we present a mathematical definition of the present value of the future guarantee charges. This is (at each projected year) the discounted value of all future guarantee charges the policyholder will pay to the insurer. Naturally, this will be multiplied by the factor giving the proportion of policies still in force, for the same reasons as discussed in the calculation of the cost of guarantees. The guarantee charge the insurer receives, say W, in year i after annuitisation is then given by

$$W_i = F_i \cdot P_i \cdot \lambda, \tag{4.7}$$

where we recall the guarantee charge, λ, is expressed as a percentage of the fund value level. The present value of all future guarantee charges can be given for each projected year in a manner similar to the calculation of the present value of the cost of all future guarantees. Firstly, for the i-th projected year after annuitisation, we obtain the discounted guarantee charge for that year, Z_i, as

$$Z_i = \frac{W_i}{C_i}, \tag{4.8}$$

where C_i gives the value at year i of £1 at annuitisation due to interest. Then the present value of all the future charges at any given year is just the sum of all the subsequent remaining year's discounted guarantee charges (up until after 55 years after annuitisation, where we assume the policy will no longer be in force).

The Path-Dependent Nature of Insurance Liability Calculation

Having demonstrated the process through which the liability on this stylised product is determined, it is clear that this is a reasonably complex, path-dependent calculation. Finding a closed-form analytical formula for the liability will not be possible. The insurance liability associated with this product depends on the cost of the guarantee for each year the product can still be in force (which we assume is a maximum of 55). Therefore, we must simulate out this full term to get the interest rates and equity and bond returns and calculate the cost of the guarantee for each year after the projection date. This illustrates why the calculation of such a product's insurance liability is a problem which requires simulation and why projecting for such liabilities will result in a nested simulation framework.

4.3 Test of LSMC method: Black-Scholes-CIR model

Having given a full introduction to the stylised variable annuity product, we will now test the LSMC method in calculating the projected liabilities on such an insurance product. For the first of these tests we will assume that the equity index returns are lognormal as in the Black-Scholes model, however we will assume that the risk-free rate is stochastic and follows a CIR model. We will denote this the Black-Scholes-CIR (BS-CIR) model. The stochastic process governing the CIR process for r is the same as in the Heston-CIR model described in Section 8.3. The parameters of the

CIR process in the real-world component of the LSMC method are set as follows: the mean reversion rate $\kappa_r = 0.3$, the mean reversion level $\theta_r = 0.05$ and the volatility of the process $\sigma_r = 0.1$. The correlation between the Brownian motions driving the equity index and risk-free rate processes is set as -0.3 and the initial value of the short-rate $r_0 = 0.04$. In the risk-neutral scenarios, the CIR process has parameters set as $\kappa_r = 0.4$, $\theta_r = 0.04$ and $\sigma_r = 0.1$, with the correlation equal to the level in the outer scenarios. Naturally, the value of risk-free rate r at the last timestep of each of the outer scenarios will be taken as the initial value of the process for r in the corresponding inner scenario stages. The volatility in the lognormal stochastic process for the equity returns is 20% and μ, the real-world growth of the equity asset, is 0.03. In these tests we consider a projection date $t = 1$ year on from annuitisation.

The first issue in applying the LSMC method in this example is to determine what the explanatory variables which influence the liability valuation at the projection date are. Naturally, the fund value at projection will be one such variable, as it is from this fund that the policyholder withdrawals are subtracted. As the projection date is at year one, the first withdrawal date, the guarantee base will not be an explanatory variable here. The guarantee base could be used as an explanatory variable if a projection horizon of greater than one year was considered. It would essentially capture the path-dependency of the fund value from annuitisation until the projection year. In the current example there is no such path-dependency to concern us, since G_1 is given by a deterministic function of F_1. We will, however, consider a 5-year projection for this product in Section 4.4. The only other explanatory variable influencing the liability at the projection is then the value of the short-term risk-free rate of return, r_1. The liabilities are obviously dependent on the levels of interest rates throughout the term of the product, because the calculated cost of the guarantee at each future year is discounted back to year one and also the level of interest rates is one of the inputs into the dynamic policyholder lapse function. With a CIR model, the expected future levels of interest rates explicitly depend on the current value of the short-rate, hence r_1 is the only explanatory variable needed in the regression to fully capture the dependence on the future levels of interest rate.

Having determined the influential explanatory variables in estimating the liability L_1, the next task is to find the form of the regression model involving these variables.

Following the findings of Chapter 3, we will not generate the outer scenarios used to calibrate the regression model in the LSMC method using the real-world Black-Scholes-CIR model, however. It was argued there that sampling these scenarios using Sobol quasi-random sampling would yield much more accurate projected liability estimates in the upper tail of the distribution. Thus, we shall adopt this approach in the implementation of the technique here. Looking at a large number of real-world paths out to the projection date at year one, it was concluded that it was extremely unlikely that the fund value, F_1, would be out-with the interval $[70,000, 160,000]$ and the risk-free rate at $t = 1$ out-with $[0, 0.12]$. Thus, pairs of points were generated inside the hypercube spanned by these two intervals using Sobol sampling and these pairs were taken as the outer fitting scenarios for the following LSMC tests. Also, following the results of Chapter 3, we employ just one pair of antithetic valuation scenarios for each of the 10,000 outer fitting scenarios generated. This follows from the argument given in that chapter that the computational effort should be employed generating as large a number of outer scenarios as possible, since the subsequent regression will greatly improve the accuracy of each of the inner scenarios. These fitting scenarios generated will then be used in the cashflow calculations outlined in Section 4.1. Ultimately, this will lead to a valuation of the liability at the projection date for each of the inner scenarios simulated. These valuations will be the response variables in the regression component of the LSMC algorithm.

The next stage is to determine the form of the regression function of the two explanatory variables. As in Chapter 3 the method we will use for this will be the stepwise AIC algorithm. This seemed to be reasonably successful in selecting a regression function for the projected put valuation problem discussed in that chapter, however there are other approaches which could be used here. Bauer, Bergmann and Reuss [Bau10] advocate using an model selection algorithm based on a generalised Mallows' C_p statistic. This will be discussed in more detail in Section 4.6. Implementing the stepwise AIC algorithm in this example, yields the following form of regression model for the projected liability at year one:

$$\begin{aligned}\hat{L} \quad &= c_0 + c_1 F_1 + c_2 r_1 + c_3 F_1^2 + c_4 r_1^2 + c_5 F_1 r_1 + c_6 F_1^3 + c_7 F_1^2 r_1 + c_8 F_1 r_1^2 \\ &\quad + c_9 F_1^4 + c_{10} F_1^3 r_1 + c_{11} F_1^5 + c_{12} F_1^4 r_1 + c_{13} F_1^6 + c_{14} F_1^5 r_1. \end{aligned} \tag{4.9}$$

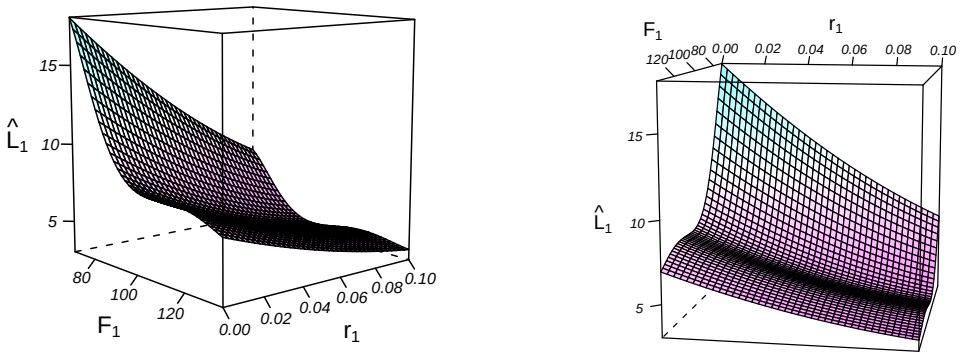

Figure 4.2: Two views of the regression surface of $\hat{L}_1$ varying with F_1 and r_1 for BS-CIR example. The units of $\hat{L}_1$ and F_1 in the plots given above is £1000.

A plot of this regression function in terms of the two explanatory variables, F_1 and r_1, is given in Figure 4.2. This plot suggests that, generally, for lower levels of year one fund value, the LSMC liability estimates are larger. This behaviour is expected for this type of guarantee. If the fund value has fallen significantly by the end of the first year after annuitisation, then it is more likely that the fund value will be unable to meet the guaranteed income level at some stage early in the product's lifetime. When this situation occurs, the insurer is then liable to meet this and all subsequent annual incomes to the annuitant out of its own reserves. Therefore, low levels of fund value at the projection date are likely to give a larger liability to the insurer from selling such a product on the market.

The plot also shows that the spot-rate at projection, r_1, is an equally crucial factor in determining the regression estimate of the VA liabilities. The regression-estimated liabilities increase with decreasing short-rate, which is, again, expected for this type of guarantee. The short-rate gives us an indication of the future expected levels of interest rate. If future interest rates are large, then any future incomes to the policyholder which must be met out of the insurer's own reserves will be discounted back to the projection year by a greater amount. This means that large levels of future interest rates will give smaller liability valuations at the projection date. Furthermore, if interest rates are low, then policyholder lapse is less likely, as a result of the dynamic lapse function being employed. This results in the guarantee remaining active on the product and a policyholder continuing to draw income from

the fund value. This is also consistent with lower future interest rates yielding a larger liability valuation at the projection date. Thus, overall, this surface makes sense in comparison with our intuition on the behaviour of this stylised product.

The dependency of the regression function on the two explanatory variables can be examined further by considering a surface contour plot. In Figure 4.3 the surface contour plot is given for $\hat{L}$ in terms of F_1 and r_1. Also shown here is the regression surface with the contour plot projected onto the upper surface of the cube which borders the surface plot. This helps visualise the connection between the regression surface and the two-dimensional contour plot. The contour plot further illustrates the dependency of the regression surface on the two explanatory variables, as described above. On closer inspection, the contour plot also suggests that for fixed short-rates, a fund value level of around £115,000 at year one will lead to a larger liability estimate than a fund value level of around £100,000. This local behaviour is contrary to the discussion of the overall trend of the surface just given. Is this simply down to a poorly fitting regression surface and some evidence of over-fitting occurring, or is this behaviour actually consistent with the features of the stylised variable annuity product? Well, if the fund value grows to £115,000 by year one, then the guarantee base will also grow by the same amount. Then, for the remainder of the product lifetime, the policyholder will withdraw at least $£115,000 \times 5\% = £5,750$ each year whilst the guarantee feature is active. On the other hand, if the fund value remains around £100,000 at year one, the guarantee base will remain at its initial level and the policyholder will only be guaranteed an annual withdrawal of at least £5,000 for the remaining lifetime of the product.

Therefore, perhaps this resultant increase in the minimum level of income the policyholder will withdraw from the fund in each of the subsequent years after year one will be of greater importance in the calculation of L_1 than the fact that the fund value level, F_1, is around £15,000 larger. Obviously, if the fund value grows by a huge amount over the first year, we would still expect a very small liability to the insurer since the fund would likely cover all the future annual incomes to the policyholder. And similarly, if the fund value level falls drastically by year one, we would still expect this to lead to a large liability. But locally, for values of F_1 of around £100,000 to £120,000, there is no reason why a larger fund value level could not lead to a *larger* expected liability.

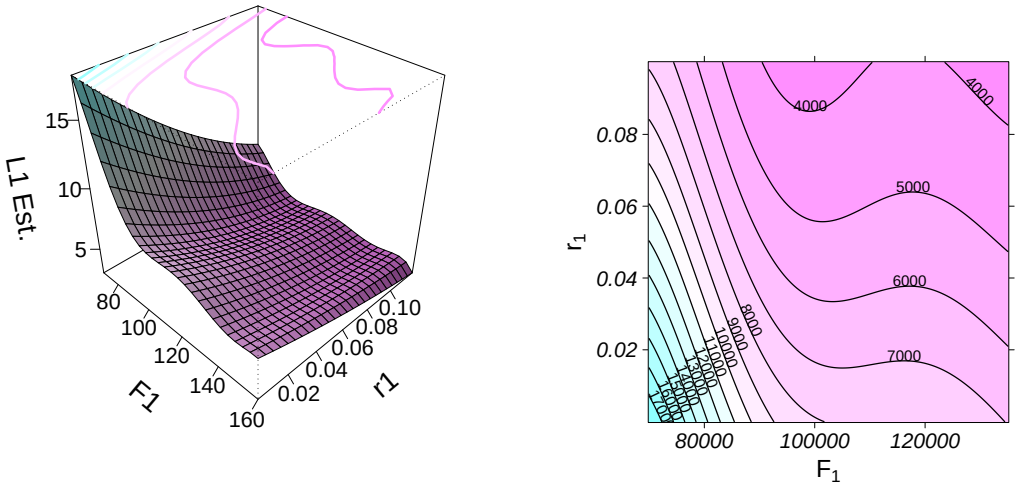

Figure 4.3: Left: BS-CIR example regression surface of $\hat{L}_1$ varying with F_1 and r_1 with contour lines projected onto roof of surrounding box. Right: Regression surface contour plot. The units of $\hat{L}_1$ and F_1 in the plots given above is £1000.

Of course, it could still just be the result of a poorly fitting regression surface, so some tests need to be conducted to see which of these explanations holds. These tests will be explained shortly and the results made clear. For now, we turn our attention to investigating how successfully the LSMC regression function approximates different percentiles of the true projected liability value distribution. Naturally, we will pay closer attention to the fit of the regression function to the upper percentiles of this distribution as this is the region which is of greatest concern to an insurer. In these tests we would ideally like to select a few key percentiles of the true liability distribution, conditional on some simulated real-world outer scenarios and compare how well the LSMC estimates for these outer scenarios match these true liabilities. But there is no analytical formula for L_1 in terms of F_1 and r_1. Also, attempting to obtain accurate estimates via simulation is computationally infeasible – this is why we need to use approximation techniques such as the LSMC method to estimate the projected liability distribution. Therefore, in these tests we will investigate how accurately the LSMC regression function approximates the projected liability value for percentiles of the regression-estimated liability distribution. This is the distribution which is given by inputting the $10,000$ real-world test scenarios from the Black-Scholes-CIR model into the regression function found by the LSMC method. Such an approach will give a reasonable test of the fit to the true percentiles if the regression-estimated liability distribution has percentiles which are reasonably

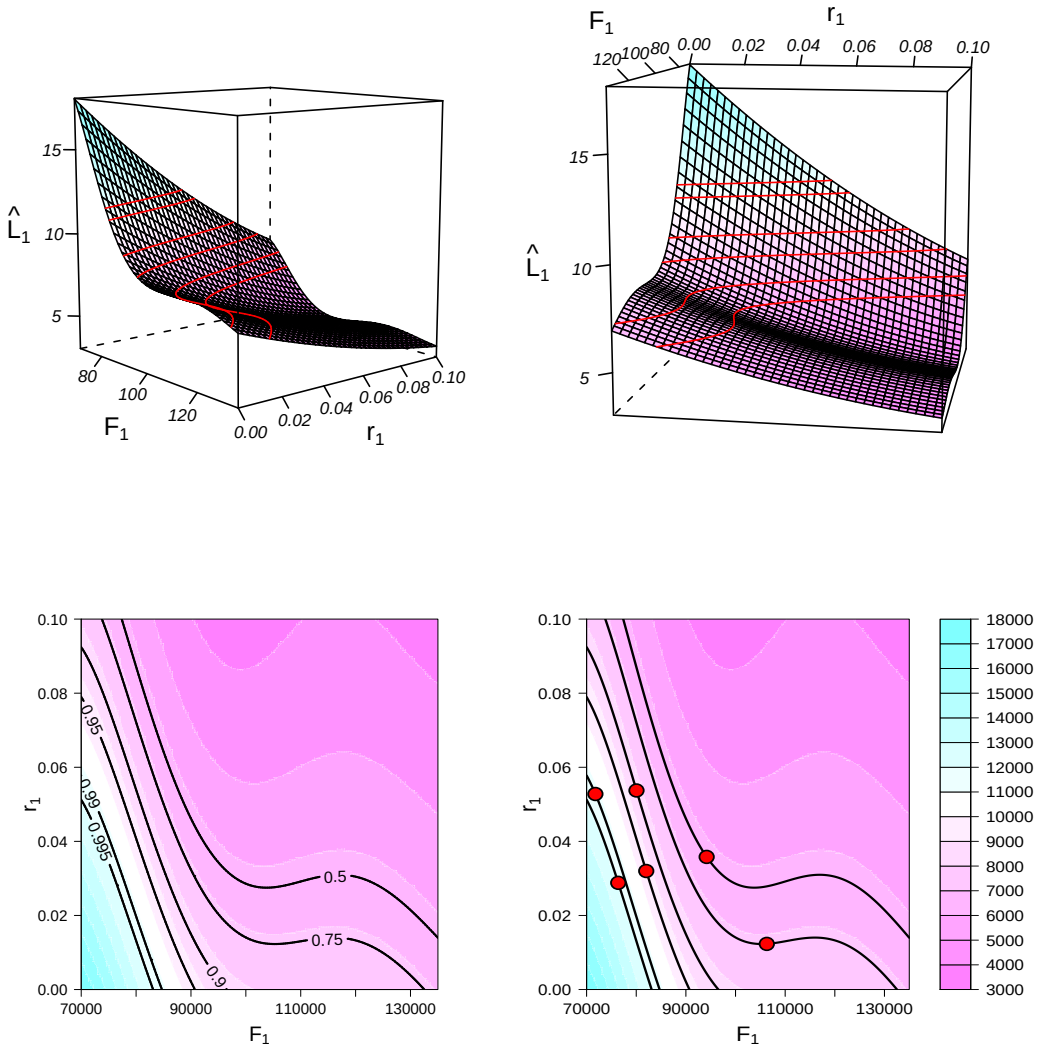

Figure 4.4: Top row: two views of the regression surface with regression-estimated distribution contour lines superimposed (red). Bottom row: surface contour plot with these contour lines superimposed. Red circles in bottom-right plot show the values of F_1 and r_1 from the outer scen. which corresponds to exact percentiles.

close to those of the true projected liability distribution. This seems a reasonable assumption to make for the purposes of testing the LSMC technique. The outer scenarios corresponding to the key percentiles of the regression-estimated projected liability distribution will then be taken and a large number of inner scenarios will be simulated to obtain an accurate valuation of this percentile. These 'full simulation' valuations can then be compared to the regression estimates as a test of the accuracy of the LSMC method.

The percentiles which we shall consider in these tests are the 50-th, 75-th, 90-th, 95-th, 99-th and 99.5-th. This will give a thorough examination of the upper percentile of the projected liability distribution, but also check the fit around the median

of the distribution and at the 'one-in-four year' worst case scenario. The contour lines at the values of the percentiles of the LSMC approximated distribution are shown superimposed on top of the surface plot of the regression function in the top line of Figure 4.4. These are just the surface plots that were shown in Figure 4.2 with the six LSMC estimated projected liability distribution percentiles shown on the surface as red lines. In the bottom line of Figure 4.4 the surface contour plots are shown with the percentile contours superimposed as black lines. In the bottom-right hand corner plot the red circles show the exact levels of fund value and short-rate corresponding to each of the six percentiles which we will be testing against. The results of the tests for each of the percentiles that we consider are given in Table 4.2 on page 109. The results show that the LSMC method seems to give a reasonably accurate approximation of the projected liabilities in the upper tail of this distribution. Indeed, the regression estimates are fairly close to the full simulation estimates for the 90-th, 95-th, 99-th and 99.5-th percentiles. These full simulation estimates were given by taking the outer scenario corresponding to each of these percentiles and performing 20,000 valuation scenarios. This resulted in liability valuations with standard errors of around £50, as given in the table. The LSMC estimates of the projected liability for the median and 75-th percentile of the distribution also seem to be reasonably accurate in comparison with the full simulation valuations performed for the corresponding real-world outer scenarios. It should be noted that we only consider the single scenario which corresponds to the percentile of the projected liability distribution. As discussed in Chapter 3 we should really consider an average of the accuracy of fit for a number of scenarios around each percentile, because many different possible pairs of F_1 and r_1 can give a liability around a given percentile. Therefore, some other scenarios around each percentile were also tested against the LSMC estimates and these gave reasonably similar levels of accuracy of fit to the results in Table 4.2. Overall, these tests provide evidence that the LSMC method gives a good approximation in both the upper tail and around the centre of the projected liability distribution.

The other issue which needs to be investigated is whether the local increase in expected liability for increasing year one fund values (in the vicinity of year one fund values of around £100,000 to £120,000) is an artifact of a poor fitting regression function from the LSMC method. Or is this actually representative of how the

liabilities of this stylised VA product behave? In order to test which of these explanations is correct, four pairs of scenarios around this region were taken and full simulations of 20,000 valuations were performed to obtain accurate estimates of the liability in each case. The results of the full simulation valuations for these four pairs of scenarios are given in Table 4.1 at the bottom of this page.

If we consider outer scenario numbers 8,524 and 8,126 we will now explain what these tests show. The fund value at year one for real-world outer scenario number 8,524 is £100,778, whilst for outer scenario 8,126 the fund value is £117,120. Given our intuition about the overall trend of this product we would expect that (if both these scenarios had the same value of short-rate, r_1), the projected liability at year one would be larger for simulation path number 8,524. Furthermore, since the short rate for path number 8,524 is less than the short rate for path number 8,126, this should increase the projected liability for path number 8,524 relative to the liability for path number 8,126. However, the LSMC regression function estimated that the liability for path 8,524 is £5,176, which is *smaller* than the LSMC estimate for path number 8,126 of £5,426. The full valuation for outer scenario number 8,524 is £5,176 with standard error of £32, whilst the full valuation of scenario 8,126 is £5,788 with a standard error of £36. Therefore, the full simulation estimates for these paths agree with the order of the LSMC regression estimates for these scenarios. The other three pairs of scenarios in Table 4.1 also confirm that this local anomaly in the overall trend of the distribution is a result of the complexity of the ratchet feature. The ability of the LSMC method to capture this rich structure is further evidence that it can accurately and robustly model complex insurance liabilities.

Sim. Num.	F_1	r_1	Reg. Est.	Full Sim.	St. Err.
8524	100,778	5.15	5,176	5,182	32
8126	117,120	5.18	5,426	5,788	36
7820	99,366	5.11	5,212	5,207	32
94	119,922	5.16	5,423	5,610	35
44	102,344	5.44	5,055	5,168	32
9918	117,697	5.88	5,174	5,458	34
3020	99,685	5.09	5,215	5,207	32
8318	115,720	5.17	5,423	5,879	36

Table 4.1: Full sim. liability valuations for scenarios chosen to investigate the local behaviour of the projected liability distribution for F_1 around £100,000-120,000.

Perc.	F_1	r_1	Reg. Est.	Full Sim.	St. Err.
50-th	94,186	3.58	6,304	6,467	37
75-th	106,300	1.23	7,155	7,335	42
90-th	80,049	5.37	8,385	8,431	42
95-th	82,015	3.20	9,406	9,469	45
99-th	71,771	5.28	11,268	11,251	47
99.5-th	76,372	2.88	11,908	11,990	50

Table 4.2: Full sim. liability valuations for the real-world outer scenarios which correspond to some key percentiles of the LSMC estimated projected liability distribution. Given for each of these outer trials are the LSMC corrected estimate and also the full sim. valuation together with standard error. All the values in the table are quoted in units of pounds, except for r_1 which is given as a percentage.

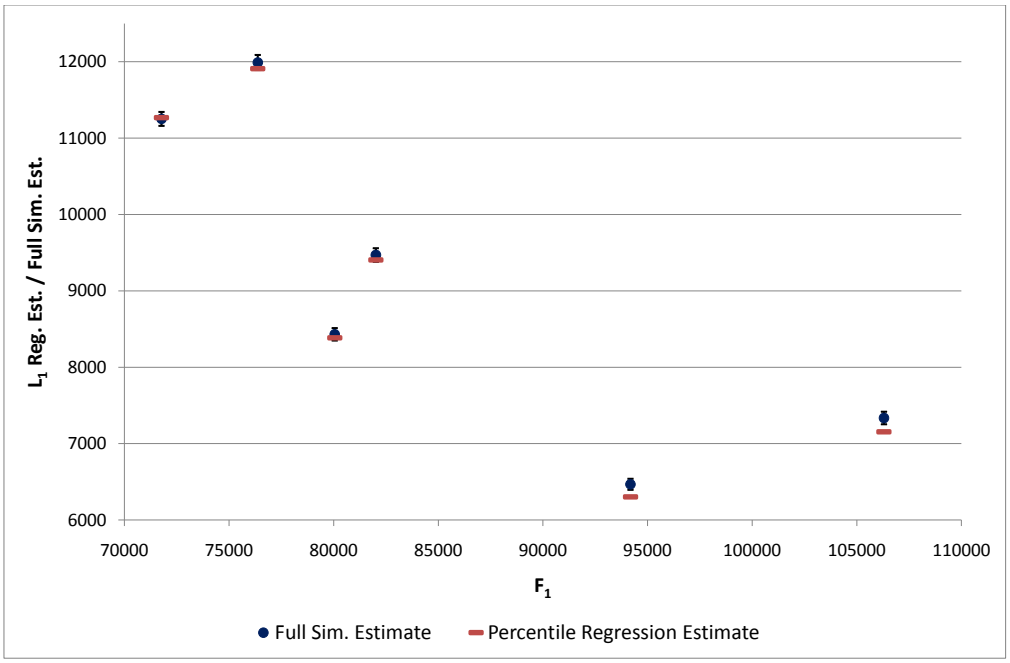

Figure 4.5: Plot of full sim. liability valuations for the real-world outer scen's which correspond to some key percentiles of the LSMC estimated projected liability distribution. Given for each percentile are the LSMC corrected estimate (red horizontal bar) and the full sim. valuation. The diameter of the small circles of the full sim. estimates are approximately representative of the 95% confidence interval of these full valuations. All values in the figure are quoted in pounds.

The LSMC estimates for some of these scenarios in Table 4.1, perhaps do not appear as accurate as those scenarios in the upper tail of the projected value distribution that were given in Table 4.2. This finding is consistent with some further exploratory tests involving full simulation valuations of other scenarios. On the other hand, the accuracy of the LSMC estimates around the centre of the projected liability distribution appeared to improve when the real-world Black-Scholes-CIR model was employed to generate the fitting scenarios, compared with the results from using Sobol sampling given here. The accuracy of the fit in the upper tail seemed to reduce with this form of fitting scenario sampling, however. These findings are the result of exploratory analysis, but they are consistent with the discussion and results of Section 3.7. The set-up which gives the best fit in the upper tail of the distribution will be of most interest to insurers, therefore the results from using Sobol fitting scenarios have been quoted here.

4.4 Test of LSMC method: Five-year projection

Rather than considering a one year projection of insurance liabilities, some practitioners might want to get an understanding of the longer term capital adequacy of their business and look at projections over a longer time-horizon. Therefore, in this section we will consider exactly the same stylised VA product and real-world financial model from the previous section, except now we will look at estimating the projected liability on such a product at a projection time of $t = 5$ years. This projected liability value is just the value of L_5 from the discussion of the cashflows on this VA product given in Section 4.2. In applying the LSMC method to find the projected liability at year five, however, the regression function will now consist of three explanatory variables, as will now be explained.

In the case of a one year projection, as considered in Section 4.3, there is no path dependency in the calculation of the liability. The first policyholder withdrawal occurs at year one, hence we do not care what happened to the equity index and bond returns in the period from annuitisation until then. For a five year projection, however, we cannot just only consider the fund value and short rate at year five as the only influential explanatory variables. If the fund value were to rise to a high level in the first couple of years after annuitisation and then fall back to £100,000 at year five, this would give a larger expected liability compared with if the fund value

were to fall in the first couple of years and then rise back up to £100,000. This is due to the product's ratchet feature, which controls the level of the floor of the future annual withdrawal amount made by the policyholder. Thus, we need to take into account this path dependency in the regression function for the liability projected at year five. It is not necessary to include the fund value at the end of all of years one to five, however, since it is only through this ratchet feature that the path dependency comes into the liability calculation. This ratcheting is completely controlled by the guarantee base, thus the only explanatory variable needed in addition to F_5 and r_5 for the five year liability projection is G_5, the guarantee base level at the projection date.

Therefore, in considering a five year liability projection, the regression function for the projected liability will now be three-dimensional. This will provide a more realistic test of the LSMC method in approximating projected insurance liabilities, since those found in practice can often depend on many risk-factors. To begin with, we will, again, consider generating the fitting scenarios using Sobol quasi-random sampling, following the results of Section 3.7. This will involve determining the intervals for which it is 'extremely unlikely' that F_5, r_5 and G_5 will be outside under the real-world we assume for the underlying economy. Then, fitting points will be sampled within the hypercube spanned by these intervals according to the Sobol algorithm. Sampling the fitting points in this manner brings about some additional complications, as will be discussed shortly.

In Figure 4.6 pairwise plots of the year five fund value, short rate and guarantee base levels are given for a sample of 10,000 real-world scenarios. From these plots we can see that such a standard quasi-random sampling approach will probably suffice for the fitting scenarios for the explanatory variables F_5 and r_5. However, the guarantee base at the projection date, G_5, is clearly highly correlated with the fund value F_5, which makes sense intuitively considering the structure of the product. The bottom-left hand side plot of Figure 4.6 shows that it would not be wise to sample the fitting scenarios, such that a significant number consisted of values of F_5 above £120,000 and values of G_5 below around £120,000. Similarly, there are very few scenarios for which F_5 is below around £100,000 and G_5 is above £150,000. On the other hand, it is often these extreme values of the risk-drivers which yield the key upper percentiles of the projected liability distribution. Therefore, a balance

must be drawn between sampling enough scenarios around the extreme values of the risk-drivers to obtain an accurate fit in the upper tail of the projected liability distribution, but not too many to make the method very poor at estimating the majority of the distribution.

Rather than simply sampling the outer fitting scenarios evenly across the given three-dimensional hypercube according to the Sobol algorithm, we will now try to adapt this idea to take into account the strong correlation between the fund value and guarantee base levels at the projection date. An adapted fitting scenario sampling approach will proceed as follows: firstly, standard Sobol sampling will be used to generate points evenly over the unit cube. Let the different dimensions of these Sobol samples be labeled $\mathbf{u}_1^s$, $\mathbf{u}_2^s$ and $\mathbf{u}_3^s$. The next step is to give the Sobol samples $\mathbf{u}_1^s$ and $\mathbf{u}_2^s$ a correlation of $\rho_{\text{F-G}}$. This is achieved by treating these Sobol samples as uniformly distributed random variables. We then convert these to samples from an equivalent 'normal distribution' using the inverse cumulative distribution function, $\Phi^{-1}(x)$, as $\mathbf{z}_1^s = \Phi^{-1}(\mathbf{u}_1^s)$ and $\mathbf{z}_2^s = \Phi^{-1}(\mathbf{u}_2^s)$. These 'equivalent normal samples' can then be correlated using a Cholesky decomposition. Setting

$$\mathbf{z}_1^{s\star} = \mathbf{z}_1^s \tag{4.10}$$

$$\mathbf{z}_2^{s\star} = \rho_{\text{F-G}}\,\mathbf{z}_1^s + \sqrt{1-\rho_{\text{F-G}}^2}\,\mathbf{z}_2^s \tag{4.11}$$

will give correlated 'equivalent normal samples' of the first and second dimensions of the Sobol sample. By employing the cumulative normal distribution function, these can be converted back to a Sobol sample which is 'uniform', but with correlated first and second dimensions: $\mathbf{u}_1^{s\star} = \Phi(\mathbf{z}_1^{s\star})$ and $\mathbf{u}_2^{s\star} = \Phi(\mathbf{z}_2^{s\star})$. The third dimension of the Sobol sample is unchanged since being initially generated, thus we just set $\mathbf{u}_3^{s\star} = \mathbf{u}_3^s$. Then $\mathbf{u}_1^{s\star}$, $\mathbf{u}_2^{s\star}$ and $\mathbf{u}_3^{s\star}$ gives a three-dimensional Sobol sample over the unit cube, such that the first and second dimensions have a correlation which is close to $\rho_{\text{F-G}}$. These Sobol sampled uniform-like vectors are then scaled and shifted to define a cube in risk-driver space for which it is 'extremely unlikely' that the values of these three influential variables will lie outside. For F_5 we set this interval as [30,000,200,000], for r_5 as [0,20%] and for G_5 it is defined to be [100,000,200,000]. From the sample of 10,000 real-world scenarios, the correlation between the values of F_5 and G_5 was very close to 0.8, thus this is the value that was chosen for $\rho_{\text{F-G}}$.

The plot in the bottom-right hand corner of Figure 4.6 shows the 10,000 Sobol fitting points for F_5, plotted against the 10,000 Sobol fitting points for G_5. This plot clearly shows the effect of the correlation which has been built into this Sobol sampling algorithm. It certainly appears to be more consistent with the realisation of the real-world model for F_1 and G_1 plotted in the bottom-left corner of Figure 4.6, in comparison with a standard Sobol sampling approach (which would just place points evenly over the whole of this square region of risk-driver space). This approach can be generalised to obtain Sobol sampled fitting scenarios for a general number of risk drivers, such that each pair of risk-drivers have a specific correlation. Further discussion of the Cholesky decomposition technique is given in Section 8.3. Correlating the Sobol sampled fitting scenarios is only likely to give an improved fit in the LSMC regression function for cases where the correlation between a pair of risk-drivers is fairly large in magnitude. Thus, in Section 4.3 we just used standard Sobol sampling, since the correlation between F_1 and r_1 was only around -0.3.

Employing the correlated Sobol sampling approach for F_5 and G_5 will still give fairly poor LSMC estimates for the five year projection example, however: If the fund level diminishes over the first five years of the product (which is likely to happen due to the policyholder withdrawals), the guarantee base will remain at, or be only slightly above, £100,000. As such, nearly half of the 10,000 real-world BS-CIR samples generated were such that $G_5 = £100,000$ and over 75% of these scenarios were such that $G_5 < £115,000$. Clearly, the correlated Sobol sampling, shown in the bottom-right plot of Figure 4.6 is actually not very representative of the real-world samples F_5 and G_5. An even distribution of points across the whole range of likely values of G_5 is not what is sought here. Rather, a far larger proportion of fitting scenarios with $\widetilde{G}_5 < £115,000$ is required to realistically capture the behaviour of the projected liability distribution. We would still, however, like to have some scenarios such that $\widetilde{G}_5$ takes an extreme value and also still be able to capture the correlation between F_5 and G_5. But adapting Sobol sampling to achieve this would be challenging and any strategy developed may be too specific to this stylised VA product and not generalise to other forms of insurance liability. The natural approach to capture this complex structure is just to use the real-world model for the fitting scenarios, although in Section 3.7 it was argued that this approach is likely to give a poor fit in the upper tail. In the tests which follow, the LSMC method will still be found

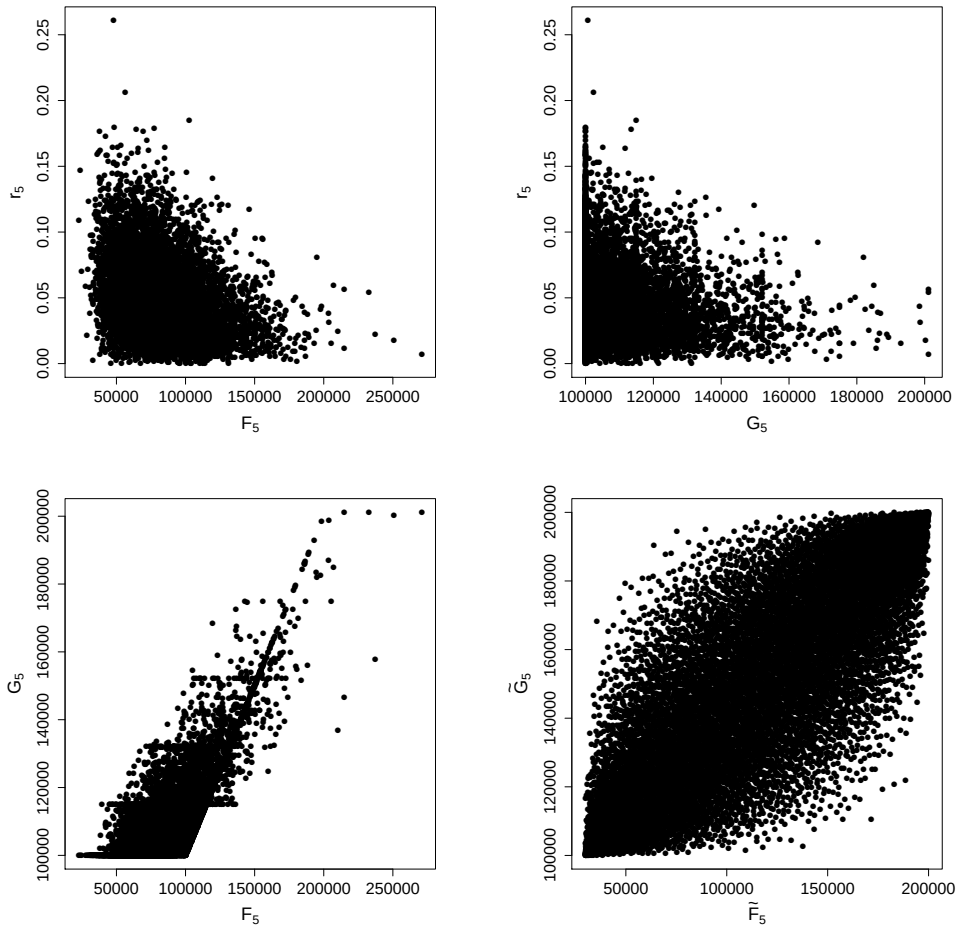

Figure 4.6: Top row and bottom-left: pair-wise plots of F_5, G_5 and r_5 from 10,000 sim's of the real-world Black-Scholes-CIR model. Bottom-right: plot of 10,000 fitting scenarios for F_5 and G_5 under the correlated Sobol sampling approach.

to perform fairly well in estimating this region of the projected liability distribution when the fitting scenarios are sampled from the real-world distribution, however. Furthermore, in Section 4.6 we will outline some lines of further research which could improve the fit in the upper tail of the projected liability distribution whilst still employing real-world sampling of the fitting scenarios.

Having discussed the issue of fitting point sampling, let us now study the form of the regression model found by the LSMC method. As in Section 4.3 10,000 outer fitting scenarios were employed and one antithetic valuation pair was simulated per outer fitting scenario. This choice for the outer and inner scenario allocation follows the results of Section 3.9. The form of the regression model for the LSMC was then chosen by the stepwise AIC approach, as was the case in Section 4.3. This gave a estimated projected liability model which included the following terms, in addition to a constant: F_5, r_5, G_5, F_5^2, r_5^2, G_5^2, $F_5 r_5$, $F_5 G_5$, $r_5 G_5$, F_5^3, r_5^3, $F_5^2 r_5$, $r_5^2 G_5$, $F_5 r_5 G_5$,

$F_5 r_5^3$, $r_5 F_5^3$, $G_5 F_5^3$, $G_5^2 F_5^3$ and $F_5^3 r_5^3$. Since, $\hat{L}_5$ is given in terms of three variables now, we cannot make the same three-dimensional plot of the regression surface as was given in the last section. Similar plots can be made if we fix the value of one of the three risk-drivers, however. Such a surface can be thought of as a three-dimensional slice of the four-dimensional regression surface for $\hat{L}_5$.

Firstly, then, let us consider fixing the value of r_5 and look at how the LSMC estimated liability depends marginally on F_5 and G_5. In Figure 4.7, the regression surface is plotted in terms of F_5 and G_5 for fixed levels of r_5 of 1%, 4.3% and 10%. These values of r_5 correspond approximately to the 5-th, 50-th and 95-th percentiles of the marginal real-world distribution for r_5, respectively. In the left-hand and central plots, we can see that the LSMC estimated liability increases with a decreasing fund value level and an increasing guarantee base level at year five. This is consistent with our intuition about the behaviour of this stylised VA product: If the fund level has fallen to a low level by year five, then it is more likely that at some point in the future this underlying VA fund will no longer be able to meet one or more policyholder withdrawals. This results in the insurer having to cover the shortfall – what we term a liability to the insurer from this product.

Similarly, if the guarantee base level is high at year five, then all the subsequent year's withdrawals made by the policyholder will be at a relatively large level. If the size of the withdrawals from the VA fund are large, then this fund is much more likely to be unable to meet these at some point early on in the lifetime of the product. This situation would represent a large liability to the insurance company selling such a product.

In the right-hand plot of Figure 4.7, it appears that for a large fixed value of r_5 and relatively large values of F_5, the liability will *fall* with an increasing level of G_5, however. Such behaviour is not consistent with the structure of the stylised VA product liability. This anomaly in the regression surface can be easily explained: If the fund value at year five is above around £115,000 then the guarantee base at year five must also be above this level, by definition of the ratchet feature of the product. As the fund value increases further, the only values of guarantee base which are consistent with this will also be large. Thus, it does not make sense to consider combinations of relatively large F_5 with relatively low G_5, as it is impossible for such scenarios to occur. That is why the regression surface exhibits this spurious

behaviour for large values of F_5 in the right-hand plot. Of course, this is of no consequence to the success of the LSMC method, since no such scenarios are going to be produced by the real-world model in practice.

In Figure 4.8 the regression surface is plotted in terms of F_5 and r_5, for fixed levels of G_5 of £100,000, £108,327 and £135,000. These values of G_5 correspond approximately to the 5-th percentile, mean value and the 95-th percentile of the marginal distribution of G_5, respectively. The 50-th percentile was close to £100,000, so the mean value was chosen here instead to give a distinguishable second regression surface. The right-hand side plot shows the regression estimated projected liability value increases with a decreasing year five fund level and also with a decreasing short rate, for a fixed large value of G_5. This is in agreement with the discussion of the structure of the stylised VA product given in Section 4.4. For lower fixed values of G_5, however, we see some spurious effects in the regression surface due to similar reasons as discussed above. In the left-hand plot the guarantee base at year five is fixed at £100,000, so it is impossible for the fund level at year five to be in excess of this level, given the structure of the product. Thus, it does not make any sense to consider this regression surface for values of F_5 beyond £100,000 and this region can be ignored. A similar argument can also explain the spurious behaviour of the middle plot of Figure 4.8.

Finally, in Figure 4.9 the regression surface is plotted in terms of r_5 and G_5, for fixed levels of F_5 of £48,000, £78,297 and £135,000. These values of F_5 correspond approximately to the 5-th, 50-th and 95-th percentiles of the marginal distribution of F_5, respectively. The middle-plot shows the LSMC projected liability distribution increases with an increasing guarantee base level and with a decreasing short rate at year five. This behaviour matches our intuition of the structure of the liabilities of the stylised VA product. Similar behaviour is exhibited for the corresponding surface for F_5 fixed at £48,000, except the whole surface is shifted upward. This also makes sense, since low levels of year five fund value level should give larger liability valuations. The right-hand side plot does not really offer any information, since if F_5 is at £128,000, then the guarantee base is extremely unlikely to be below around £120,000. Therefore, little inference can be made from this plot as the majority of the surface is so improbable.

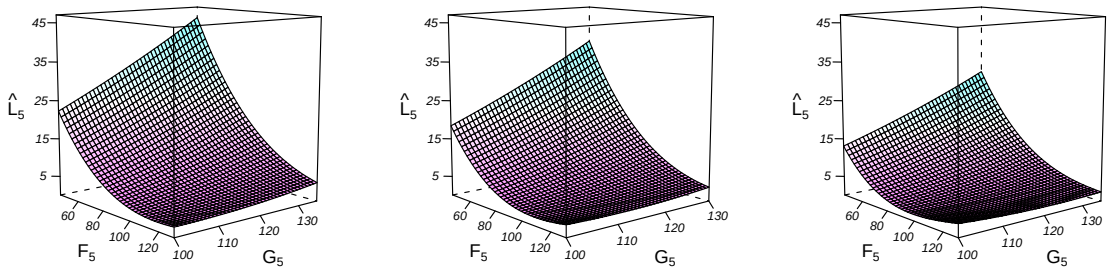

Figure 4.7: Regression surface L_5 in terms of F_5 and G_5 for fixed level of r_5. Variable r_5 fixed at 1% (left plot), 4.3% (middle plot) and 10% (right plot). Note that some regions of the surfaces correspond to essentially impossible combinations of explanatory variables leading to spurious behaviour. See Section 4.4.

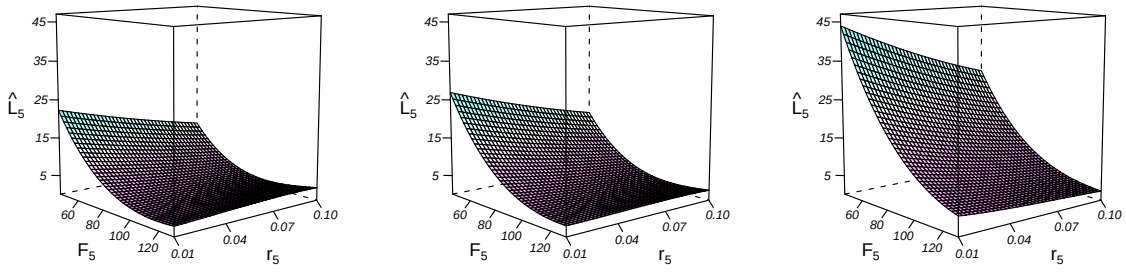

Figure 4.8: Regression surface L_5 in terms of F_5 and r_5 for fixed level of G_5. Term G_5 fixed at £100,000 (left plot), £108,327 (middle plot) and £135,000 (right plot). Note that some regions of the surfaces correspond to essentially impossible combinations of expl. variables leading to spurious behaviour. See Section 4.4.

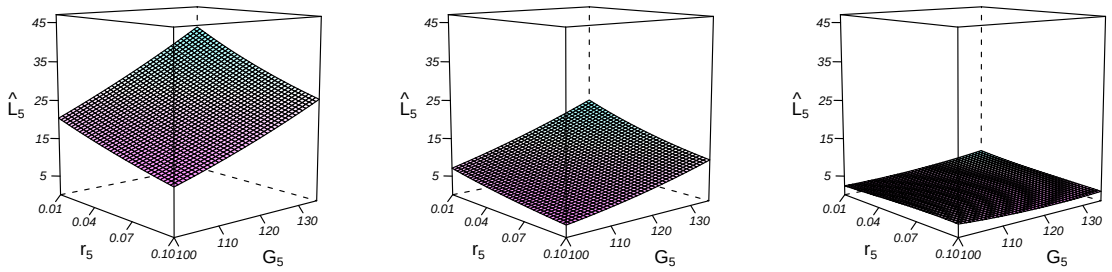

Figure 4.9: Regression surface L_5 in terms of r_5 and G_5 for fixed level of F_5. Term F_5 fixed at £48,000 (left plot), £78,297 (middle plot) and £128,000 (right plot). Note that some regions of the surfaces correspond to essentially impossible combinations of expl. variables leading to spurious behaviour. See Section 4.4.

Overall, these three-dimensional surfaces make sense in terms of our intuition on the behaviour of the liabilities of this stylised VA product. Therefore, let us turn our attention to testing how accurately the LSMC method estimates some key percentiles of the projected liability distribution for a five year projection. As in Section 4.3 we will consider percentiles of the LSMC estimated liability distribution and make the assumption that these will correspond closely with the percentiles of the true liability distribution at the projection year $t = 5$.

In Table 4.3, some key percentiles of the LSMC regression-estimated projected liability distribution are quoted. Also given in this table are full simulation valuations of the year five liabilities (with standard errors) for the outer scenarios which correspond to these percentiles of the regression-estimated liability distribution. For reference the values of F_5, G_5 and r_5 for these outer scenarios are also stated. The percentiles for these tests are, again, chosen to be the 50-th, 75-th, 90-th, 95-th, 99-th and 99.5-th. The results quoted show that the LSMC method gives fairly accurate approximations of the year five projected liability, both in the centre and upper tail of this distribution. The LSMC estimate for the 50-th percentile is only around £60 out from the full simulation estimate. For the 75-th this difference is also only around £120.

As we look further into the tail the LSMC are slightly less accurate with these differences being approximately £400 and £440 at the 95-th and 99-th percentile, respectively. The 99.5-th percentile of the LSMC projected liability distribution then gets within £100 of the full simulation valuation, however. This is, of course, just one of many possible outer scenarios which give an estimated liability valuation in the extreme upper tail of this distribution. If we consider a few other outer scenarios which correspond to other percentiles in the immediate vicinity of this 99.5-th percentile, the LSMC regression estimated distribution is unlikely to give such a good estimate compared to their respective full valuations.

To make this more concrete, in Table 4.4 full simulation valuations for seven percentiles of the LSMC estimated distribution at year five ranging from the 99.47-th to the 99.53-rd. Clearly, some of these percentiles lie within about £100 of the full valuation, while others are out by around £1000. On average, the LSMC estimates are roughly £600 out from the full valuations around the 99.5-th percentile of the projected liability distribution. While, similar analysis for the centre of the distribution

indicates this average difference between LSMC estimate and full valuation is only around £100. This is not surprising, since it is likely to be more difficult to obtain an accurate valuation in the upper tail of the liability distribution. Furthermore, it was argued in Section 3.7 that sampling the fitting points from the real-world distribution would often lead to a poor fit in this region. The results of the LSMC estimates in the upper tail are still very encouraging, however, and interesting lines of further research which could improve this accuracy significantly will be outlined in Section 4.6.

Perc.	F_5	r_5	G_5	Reg. Est.	Full Sim.	St. Err.
50-th	94,299	2.06	114,573	6,220	6,158	34
75-th	69,634	0.98	100,000	9,921	10,042	39
90-th	49,938	6.38	100,943	14,047	13,905	34
95-th	43,387	7.01	100,000	16,624	16,223	31
99-th	42,697	2.94	100,000	21,421	20,981	33
99.5-th	32,405	7.47	100,000	23,289	23,206	28

Table 4.3: Full sim. liability valuations for the real-world outer scenarios which correspond to some key percentiles of the LSMC estimated projected liability distribution. Given for each of these outer trials are the LSMC corrected estimate and also the full sim. valuation together with standard error. All the values in the table are quoted in units of pounds, except for r_5 which is given as a percentage.

Perc.	F_5	r_5	G_5	Reg. Est.	Full Sim.	St. Err.
99.47-th	39,112	3.67	100,000	22,968	22,382	31
99.48-th	39,778	3.25	100,000	23,016	21,695	30
99.49-th	41,254	2.25	100,000	23,229	23,124	33
99.50-th	32,405	7.47	100,000	23,289	23,206	28
99.51-st	39,744	3.01	100,000	23,331	22,590	32
99.52-nd	36,617	4.78	100,000	23,391	22,916	30
99.53-rd	34,677	5.89	100,000	23,455	22,412	28

Table 4.4: Full sim. liability valuations for the real-world outer scenarios which correspond to the close vicinity of the 99.5-th percentile of the LSMC estimated projected liability distribution. Given for each of these outer trials are the LSMC corrected estimate and the full sim. valuation together with standard error. All values in the table are quoted in pounds, except r_5 which is given as a percentage.

4.5 Test of LSMC method: Heston-CIR model

Some tests were also performed applying the LSMC method to the stylised VA product under a Heston-CIR model. This model is given a full introduction in Section 8.3. However, for a one year projection of the liabilities under this set-up there should now be three explanatory variables: F_1, r_1 and v_1, the latter being the level of the equity variance process at the projection time. This analysis showed that this additional explanatory variable v_1 did not have a very significant effect on the LSMC regression model which was found and the results were similar to those given in Section 4.3. Hence, the results will not be reported here.

4.6 Conclusion and further research

Overall, the tests of the LSMC method being applied to the stylised VA product show the technique is successful in modelling projected insurance liabilities. The guidance on how best to configure the method, given in Chapter 3, also seems to extend to the insurance setting, with reasonably accurate LSMC estimates found across the projected liability distribution. This analysis should interest practitioners who are looking to apply novel techniques for estimating their projected liability distribution or calculate a SCR as part of their preparations for the impending the Solvency II regulatory framework.

To conclude this chapter, some interesting lines of further research related to the LSMC method will be outlined. The first of these is to consider a larger number of risk factors and investigate how this affects the accuracy of the LSMC estimates. In the five year projection there were three influential explanatory variables in the LSMC regression model, but if a more sophisticated set-up is considered, additional factors could be included. This could be through extending the modelling to consider stochastic mortality, incorporating an extra stochastic risk driver in the policyholder lapsation model or by including some form of credit risk.

Another line of interesting further research is to investigate how appropriate a statistic the Akaike information criterion (AIC) is for selecting the regression model in the LSMC algorithm. Although the results reported in the thesis appear reasonably promising, Bauer, Bergmann and Reuss [Bau10] note that the AIC relies on "the

rather restrictive assumptions of homoscedasticity and/or normally distributed errors" and that "these assumptions are likely to be violated in [a realistic insurance] setting". They, instead, suggest using a generalised version of Mallows' complexity parameter, as originally proposed by Baek, Karaman, and Ahn [Bae05]. Under this approach the data is divided into smaller groups such that homoscedasticity can be assumed within any one group. Then, the variances are estimated within each group and a generalised version of Mallows' complexity parameter is derived from the weighted least-squares estimators which result. It would be interesting to investigate if such an approach for choosing the form of the regression model will give more accurate LSMC estimates for the stylised VA product, than those resulting from a regression model selected by the stepwise AIC algorithm.

The final area of further research which will be discussed is possible approaches for improving the accuracy of the estimates in the upper percentiles of the projected liability distribution. One reasonably successful approach found in the thesis is to sample the fitting scenarios using a quasi-random scheme, but in some circumstances this can be particularly challenging to implement.

Therefore, it is important to investigate approaches which could improve the fit in the upper tail of the projected liability distribution, whilst still sampling the outer fitting scenarios from the assumed real-world distribution. A general method for achieving this is to sample a larger number of valuation scenarios for the real-world outer scenarios which will give estimates which are in the upper tail of the liability distribution at the projection year. Then, the weight given to each outer fitting scenario in the LSMC regression will be proportional to the number of subsequent inner valuation scenarios which are performed. Such an approach should lead to more accurate LSMC projected liability valuations.

There are a few papers which have offered approaches for obtaining more accurate valuations in the tail of a conditional distribution within a nested simulation set-up. Broadie, Du and Moallemi [Bro11a] develop an algorithm which can dramatically improve the accuracy of estimating the probability of incurring a large loss by sequentially allocating the inner scenario simulations based on marginal changes to some risk-measure for each outer scenario. Juneja and Ramprasath [Jun09] also investigate calculating the probability of a large loss on a portfolio value within a nested simulation framework. They determine a asymptotically optimal budget

allocation based on the theory of large deviations to efficiently allocate the computational budget for such a calculation. Liu, Nelson and Staum [Liu10b] present an efficient simulation procedure for the point estimation of expected shortfall within a nested simulation framework. The approach uses ranking and selection to allocate greater computational resource to estimate the largest losses resulting in improved accuracy in the estimate of expected shortfall. Finally, Liu and Staum [Liu09] investigate using stochastic kriging, a technique used in meta-modelling, to improve the accuracy of estimates of expected shortfall in a nested simulation set-up. This approach improves accuracy in the valuation of a particular scenario, based on information from other scenarios which are 'close' in risk-driver space, allowing for greater computational efficiency.

Investigating whether any of these approaches could be successfully incorporated into the LSMC framework to give more accurate upper tail projected liability valuations is a very interesting line of future research. Some recent papers given by Broadie, Du and Moallemi [Bro11b] and [Bro11c] begin to explore the application of some of these ideas.

Part II

Estimating insurance liability sensitivities

Chapter 5

Heston and SVJD models

In Part I of the thesis the least-squares Monte Carlo simulation technique for estimating the projected value of insurance liabilities was introduced and investigated. As well as being able to accurately value and project complex insurance liabilities, many insurance companies wish to employ a hedging strategy to mitigate some of the risk they are exposed to from offering unit-linked products featuring guarantees on the market.

In order to construct an effective hedging strategy for an option, one needs to know the sensitivities of the option value to the key risk-drivers on which this quantity depends. To hedge some of the risk faced in selling unit-linked insurance products, practitioners must similarly determine the sensitivity of the liability value to its key risk-drivers. Investigating approaches for the calculation of these insurance Greeks is the main purpose of Part II of this thesis. An outline of the structure of the remainder of the thesis and a review of relevant literature was given in Section 1.1. We now introduce the equity models which will be assumed in the constructing the estimators of the liability sensitivities.

5.1 Heston's Model

With unit-linked insurance products it is imperative that we have a realistic model for an equity index, as a common feature of these products is their fund's participation in the stock market. One popular choice of model for an equity index is the standard Black-Scholes model, which offers great tractability and a vast amount of literature employing this model. Much research, however, indicates that despite these benefits, this model does not capture the true dynamics of a typical equity index, such as the FTSE-100. This is especially problematic when considering the challenge of risk management, as the extreme behaviour of the true index returns can act very differently to that predicted by the Black-Scholes model, as discussed and examined by many authors over the years. Indeed Duffie and Pan [Duff97] argue that "many markets return shocks have fatter than normal tails measured ei-

ther by kurtosis or tail critical values at typical confidence levels" and they provide evidence for this fact using the S&P 500 index. Therefore, in the interests of sound risk management one must employ a model of equity dynamics which captures the extremal behaviour of the distribution of returns much more realistically than the Black-Scholes model. While there are various adaptions which can help model these "fatter than normal tails", there are two commonly used approaches which allow some of the structure and tractability of the Black-Scholes model to be kept. These two approaches will be introduced in this section, together with a model which combines them both together to provide a realistic model for describing equity index dynamics.

The first of the two approaches for increasing the probability of extreme returns relative to the Black-Scholes model is to introduce stochastic volatility (SV). This is where the volatility of the asset is assumed to vary randomly in time according to some prescribed model, rather than remain constant as it does under the Black-Scholes model. Thus, under a SV model we now have a stochastic process describing the behaviour of the volatility (or equivalently, the variance) in time, as well as one describing the behaviour of the asset price. We now introduce an example of a model which introduces SV, yet retains much of the analytical tractability of the Black-Scholes model. Indeed, in Chapter 6 we derive an analytical formulae for the European option price and delta and gamma sensitivities under this model.

The Heston stochastic volatility model [Hes93] is defined by the coupled two-dimensional stochastic differential equation (SDE):

$$\begin{cases} dS_t &= rS_t dt + \sqrt{v_t} S_t dW_t^S \\ dv_t &= \kappa(\theta - v_t)dt + \sigma_v \sqrt{v_t} dW_t^v, \end{cases} \tag{5.1}$$

where $\kappa > 0$ represents the mean reversion speed, $\theta > 0$ denotes the mean reversion level and $\sigma_v > 0$ is the volatility of the variance process. The Brownian motions for the asset and volatility processes have correlation ρ, or, $dW_t^S dW_t^v = \rho dt$. The stochastic process S_t represents the prices of an equity asset or index in time and v_t models the instantaneous variance of relative changes to S_t. The dynamics of the variance under the Heston model are the same as the dynamics of interest rates under the famous Cox-Ingersoll-Ross (CIR) model [Cox85], thus we may equally appeal to any literature and results which exist for this model.

One well known result for the Heston model (see for example [And07]) which will be insightful for much of the analysis which follows in this chapter is given in the following proposition:

Proposition 5.1. *Assume that $v_0 > 0$. If $2\kappa\theta \geq \sigma_v^2$ then the process for v is such that $v > 0$ almost surely. If $2\kappa\theta < \sigma_v^2$, then the origin is accessible and strongly reflecting.*

Proof. See Cairns [Cai04] for a proof in the context of the CIR model. □

When calibrating a Heston model to market data it is common for $2\kappa\theta$ to take a value significantly below σ_v^2. Thus, the likelihood of the variance hitting zero is often quite high and the process for v can have a strong affinity for the region around zero. A plot of the Cumulative Distribution Function (CDF) for v, with the parameters the same as those used by Andersen [And07] for a similar plot, demonstrates this affinity to the area close to the origin. This is shown in Figure 5.1. This affinity for the region around zero produces some complications when discretising the variance process. This will be discussed more in Section 5.3.

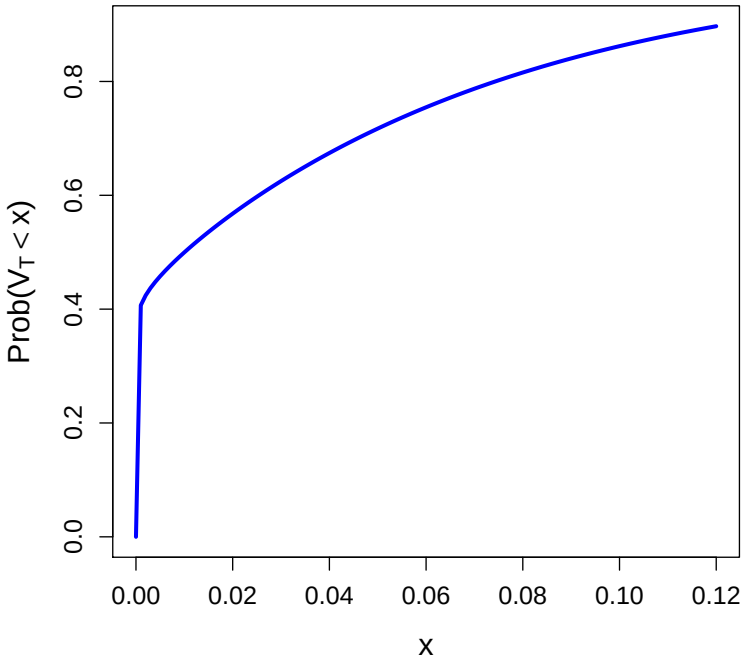

Figure 5.1: The theoretical cumulative distribution function for v_T given v_0, which is given by a non-central chi-squared distribution, as described in the main body of the report. Here $T = 0.1$,$v_0 = \theta = 4\%$, $\kappa = 0.5$ and $\epsilon = 1$. This gives an appreciation of the affinity the variance process has with the region around zero.

5.2 Stochastic volatility jump diffusion (SVJD) model

The second approach which is often used to increase the thickness of a tail of the distribution of equity returns is to incorporate jumps into the stochastic process governing these returns. One simple model which does this is the Merton Jump Diffusion model introduced in 1975 [Mer76]. In recent years many researchers have investigated using pure-jump Levy processes as a more sophisticated model of the jumps found in market equity indices. See Glasserman [Gla03] for more details. In this report we will use the Merton model, but not in isolation. Instead we will discuss a model incorporating the jump component of Merton's model together with the stochastic volatility dynamics of the Heston model, introduced in the previous section. We will refer to this model as the stochastic volatility jump diffusion (SVJD) model, although it is sometimes denoted the Bates model after one of the first researchers to propose it [Bat96].

Under the risk-neutral measure the SVJD model is defined by the following coupled system of SDEs:

$$\begin{cases} dS_t &= (r - \lambda\bar{\mu})S_t dt + \sqrt{v_t}S_t dW_t^S + S_t(J^S - 1)dN_t \\ dv_t &= \kappa(\theta - v_t)dt + \sigma_v\sqrt{v_t}dW_t^v. \end{cases} \tag{5.2}$$

Here N_t denotes a Poisson process independent of W_t^s and W_t^v with constant intensity λ dictating when in time these discontinuous jumps in the equity return process occur and J^s is the relative size of this jump, that is $S_{t^+} = J^s S_{t^-}$ when a jump occurs. This jump size J^s takes a lognormal distribution with mean logarithmic jump size μ_s and variance σ_s^2. The parameter $\bar{\mu} = \exp\left(\mu_s + \frac{1}{2}\sigma_s^2\right) - 1$ then gives the mean jump size of the asset, for example $\bar{\mu} = -0.25$ would represent a fall in equity price of 25% on average, as a result of a jump. Again, the Brownian motions for the asset and volatility processes are assumed to have correlation ρ. By setting $\sigma_v = 0$ and $\theta = v_0$, the original Merton Jump Diffusion model is recovered.

Incorporating both discontinuous jumps and a stochastic process for the volatility in the Black-Scholes model will give a much more realistic depiction of the equity asset returns, particularly in the extremities of the distribution. For a more detailed discussion of this, see, for example, Duffie and Pan [Duff97].

5.3 Simulating from Heston's model

In Chapter 6 we will show how an analytical value can be found for a European option and simple unit-linked insurance liability under the Heston model (and also under the SVJD model). However, for all but simple insurance liabilities such analytical formulae are not available and some numerical technique is required for valuation. In this section, we turn our attention to the problem of simulating the returns from the Heston and SVJD models. As the jump component of the SVJD model is independent of the variance process, this can be simulated in isolation (and without bias) from the Poisson distribution (for the occurrences of the jumps) and the lognormal distribution (for each of these jump's size). Therefore, in the following analysis we will discuss some methods for simulating from the Heston model, but by simply appending an independent jump process the SVJD model can also be simulated by all these approaches. We also note that the CIR interest rate model takes exactly the same form as the Heston model, thus it can be discretised by all of the same methods which apply to the Heston model. This is an important point with regard to the analysis using the Heston-CIR model in Chapter 8.

5.3.1 Full truncation Euler scheme

Perhaps the simplest scheme one could attempt to use to simulate a trajectory of the Heston variance process would be a simple forward difference, or Euler, discretisation scheme. This would express the variance, V_t, and the logarithm of the equity price, $X_t = \ln S_t$, at discrete points in time as

$$\hat{X}_{t+\Delta t} = \hat{X}_t + (r - \frac{1}{2}\hat{v}_t)\Delta t + \sqrt{\hat{v}_t} Z_X \sqrt{\Delta t} \tag{5.3}$$

$$\hat{v}_{t+\Delta t} = \hat{v}_t + \kappa(\theta - \hat{v}_t)\Delta t + \sigma_v \sqrt{\hat{v}_t} Z_v \sqrt{\Delta t}, \tag{5.4}$$

with Z_X and Z_v standard random normal variates with correlation ρ. This can be achieved by setting Z_v as a generated standard random normal variate, then setting $Z_X = \rho Z_v + \sqrt{1-\rho^2} Z_I$, where Z_I is a standard random normal variate independent of Z_v. Clearly this scheme can take values below zero with non-zero probability and thus would fail when attempting to calculate $\sqrt{\hat{v}_t}$ in the discretisation of $\hat{X}_{t+\Delta t}$. A number of possible ways to fix this problem have been proposed, including using

absolute value or absorption functions

$$|y| = \begin{cases} -y & \text{if } y < 0 \\ y & \text{if } y \geq 0 \end{cases} \quad \text{or} \quad y^+ = \max(y, 0). \tag{5.5}$$

Lord et al. [Lor08] compare some of these basic remedies and find the one which results in the least discretisation bias is the following scheme, which they call the Full Truncation Euler:

$$\hat{X}_{t+\Delta t} = \hat{X}_t + (r - \frac{1}{2}\hat{v}_t^+)\Delta t + \sqrt{\hat{v}_t^+} Z_X \sqrt{\Delta t} \tag{5.6}$$

$$\hat{v}_{t+\Delta t} = \hat{v}_t + \kappa(\theta - \hat{v}_t^+)\Delta t + \sigma_v \sqrt{\hat{v}_t^+} Z_v \sqrt{\Delta t}. \tag{5.7}$$

Under this approach, if the process for v falls below zero, then the stochastic process will change to become deterministic with upward drift of $\kappa\theta$, until the variance becomes non-negative. Naturally, whenever we use values from the variance process discretisation within the asset price process discretisation, those values must be truncated to be non-negative. This is the purpose of the maximum functions in Equation 5.6 above.

5.3.2 Andersen moment-matching approach

A slightly more sophisticated discretisation scheme proposed by Andersen [And07] samples the variance values from a distribution which is a superposition of two distinct distributions; one for when the previous timestep variance values are small and a different one when they are large. Andersen proposes two different versions of these composition of distributions to model the Heston variance process. The exact nature of these two moment-matching approaches will be made clear in the remainder of this section, however we note that Andersen gives evidence to suggest that using either of his two proposed methods can significantly reduce the bias in option price estimates when the underlying asset follows a Heston model, compared with using the Full Truncation Euler scheme.

To explain why the sampling approaches introduced by Andersen can provide less bias in simulating paths of the Heston variance process, we use the following proposition and corollary from Andersen [And07]:

Proposition 5.2. *Let $F_{\chi'^2}(y;\nu,\lambda)$ be the cumulative distribution function for the non-central chi-square distribution with ν degrees of freedom and non-centrality parameter λ, i.e.,*

$$F_{\chi'^2}(z;\nu,\lambda) = e^{-\frac{\lambda}{2}} \sum_{j=0}^{\infty} \frac{(\lambda/2)^j}{j!2^{\nu/2+j}\Gamma(\nu/2+j)} \int_0^z z^{\nu/2+j-1} e^{-x/2} dx. \tag{5.8}$$

For the Heston variance process define

$$d = \frac{4\kappa\theta}{\sigma_v^2} \quad \textit{and} \quad n(t,T) = \frac{4\kappa e^{-\kappa(T-t)}}{\sigma_v^2(1-e^{-\kappa(T-t)})}, \quad \text{for } T > t. \tag{5.9}$$

Conditional on v_t, v_T is distributed as $e^{-\kappa(T-t)}/n(t,T)$ times a non-central chi-square distribution with d degrees of freedom and non-centrality parameter $v_t \cdot n(t,T)$, i.e.,

$$\mathbb{P}(v_T < x | v_t) = F_{\chi'^2}\left(\frac{x \cdot n(t,T)}{e^{-\kappa(T-t)}}; d, v_t \cdot n(t,T)\right). \tag{5.10}$$

Corollary 5.3. *Conditional on v_t, v_T has the following first two moments:*

$$\mathbb{E}[v_T|v_t] = \theta + (v_t - \theta)e^{-\kappa(T-t)} \tag{5.11}$$

$$\mathrm{Var}[v_T|v_t] = \frac{v_t\sigma_v^2 e^{-\kappa(T-t)}}{\kappa}\left(1-e^{-\kappa(T-t)}\right) + \frac{\theta\sigma_v^2}{2\kappa}\left(1-e^{-\kappa(T-t)}\right)^2. \tag{5.12}$$

Proof. For a proof of Proposition 5.2 see, for example, Cairns [Cai04] (proof given in the context of the CIR interest rate model). □

Thus, conditionally on the previous timestep variance value, the Heston variance process is described by a non-central chi-squared distribution, with parameters which can be easily determined. It is this exact distribution for the variance that Andersen uses as guidance in developing a more sophisticated discretisation scheme. As the non-centrality parameter tends to infinity, it is known that the non-central chi-square distribution approaches a Gaussian, see, for example, Johnson et al. [Joh95]. Therefore, it would seem that for sufficiently large values of v_t, a good approximation for $v_{t+\Delta t}$ would be a Gaussian random variable, with first two moments matched to those of the exact non-central chi square distribution, which are given explicitly in Corollary 5.3. For small values of v_t however, the non-centrality parameter tends to zero and the distribution of $v_{t+\Delta t}$ approaches that of a (central) chi-square distribution with $d = \frac{4\kappa\theta}{\sigma_v^2}$ degrees of freedom, where the density of a chi-square distribution

with ν degrees of freedom is given by

$$f_{\chi^2}(x;\nu) = \frac{1}{2^{\nu/2}\Gamma(\nu/2)} e^{-x/2} x^{\frac{\nu}{2}-1}. \tag{5.13}$$

In practice d can be significantly less than 2 and in these cases the $x^{\nu/2-1}$ term implies that for small values of v_t the density of $v_{t+\Delta t}$ will be very large around zero. See Figure 5.1 for graphical evidence of this fact. It is fairly clear from this analysis that approximating $v_{t+\Delta t}$ with a moment-matched Gaussian random variable will not be very accurate when v_t is near zero. Finding a discretisation scheme which overcomes this problem for low values of variance is the motivation behind Andersen's two proposed methods.

Andersen's Truncated Gaussian Scheme

The first of Andersen's two proposed sampling schemes for the Heston model is the truncated Gaussian (TG) method. Under the TG method we sample from a moment-matched Gaussian distribution, but all the probability mass below zero is added as a Delta function at the origin. For large values of v_t it is unlikely that $v_{t+\Delta t}$ will reach zero and the scheme will take a value from a moment-matched Gaussian distribution. Whereas, for small values of v_t the scheme will approximate the central chi-square density by a point mass at zero combined with an upper tail density which is proportional to $e^{-x^2/2}$. That is, the TG scheme will determine the variance at time $t + \Delta t$ as

$$\hat{v}_{t+\Delta t} = (\mu + \sigma \cdot Z_v)^+, \tag{5.14}$$

where Z_v is a standard Gaussian random variable and μ and σ are constants that will depend on the size of the timestep Δt, the value of $\hat{v}_t$ and the parameters κ, θ and σ_v from the Heston model.

We will not state explicitly how to calculate μ and σ here, as Andersen states that the alternative quadratic exponential scheme, which will be introduced next, is generally more accurate than the TG scheme, without adding extra complexity or requiring much greater computation time. Thus we shall just proceed to give a full description of this alternative method instead.

The upper tail density for $\hat{v}_{t+\Delta t}$ is modelled as being proportional to $e^{-x^2/2}$ under the TG scheme. For small values of $\hat{v}_t$, however, this decay is too fast. To address this problem Andersen introduced a modified version of the TG scheme which he denotes the quadratic exponential (QE) discretisation scheme. This method will now be introduced.

Firstly we note that a non-central chi-square distribution with moderate or large non-centrality parameter can be well approximated by a power transformation of a Gaussian distribution. See, for example, Patnaik [Pat49]. Thus, in the QE method, for sufficiently large values of $\hat{v}_t$ we set

$$\hat{v}_{t+\Delta t} = a(b + Z_v)^2, \tag{5.15}$$

where Z_v is a standard Gaussian random variable and a and b are constants that will depend on Δt, $\hat{v}_t$ and the three parameters from the Heston variance process. The scheme given by Equation 5.15 does not work well for small values of $\hat{v}_t$, thus it is used in combination with a scheme for low values of $\hat{v}_t$. This low volatility regime scheme is based on the asymptotic approximation of the non-central chi-square density given in Equation 5.13. For this scheme a point mass at the origin is combined with an exponential tail similar to that found in this asymptotic approximation. Formally, the following approximate density for $\hat{v}_{t+\Delta t}$ is used:

$$\mathbb{P}(\hat{v}_{t+\Delta t} \in [x, x + dx]) \approx \big(p\delta(0) + \beta(1-p)e^{-\beta x}\big)dx, \quad x \geq 0, \tag{5.16}$$

where δ is the Dirac delta-function and p and β are constants which are to be determined. This implementation of a point probability mass at the origin is similar to the TG scheme, except the size of this mass, p, is now explicitly defined, rather than implied by other parameters. Some basic analysis shows that if $p \in [0, 1]$ and $\beta \geq 0$, then Equation 5.16 defines a valid probability density function. If we integrate this density function, then we obtain the cumulative distribution function for the low variance regime as

$$\Omega(x) = \mathbb{P}(\hat{v}_{t+\Delta t} \leq x) = p + (1-p)\big(1 - e^{-\beta x}\big), \quad x \geq 0. \tag{5.17}$$

Computing the inverse of Ω,

$$\Omega^{-1}(u;p,\beta) = \begin{cases} 0 & \text{if } 0 \leq u \leq p \\ \beta^{-1}\ln\left(\frac{1-p}{1-u}\right) & \text{if } p < u \leq 1, \end{cases} \tag{5.18}$$

and using the inversion method gives us a simple sampling scheme for $\hat{v}_{t+\Delta t}$ as

$$\hat{v}_{t+\Delta t} = \Omega^{-1}(U_v;p,\beta), \tag{5.19}$$

where U_v is a sample from a standard uniform distribution. Thus, Equations 5.15 and 5.19 define the QE discretisation scheme for large and small $\hat{v}_t$ respectively.

All that remains is to determine how to calculate the constants a, b, p and β and how to define the rule for dictating whether we are in the low- or high-volatility regime. The following two Propositions gives a means for calculating the values of a, b, p and β.

Proposition 5.4. *Let $m = \mathbb{E}[v_{t+\Delta t}|v_t]$ and $s^2 = \mathrm{Var}[v_{t+\Delta t}|v_t]$ using the definitions of these moments given in Corollary 5.3 and set $\Psi = s^2/m^2$. Provided that $\Psi \leq 2$, set*

$$b^2 = 2\Psi^{-1} - 1 + \sqrt{2\Psi^{-1}}\sqrt{2\Psi^{-1}-1} \geq 0 \tag{5.20}$$

$$a = \frac{m}{1+b^2}. \tag{5.21}$$

Let $\hat{v}_{t+\Delta t}$ be given by Equation 5.15. Then $\mathbb{E}\left[\hat{v}_{t+\Delta t}\right] = m$ and $\mathrm{Var}\left[\hat{v}_{t+\Delta t}\right] = s^2$.

Proof. See Andersen [And07]. □

Proposition 5.5. *Let m, s and Ψ be as defined in Proposition 5.4. Assume that $\Psi \geq 1$ and set*

$$p = \frac{\Psi - 1}{\Psi + 1} \in [0,1) \tag{5.22}$$

$$\beta = \frac{1-p}{m} = \frac{2}{m(\Psi+1)} > 0. \tag{5.23}$$

Let $\hat{v}_{t+\Delta t}$ be given by Equation 5.19. Then $\mathbb{E}\left[\hat{v}_{t+\Delta t}\right] = m$ and $\mathrm{Var}\left[\hat{v}_{t+\Delta t}\right] = s^2$.

Proof. See Andersen [And07]. □

In Proposition 5.4 the quadratic sampling scheme can only be moment-matched for $\Psi \leq 2$, whereas in Proposition 5.5 the exponential scheme can only be moment-matched for $\Psi \geq 1$. These domains of applicability overlap and this implies that we should choose some fixed critical level $\Psi_c \in [1, 2]$ ahead of using the QE discretisation scheme. Then, in the QE scheme, we use Equation 5.15 if $\Psi \leq \Psi_c$, and Equation 5.19 otherwise. Thus, the domains of applicability for Propositions 5.4 and 5.5 in effect give us a rule dictating whether we are in the low- or high-volatility regime in the QE method. In the interval where the domains overlap we must make a choice where to demarcate the variance into a low and high regime. Andersen states that the exact choice of $\Psi_c \in [1, 2]$ does not have a significant effect on the success of the discretisation scheme and the value $\Psi_c = 1.5$ is perfectly adequate. A summary of this scheme is now given:

Summary of the quadratic exponential discretisation scheme:

1. Given $\hat{v}_t$, set $m = \mathbb{E}[v_{t+\Delta t}|v_t]$ and $s^2 = \text{Var}[v_{t+\Delta t}|v_t]$ using Corollary 5.3.
2. Compute $\Psi = s^2/m^2$.
3. Generate a uniform random variate, U_v.
4. If $\Psi \leq \Psi_c$:
   (a) Compute a and b as defined in Proposition 5.4.
   (b) Calculate $Z_v = \Phi^{-1}(U_v)$, where $\Phi^{-1}(\cdot)$ represents the inverse of the cumulative standard normal distribution function.
   (c) Set $\hat{v}_{t+\Delta t} = a(b + Z_v)^2$
5. If $\Psi > \Psi_c$:
   (a) Compute β and p as defined in Proposition 5.5.
   (b) Set $\hat{v}_{t+\Delta t} = \Omega^{-1}(U_v; p, \beta)$, with Ω^{-1} given by Equation 5.18.

Andersen then proposes a discretisation scheme for $\ln S_{t+\Delta t}$ based on the exact integral representation

$$\begin{aligned} \ln S_{t+\Delta t} &= \ln S_t + r\Delta t + \frac{\rho}{\epsilon}(v_{t+\Delta t} - v_t - \kappa\theta\Delta t) \\ &\quad + \left(\frac{\kappa\rho}{\epsilon} - \frac{1}{2}\right)\int_t^{t+\Delta t} v_u du + \sqrt{1-\rho^2}\int_t^{t+\Delta t}\sqrt{v_u}dW_u. \end{aligned} \tag{5.24}$$

To use this result in the discretisation of $\ln S_{t+\Delta t}$, we use the approximation

$$\int_t^{t+\Delta t} v_u du \approx \Delta t(\gamma_1 v_t + \gamma_2 v_{t+\Delta t}), \tag{5.25}$$

where γ_1 and γ_2 are constants. Using $\gamma_1 = 1$ and $\gamma_2 = 0$ would give an Euler-like approximation to this integral, whereas setting $\gamma_1 = 0.5$ and $\gamma_2 = 0.5$ would employ a central discretisation.

As the Brownian motion W is independent of v, conditional on v_t and $\int_t^{t+\Delta t} v_u du$, the value of the integral

$$\int_t^{t+\Delta t} \sqrt{v_u} dW_u \tag{5.26}$$

is just a normal random variable with mean zero and variance $\int_t^{t+\Delta t} v_u du$.

Putting all this together, Andersen proposes to discretise $\ln S_{t+\Delta t}$ using

$$\ln \hat{S}_{t+\Delta t} = \ln \hat{S}_t + K_0 + K_1 \hat{v}_t + K_2 \hat{v}_{t+\Delta t} + \sqrt{K_3 \hat{v}_t + K_4 \hat{v}_{t+\Delta t}} \cdot Z, \tag{5.27}$$

where Z is a standard normal random variate, independent of $\hat{v}$, and the K_i are given by

$$\begin{aligned}
K_0 &= \left(r - \frac{\rho\kappa\theta}{\epsilon}\right)\Delta t && (5.28)\\
K_1 &= \gamma_1 \Delta t\left(\frac{\kappa\rho}{\epsilon} - \frac{1}{2}\right) - \frac{\rho}{\epsilon} && (5.29)\\
K_2 &= \gamma_2 \Delta t\left(\frac{\kappa\rho}{\epsilon} - \frac{1}{2}\right) + \frac{\rho}{\epsilon} && (5.30)\\
K_3 &= \gamma_1 \Delta t(1-\rho^2) && (5.31)\\
K_4 &= \gamma_2 \Delta t(1-\rho^2). && (5.32)
\end{aligned}$$

Andersen argues that the QE method is almost always preferable out of the two methods proposed in [And07]. Therefore, whenever the Andersen method is referred to in later parts of this thesis, such discussion will assume this version of the method. In fact the Andersen QE method is the discretisation scheme which is used for all the analysis throughout this thesis which required simulating paths from the Heston and CIR models.

5.3.3 Other possible simulation schemes

There are various other discretisation schemes in the literature and we pause briefly to mention a few of these. Perhaps the most important of these alternative schemes is that of Broadie and Kaya [Bro04]. Although the Andersen quadratic exponential method may reduce the discretisation bias significantly compared to other simple discretisation schemes, it will not eliminate bias completely from estimates of option prices under the Heston equity model. The Broadie and Kaya scheme, however, allows sampling from the Heston (and also the SVJD) model equity prices process without incorporating any discretisation bias whatsoever. It proceeds by sampling directly from the known distribution for the Heston variance process, given in Proposition 5.2, and then uses Fourier inversion techniques to obtain an unbiased sample from the equity price process using the sampled realisation of the variance process.

There are various other discretisation schemes for the Heston model, which similarly to the Andersen model, give some bias in the samples. The Kahl-Jäckel discretisation scheme [Kah05a] is similar to the Euler scheme, except they use an implicit Milstein scheme for the square-root diffusion of the variance process, together with a particular discretisation for the asset price process. Another discretisation scheme for the Heston model that shall be briefly discussed is that proposed by Zhu [Zhu08]. Under this scheme rather than attempting to discretise the mean reverting square root variance process, Zhu proposes an approach to simulate from the discretised transformed volatility process. This new Ornstein-Uhlenbeck has the benefit of not having any square root term. A fourth alternative discretisation scheme which we shall mention is that recently proposed by Halley, Malham and Wiese [Hal09]. They discuss sampling from the non-central chi-squared transition density by simulating Poisson distributed sums of powers of generalised Gaussian random variables. These generalised Gaussian random variables can be simulated efficiently using an extension of Marsaglia's polar method for generating pairs of random normal variables. Halley, Malham and Wiese argue that it provides a good alternative to Andersen's approach, with almost similar speeds, yet simulating exactly from the transition density, rather than through an approximation as in Andersen. Finally, Glasserman and Kim [Gla09] have proposed a method to circumvent the most time-consuming step in the Broadie-Kaya approach. They derive an explicit representation of the

transitions of the Heston model and use this result to give a fast and accurate simulation scheme. The integral of the variance process over an interval, conditional on the value of the variance at the end-points, is found in terms of infinite sums and mixtures of gamma random variables, with the increments of the variance process themselves given as mixtures gamma random variables. Glasserman and Kim show that using this expansion, together with the Broadie-Kaya method, allows exact sampling from the transitions of the Heston model. The analysis presented in the subsequent chapters of this thesis is independent of the specific choice of discretisation scheme.

Chapter 6

Semi-analytical liability values under the Heston model

In this chapter a detailed description of how a semi-analytical formula can be developed for the price of a European option under the Heston model will be given. This analysis will then be extended to value simple unit-linked insurance liabilities. Furthermore, in Chapter 7 we refer to analytical option prices under the Heston model in testing the conditional likelihood ratio method. Therefore, it is important to make clear precisely how such pricing formulae can be derived. Finding such formulae involves quite a few mathematical definitions and arguments, thus these will be constructed from first principles over the first couple of sections of this chapter. This will expand on the treatment given by Gatheral [Gat06] and provide a more detailed explanation of each of the steps in the derivation. Also, some typographical and minor theoretical errors which are present in Gatheral's textbook will be corrected. To begin this chapter, we show how the Fourier transform can be used to give a general option pricing formula, which is independent of the equity model.

6.1 Fourier transform pricing

The aim of this section is to show how the Fourier transform can be used in pricing European options. This analysis holds under a general stochastic equity model, whether it be the Black-Scholes model or a more advanced stochastic volatility model. In order to derive this Fourier transform-based option pricing formula, some key definitions and results are needed. These will also be useful in Section 6.2, so we will state them explicitly here for clarity of exposition. These definitions can be found in many textbooks on partial differential equations and digital signal processing. They are also given in the paper by Matsuda [Mat04].

The first definition that we recall is that of the characteristic function. Given a random variable X with probability density function (pdf) $f(x)$, the characteristic function, $\phi(\omega)$ is given by

$$\phi(\omega) = \mathbb{E}[e^{i\omega X}] = \int_{-\infty}^{\infty} e^{i\omega x} f(x) dx. \tag{6.1}$$

The Fourier transform of the random variable X has an almost equivalent definition to the characteristic function. The only difference is the sign of the exponent in the expectation. In full, the Fourier transform of the random variable X is

$$\mathcal{F}(\omega) = \mathbb{E}[e^{-i\omega X}] = \int_{-\infty}^{\infty} e^{-i\omega x} f(x) dx. \tag{6.2}$$

The pdf of the random variable X can be recovered from the Fourier transform using the inverse Fourier transform formula:

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\omega x} \mathcal{F}(\omega) d\omega. \tag{6.3}$$

These are the main definitions which we will make use of throughout this chapter. It should be noted that there are alternative definitions of Fourier transform and inverse transform that are often used in the literature. Some of the results that follow in this section can appear different under these alternative definitions.

It is now useful to state a couple of key results which will also be needed later. Firstly, we give some general results for taking the Fourier transform of some partial derivatives of a function $f(t, x)$, where the Fourier transform is a mapping $\mathcal{F} : (t, x) \mapsto (t, \omega)$:

$$\mathcal{F}\left(\frac{\partial f(t,x)}{\partial t}\right) = \frac{\partial}{\partial t}\mathcal{F}(f(t,x)) \tag{6.4}$$

$$\mathcal{F}\left(\frac{\partial f(t,x)}{\partial x}\right) = i\omega\mathcal{F}(f(t,x)) \tag{6.5}$$

$$\mathcal{F}\left(\frac{\partial^2 f(t,x)}{\partial x^2}\right) = -\omega^2\mathcal{F}(f(t,x)). \tag{6.6}$$

Proof. Equation 6.4 just follows from the definition of the Fourier transform and exchanging the order of differentiation and integration. Differentiating the inverse Fourier transform, Equation 6.3, and bringing the derivative inside the integral gives $\frac{\partial f}{\partial x} = \frac{1}{2\pi}\int_{-\infty}^{\infty} i\omega\mathcal{F}(\omega)d\omega$. Applying the Fourier transform to both sides of this equation gives us Equation 6.5. Equation 6.6 can be found similarly. □

Next, let us define the Heaviside step function,

$$\theta(x - x_0) = \begin{cases} 1 & \text{if } x > x_0 \\ 0 & \text{if } x \leq x_0. \end{cases} \tag{6.7}$$

The derivative of the Heaviside step function is given by the Dirac delta function, $\delta(x - x_0)$. We can think of this function as being zero everywhere except at $x = x_0$, where there is a 'spike'. Furthermore, if we integrate the Dirac delta function from $x = -\infty$ to $x = \infty$, we get unity. This means the spike at $x = x_0$ must be infinite in magnitude. We will consider an alternative definition of the Dirac delta function:

$$\delta(x - x_0) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\omega(x-x_0)} d\omega. \tag{6.8}$$

See Matsuda [Mat04] for a proof (based upon an argument in Carr [Car]) that this is indeed an identity for the Dirac delta function.

With these definitions made, we can now give a couple of important results which will be of use throughout this and the next section:

Firstly, the Fourier transform of the Heaviside step function is given by

$$\mathcal{F}_\theta(\omega) = \frac{1}{i\omega} + \pi\delta(\omega). \tag{6.9}$$

For a proof of this formula refer to James [Jam11].

The second useful result is often referred to as Parseval's identity. It states that

$$\int_{-\infty}^{\infty} f(x)g(x)dx = \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathcal{F}(-\omega)\mathcal{G}(\omega)d\omega, \tag{6.10}$$

where $\mathcal{F}(\omega)$ and $\mathcal{G}(\omega)$ are the Fourier transforms of the random variables with densities $f(x)$ and $g(x)$, respectively. A short proof of this result will now be given.

Proof. The following argument follows the proof given in Matsuda [Mat04]:

$$\begin{aligned}
\int_{-\infty}^{\infty} f(x)g(x)dx &= \int_{-\infty}^{\infty} \left(\frac{1}{2\pi}\int_{-\infty}^{\infty} e^{i\alpha x}\mathcal{F}(\alpha)d\alpha\right)\left(\frac{1}{2\pi}\int_{-\infty}^{\infty} e^{i\eta x}\mathcal{G}(\eta)d\eta\right)dx \\
&= \left(\frac{1}{2\pi}\right)^2 \int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty} e^{-i\omega x}\mathcal{F}(-\omega)d\omega\right)\left(\int_{-\infty}^{\infty} e^{i\eta x}\mathcal{G}(\eta)d\eta\right)dx \\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty}\mathcal{F}(-\omega)\int_{-\infty}^{\infty}\mathcal{G}(\eta)\left(\frac{1}{2\pi}\int_{-\infty}^{\infty} e^{i(\eta-\omega)x}dx\right)d\eta d\omega \\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty}\mathcal{F}(-\omega)\frac{1}{2\pi}\int_{-\infty}^{\infty}\Big(\mathcal{G}(\eta)\delta(\eta-\omega)\Big)d\eta d\omega \\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty}\mathcal{F}(-\omega)\mathcal{G}(\omega)d\omega. \qquad (6.11)
\end{aligned}$$

In the first line above we have made use of the inverse Fourier transform, Equation 6.3, and in the second line we have changed variable. In the third line the order of integration has been changed and in the fourth line we have invoked our definition of the Dirac delta function, given in Equation 6.8. In the final line we make use of the standard 'sifting' property of the Dirac delta function. □

This brings us to the key result of this section. It has appeared previously in the literature, for example in Carr and Madan [Car99]:

Proposition 6.1. *The expectation of the payoff function under any model for equity returns (whether incorporating stochastic volatility or not) can be expressed as:*

$$\mathbb{E}[\max(S_T - K, 0)] = \frac{1}{2}(F - K) + \frac{1}{\pi}\int_0^{\infty}(F \cdot f_1 - K \cdot f_0)d\omega, \qquad (6.12)$$

where

$$f_1 = \mathrm{Re}\left(\frac{e^{-i\omega\ln(K)}}{i\omega F}\phi_X(\omega - i)\right) \quad , \quad f_0 = \mathrm{Re}\left(\frac{e^{-i\omega\ln(K)}}{i\omega}\phi_X(\omega)\right) \qquad (6.13)$$

and $F = \mathbb{E}[S_T]$ is the forward price of the underlying and $\mathrm{Re}(z)$ denotes the real part of the complex number z.

Proof. This expectation is given in terms of an integral by

$$\mathbb{E}[\max(S_T - K, 0)] = \int_0^{\infty}\max(x - K, 0)f_{S_T}(x)dx, \qquad (6.14)$$

where $f_{S_T}(x)$ is the pdf of the equity returns S_T. If we make the change of variable

$y = \ln(x/K)$ and define $g_Y(y) := f_{S_T}(x(y))\frac{dx(y)}{dy}$, then this expectation can be written as

$$\begin{aligned}
\mathbb{E}[\max(S_T - K, 0)] &= \int_{-\infty}^{\infty} \max(Ke^y - K, 0) f_{S_T}(x) \frac{dx(y)}{dy} dy \\
&= K \int_{-\infty}^{\infty} \max(e^y - 1, 0) g_Y(y) dy \\
&= K \int_{-\infty}^{\infty} (e^y - 1)\theta(y) g_Y(y) dy \\
&= K \left(\int_{-\infty}^{\infty} e^y \theta(y) g_Y(y) dy - \int_{-\infty}^{\infty} \theta(y) g_Y(y) dy \right), \qquad (6.15)
\end{aligned}$$

where $\theta(y)$ is the Heaviside step function, given in Equation 6.7.

Let us consider the second integral from above first. This integral can be represented as

$$\begin{aligned}
\int_{-\infty}^{\infty} g_Y(y)\theta(y) dy &= \frac{1}{2\pi} \int_{\infty}^{\infty} \mathcal{G}_Y(-\omega)\mathcal{F}_\theta(\omega) d\omega \\
&= \frac{1}{2\pi} \int_{-\infty}^{\infty} \phi_Y(\omega) \left(\frac{1}{i\omega} + \pi\delta(\omega) \right) d\omega \\
&= \frac{1}{2} \int_{-\infty}^{\infty} \phi_Y(\omega)\delta(\omega) d\omega + \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{\phi_Y(\omega)}{i\omega} d\omega \\
&= \frac{1}{2} + \frac{1}{2\pi} \int_{0}^{\infty} \left(\frac{\phi_Y(\omega)}{i\omega} + \frac{\phi_Y(-\omega)}{-i\omega} \right) d\omega \\
&= \frac{1}{2} + \frac{1}{\pi} \int_{0}^{\infty} \mathrm{Re}\left(\frac{\phi_Y(\omega)}{i\omega} \right) d\omega. \qquad (6.16)
\end{aligned}$$

The steps of this calculation would benefit from some clarification: In the first line here we have made use of Parserval's identity, Equation 6.10. In the second line we have used the fact that changing the sign of the argument of a Fourier transform gives the characteristic function and we have also used the formula for the Fourier transform of the Heaviside step function, which was given in Equation 6.9. In the fourth line, we have applied the sifting property of the Dirac delta function which, together with the fact that $\phi_Y(0) = 1$, gives the first integral as $1/2$. For the second integral in the fourth line, we have just applied some simple algebraic manipulation. In the final line above, we have used the fact that for any complex number z, we have $z + \bar{z} = 2\mathrm{Re}(z)$, where $\bar{z}$ denotes the complex conjugate of z and $\mathrm{Re}(z)$ denotes the real part of z.

The second integral in the formula for $\mathbb{E}[\max(S_T - K, 0)]$ can be manipulated similarly:

$$\begin{aligned}\int_{-\infty}^{\infty} e^y g_Y(y)\theta(y)dy &= \frac{1}{2\pi}\int_{-\infty}^{\infty}\phi_Y(\omega - i)\left(\frac{1}{i\omega} + \pi\delta(\omega)\right)d\omega \\ &= \frac{1}{2}\int_{-\infty}^{\infty}\phi_Y(\omega - i)\delta(\omega)d\omega + \frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{\phi_Y(\omega - i)}{i\omega}d\omega \\ &= \frac{\frac{F}{K}}{2} + \frac{1}{\pi}\int_0^{\infty}\mathrm{Re}\left(\frac{\phi_Y(\omega - i)}{i\omega}\right)d\omega. \end{aligned} \tag{6.17}$$

In the first line above we have used

$$\int_{-\infty}^{\infty} e^{i\omega y}(e^y g_Y(\omega))d\omega = \int_{-\infty}^{\infty} e^{y+i\omega y} g_Y(y)dy = \int_{-\infty}^{\infty} e^{i(\omega - i)y} g_Y(y)dy = \phi_Y(\omega - i). \tag{6.18}$$

Also, in the first integral in the third line we have used the sifting property of the delta-Dirac function and the fact that $\phi_Y(-i) = \mathbb{E}[e^Y] = \mathbb{E}[\frac{S_T}{K}] = \frac{F}{K}$.

If we let $X_T = \ln(S_T)$, then

$$\phi_Y(\omega) = e^{-i\omega\ln(K)}\phi_X(\omega - i) \quad \text{and} \quad \phi_Y(\omega - i) = \frac{e^{-i\omega\ln(K)}}{K}\phi_X(\omega - i). \tag{6.19}$$

Substituting everything back into Equation 6.15 yields the formula for the expected payoff given in Proposition 6.1. □

6.2 Heston valuation equation

In this section we will now proceed to derive the semi-analytical formula for the price of a European option under the Heston model. The method we will use to find such a formula will involve solving a partial differential equation governing the value of the option under the Heston model dynamics for the underlying asset. The first stage in this process is therefore deriving this partial differential equation for the option value. This equation will then be solved to find the semi-analytical formula for the price in Section 6.2.2.

6.2.1 The valuation equation under stochastic volatility

The partial differential equation (pde) which the option value must satisfy under a general stochastic volatility model for the underlying equity returns will now be derived. This general model includes the Heston model as a special case and we will make use of this pde when we derive a semi-analytical price for a European option under the Heston model dynamics in Section 6.2.2. The following analysis closely follows that of Gatheral [Gat06], which in turn is based on Wilmott [Wil00].

Proposition 6.2. *Under the general stochastic volatility model given by the following stochastic processes*

$$\begin{cases} dS_t &= \mu_t S_t dt + \sqrt{v_t} S_t dW_t^S \\ dv_t &= \alpha(S_t, v_t, t)dt + \eta\beta(S_t, v_t, t)\sqrt{v_t} dW_t^v, \end{cases} \tag{6.20}$$

where W_t^S and W_t^v have correlation ρ, an option with value $V(t, S_t, v_t)$ must satisfy the following valuation pde:

$$\frac{\partial V}{\partial t} + \frac{1}{2}\frac{\partial^2 V}{\partial S^2} v S^2 + \frac{1}{2}\frac{\partial^2 V}{\partial v^2}\eta^2\beta^2 v + \frac{\partial^2 V}{\partial S \partial v}\eta\beta S v\rho + rS\frac{\partial V}{\partial S} - rV = -(\alpha - \phi\beta\eta\sqrt{v})\frac{\partial V}{\partial v}. \tag{6.21}$$

In the above pde ϕ is the market price of volatility risk, which gives the extra expected return on the asset for each extra unit of volatility risk borne. This will be explained in more detail shortly. A full proof of Proposition 6.2 will now be given.

Proof. Following Gatheral, firstly let $V(t, S_t, v_t)$ denote the value of some option under the stochastic equity model described in Equation 6.20. By applying the two-dimensional form of Itô's lemma, the instantaneous change in value of the option can be represented by the following stochastic process:

$$\begin{aligned} dV &= \frac{\partial V}{\partial t}dt + \frac{\partial V}{\partial S}dS + \frac{\partial V}{\partial v}dv + \frac{1}{2}\left(\frac{\partial^2 V}{\partial S^2}dS^2 + \frac{\partial^2 V}{\partial v^2}dv^2 + 2\frac{\partial^2 V}{\partial S \partial v}dSdv\right) \\ &= \frac{\partial V}{\partial t}dt + \frac{\partial V}{\partial S}dS + \frac{\partial V}{\partial v}dv + \frac{1}{2}\left(\frac{\partial^2 V}{\partial S^2}vS^2 + \frac{\partial^2 V}{\partial v^2}\eta^2\beta^2 v + 2\frac{\partial^2 V}{\partial S \partial v}\eta\beta S v\rho\right)dt \\ &= \frac{\partial V}{\partial S}dS + \frac{\partial V}{\partial v}dv + \left(\frac{\partial V}{\partial t} + \frac{1}{2}\frac{\partial^2 V}{\partial S^2}vS^2 + \frac{1}{2}\frac{\partial^2 V}{\partial v^2}\eta^2\beta^2 v + \frac{\partial^2 V}{\partial S \partial v}\eta\beta S v\rho\right)dt. \end{aligned} \tag{6.22}$$

Now, in order to find the price of an option, we aim to construct a replicating portfolio which perfectly hedges changes in the option value over an instantaneous timestep. Under the Black-Scholes model this (theoretical) continuous-time hedging can be achieved by just investing in the underlying stock. However, under the stochastic volatility model of Equation 6.20 random changes in the volatility must also be hedged. Thus, the risk-less portfolio in this case will contain the option which we we are interested in pricing, some holding, $-\Delta$ units say, in the underlying equity asset and now also some holding, $-\Delta_1$ units say, in another asset with value V_1 which is dependent on the volatility of the equity asset. By a negative holding in an asset we mean we hold a short position of this number of units in that asset. Here we are both 'delta hedging' the equity risk factor by investing in the underlying asset and 'vega hedging' the volatility risk by investing in the asset with value V_1. The value of this portfolio is given by

$$\Pi = V - \Delta S - \Delta_1 V_1, \tag{6.23}$$

where we have stopped explicitly denoting the dependence on t, S_t and v_t in the notation to keep the forthcoming analysis easy to follow. Using the two-dimensional form of Itô's lemma, the instantaneous change in the value of this portfolio can be given by

$$\begin{aligned} d\Pi &= dV - \Delta dS - \Delta_1 dV_1 \\ &= \left(\frac{\partial V}{\partial t} + \frac{1}{2}\frac{\partial^2 V}{\partial S^2} v S^2 + \frac{1}{2}\frac{\partial^2 V}{\partial v^2}\eta^2\beta^2 v + \frac{\partial^2 V}{\partial S \partial v}\eta\beta S v\rho\right)dt \\ &\quad - \Delta_1\left(\frac{\partial V_1}{\partial t} + \frac{1}{2}\frac{\partial^2 V_1}{\partial S^2} v S^2 + \frac{1}{2}\frac{\partial^2 V_1}{\partial v^2}\eta^2\beta^2 v + \frac{\partial^2 V_1}{\partial S \partial v}\eta\beta S v\rho\right)dt \\ &\quad + \left(\frac{\partial V}{\partial S} - \Delta_1\frac{\partial V_1}{\partial S} - \Delta\right)dS + \left(\frac{\partial V}{\partial v} - \Delta_1\frac{\partial V_1}{\partial v}\right). \end{aligned} \tag{6.24}$$

Now, this portfolio must be instantaneously risk-free, thus the coefficients before the dS and dv here must be set equal to zero. This gives the following two pde's which must hold:

$$\frac{\partial V}{\partial S} - \Delta_1\frac{\partial V_1}{\partial S} - \Delta = 0, \tag{6.25}$$

$$\frac{\partial V}{\partial v} - \Delta_1\frac{\partial V_1}{\partial v} = 0. \tag{6.26}$$

The conditions in Equations 6.25 and 6.26 ensure the portfolio Π will now grow deterministically over an instantaneous timestep. In order to avoid creating arbitrage opportunities, this deterministic growth must be at the same rate of return, r, as some risk-free reference asset. Thus, in order to avoid arbitrage, the following pde must hold:

$$\begin{aligned}&\left(\frac{\partial V}{\partial t}+\frac{1}{2}\frac{\partial^2 V}{\partial S^2}vS^2+\frac{1}{2}\frac{\partial^2 V}{\partial v^2}\eta^2\beta^2 v+\frac{\partial^2 V}{\partial S\partial v}\eta\beta Sv\rho\right)dt\\&-\Delta_1\left(\frac{\partial V_1}{\partial t}+\frac{1}{2}\frac{\partial^2 V_1}{\partial S^2}vS^2+\frac{1}{2}\frac{\partial^2 V_1}{\partial v^2}\eta^2\beta^2 v+\frac{\partial^2 V_1}{\partial S\partial v}\eta\beta Sv\rho\right)dt=r\Pi dt.\end{aligned}\tag{6.27}$$

The right-hand side of this equation can be expanded as

$$\begin{aligned}r\Pi dt&=r(V-\Delta S-\Delta_1 V_1)dt\\&=rVdt-r\left(\frac{\partial V}{\partial S}-\Delta_1\frac{\partial V_1}{\partial S}\right)Sdt-r\Delta_1 V_1 dt,\end{aligned}\tag{6.28}$$

where in the first line we have substituted in the definition of Π given in Equation 6.23 and in the second line we have substituted in the expression for Δ given by rearranging Equation 6.25. Substituting this expression for $r\Pi dt$ back into Equation 6.27 and rearranging allows us to express the pde in the form:

$$\begin{aligned}&\frac{\partial V}{\partial t}+\frac{1}{2}\frac{\partial^2 V}{\partial S^2}vS^2+\frac{1}{2}\frac{\partial^2 V}{\partial v^2}\eta^2\beta^2 v+\frac{\partial^2 V}{\partial S\partial v}\eta\beta Sv\rho+rS\frac{\partial V}{\partial S}-rV\\&=\Delta_1\left(\frac{\partial V_1}{\partial t}+\frac{1}{2}\frac{\partial^2 V_1}{\partial S^2}vS^2+\frac{1}{2}\frac{\partial^2 V_1}{\partial v^2}\eta^2\beta^2 v+\frac{\partial^2 V_1}{\partial S\partial v}\eta\beta Sv\rho+rS\frac{\partial V_1}{\partial S}-rV_1\right).\end{aligned}\tag{6.29}$$

Now, Equation 6.26 implies that

$$\Delta_1=\frac{\frac{\partial V}{\partial v}}{\frac{\partial V_1}{\partial v}}\tag{6.30}$$

and substituting this expression for Δ_1 into Equation 6.29 results in

$$\begin{aligned}&\frac{1}{\frac{\partial V}{\partial v}}\left(\frac{\partial V}{\partial t}+\frac{1}{2}\frac{\partial^2 V}{\partial S^2}vS^2+\frac{1}{2}\frac{\partial^2 V}{\partial v^2}\eta^2\beta^2 v+\frac{\partial^2 V}{\partial S\partial v}\eta\beta Sv\rho+rS\frac{\partial V}{\partial S}-rV\right)\\&=\frac{1}{\frac{\partial V_1}{\partial v}}\left(\frac{\partial V_1}{\partial t}+\frac{1}{2}\frac{\partial^2 V_1}{\partial S^2}vS^2+\frac{1}{2}\frac{\partial^2 V_1}{\partial v^2}\eta^2\beta^2 v+\frac{\partial^2 V_1}{\partial S\partial v}\eta\beta Sv\rho+rS\frac{\partial V_1}{\partial S}-rV_1\right).\end{aligned}\tag{6.31}$$

Clearly, the left-hand side of Equation 6.31 is a function of V only, while the right-hand side is a function involving only V_1. For this equality to hold, it must therefore

be the case that both sides are equal to some function involving only the independent variables t, S and v. Thus, the value, V, of the option must satisfy the following pde:

$$\frac{1}{\frac{\partial V}{\partial v}}\left(\frac{\partial V}{\partial t}+\frac{1}{2}\frac{\partial^2 V}{\partial S^2}vS^2+\frac{1}{2}\frac{\partial^2 V}{\partial v^2}\eta^2\beta^2 v+\frac{\partial^2 V}{\partial S\partial v}\eta\beta Sv\rho+rS\frac{\partial V}{\partial S}-rV\right)=f(S,v,t). \tag{6.32}$$

Now, without loss of generality the function $f(S,v,t)$ can be expressed as

$$f(S,v,t)=-\big(\alpha(S,v,t)-\phi(S,v,t)\beta(S,v,t)\eta(S,v,t)\sqrt{v}\big). \tag{6.33}$$

We note that this form for f is slightly different to that given in Gatheral. The reasons for choosing to express this arbitrary function in this form will be discussed in a moment, along with an explanation of why the form for f given by Gatheral is incongruent with his subsequent discussion. However, we now complete the proof of the valuation equation: The value, V, of an option under the general stochastic volatility model, Equation 6.20, must satisfy the following pde:

$$\frac{\partial V}{\partial t}+\frac{1}{2}\frac{\partial^2 V}{\partial S^2}vS^2+\frac{1}{2}\frac{\partial^2 V}{\partial v^2}\eta^2\beta^2 v+\frac{\partial^2 V}{\partial S\partial v}\eta\beta Sv\rho+rS\frac{\partial V}{\partial S}-rV=-(\alpha-\phi\beta\eta\sqrt{v})\frac{\partial V}{\partial v}. \tag{6.34}$$

□

An Aside: The Market Price of the Volatility of Risk

So why do we choose to express the arbitrary function $f(S,v,t)$ in the form of Equation 6.33? In the next couple of paragraphs we show that choosing this representation for f is useful in terms of pricing options using the valuation equation (which is given for general f by Equation 6.32).

For the moment, let us consider an alternative portfolio. This will consist of the option which we want to price and $-\Delta$ units of the underlying equity asset. In other words, we are now delta-hedging, but not vega-hedging the option. The value of this portfolio is given by

$$\Pi^\star = V-\Delta S. \tag{6.35}$$

Using Itô's lemma, the instantaneous change in the value of this portfolio can be

expressed as

$$d\Pi^\star = \left(\frac{\partial V}{\partial t}+\frac{1}{2}\frac{\partial^2 V}{\partial S^2}vS^2+\frac{1}{2}\frac{\partial^2 V}{\partial v^2}\eta^2\beta^2 v+\frac{\partial^2 V}{\partial S\partial v}\eta\beta Sv\rho\right)dt+\left(\frac{\partial V}{\partial S}-\Delta\right)dS+\left(\frac{\partial V}{\partial v}\right)dv. \tag{6.36}$$

Of course, under a delta-hedging strategy we choose $\Delta = \frac{\partial V}{\partial S}$ which will eliminate the risk from the dS term. Then

$$\begin{aligned} d\Pi^\star - r\Pi^\star dt &= \left(\frac{\partial V}{\partial t} + \frac{1}{2}\frac{\partial^2 V}{\partial S^2}vS^2 + \frac{1}{2}\frac{\partial^2 V}{\partial v^2}\eta^2\beta^2 v \right. \\ &\qquad \left. + \frac{\partial^2 V}{\partial S\partial v}\eta\beta Sv\rho + rS\frac{\partial V}{\partial S} - rV\right)dt + \left(\frac{\partial V}{\partial v}\right)dv \\ &= -(\alpha - \phi\beta\eta\sqrt{v})\frac{\partial V}{\partial v}dt + \frac{\partial V}{\partial v}\big(\alpha dt + \eta\beta\sqrt{v}dW^v\big) \\ &= \beta\eta\sqrt{v}\frac{\partial V}{\partial v}\big(\phi dt + dW^v\big), \end{aligned} \tag{6.37}$$

where in going from the first to the second line above we have used Equation 6.21 to replace the terms inside the first bracket and Equation 6.20 to substitute in an expression for dv.

From Equation 6.37 we can see that for every unit of volatility risk, dW^v, that we take on, we get an extra $\phi(S, v, t)dt$ units of expected return. In an analogous manner to the capital asset pricing model (CAPM), we call the term ϕ the market price of volatility risk.

If we define the risk-neutral drift in the general variance process by

$$\alpha' = \alpha - \beta\eta\sqrt{v}\phi, \tag{6.38}$$

then starting with a risk-neutral drift in the variance process of Equation 6.20 would have given identical results to choosing to set the market price of volatility risk term to zero. This makes sense as setting the drift as α' ensures we are in a risk-neutral world. In Section 6.2.2 we will assume that we are always working under the risk-neutral measure, as is required in order to price options in a consistent manner in the market. We will drop the prime on the alpha to explicitly state this fact.

If one is studying the above alongside the derivation given by Gatheral [Gat06] an important point should be heeded. Gatheral uses $f = -(\alpha - \phi\beta\sqrt{v})$, instead of our form of f in Equation 6.33. After going through the same calculations to those above,

he ends up with a similar form to the bottom line of Equation 6.37. Unfortunately, he has made a mistake in getting to this equation (which is acknowledged in the errata list for the textbook, which is available online). If one corrects for this error, then the final form of his equation is not compatible with the discussion which is given subsequently in his text. Thus, taking into account this error in the calculation, one must really start with our form of the function $f(S, v, t)$ and not the form given by Gatheral.

6.2.2 Semi-analytical option price under the Heston model

In this section we will now give a derivation of the semi-analytical formula for the price of a European call option under the Heston stochastic volatility model. As there are quite a few steps involved in finding this pricing formula, we will describe these in the form of a series of propositions. After outlining how the construction of the pricing formula proceeds, each of these propositions will then be proved. Later in the chapter semi-analytic formulae will be derived for valuing simple liabilities under the Heston model. Therefore, making clear exactly how such pricing formulae for European options under this model are derived is a worthwhile exercise. The formula for the price of a European call option under the Heston model is stated in the following proposition:

Proposition 6.3. *Suppose the value of an equity asset follows the Heston stochastic volatility model, described by the following stochastic differential equations,*

$$\begin{cases} dS_t &= \mu_t S_t dt + \sqrt{v_t} S_t dW_t^S \\ dv_t &= \kappa(\theta - v_t)dt + \sigma_v \sqrt{v_t} dW_t^v, \end{cases} \tag{6.39}$$

where the Brownian motions driving the two stochastic processes, W_t^S and W_t^v, have correlation ρ. Then, the price of a European call option written on this asset at time t can be expressed semi-analytically as

$$c^{SV} = e^{-r\tau} K(e^x P_1(x, v, \tau) - P_0(x, v, \tau)), \tag{6.40}$$

where K is the strike price of the option, r is the risk-free rate, $x = \ln(S_t e^{r\tau}/K)$ and $\tau = T - t$.

The terms P_0 and P_1 are given by

$$P_j(x, v, \tau) = \frac{1}{2} + \frac{1}{\pi} \int_0^\infty \mathrm{Re}\left(\frac{\exp(i\omega x + C_j(\omega, \tau)\theta + D_j(\omega, \tau)v)}{i\omega} \right) d\omega, \tag{6.41}$$

where

$$C_j(\omega, \tau) = \kappa \left(r_- \tau - \frac{2}{\sigma_v^2} \ln \left(\frac{1 - g e^{-d\tau}}{1 - g} \right) \right) \tag{6.42}$$

$$D_j(\omega, \tau) = r_- \left(\frac{1 - \exp(-d\tau)}{1 - g \exp(-d\tau)} \right), \tag{6.43}$$

with $g = r_-/r_+$, $r_\pm = (\beta \pm d)/\sigma_v^2$, $d = \sqrt{\beta^2 - 4\alpha\gamma}$, $\beta = \kappa - \rho\sigma_v j - \rho\sigma_v i\omega$, $\alpha = -(\omega^2 + i\omega)/2 + ij\omega$ and $\gamma = \sigma_v^2/2$. The κ, θ, σ_v and ρ are the parameters of the Heston variance process. Recall, the function $\mathrm{Re}(z)$ gives the real part of z.

The integrals in the definition of P_0 and P_1 can be solved numerically and the results can then be used to calculate the option price c^{SV}. As numerical quadrature is required to evaluate these integrals, this is why we refer to the formula for the European option price under the Heston model as semi-analytical.

Rather than giving a proof of Proposition 6.3 as a whole, it is easier to follow such an argument if we divide this into a number of smaller propositions and then prove each of these in turn. Thus, we will now state these propositions, which give an overview of how such a pricing formula is constructed:

Proposition 6.4. *The valuation equation under a general stochastic volatility model, derived in the last section, Equation 6.21, becomes*

$$\frac{\partial V}{\partial t} + \frac{1}{2}\frac{\partial^2 V}{\partial S^2} v S^2 + \frac{1}{2}\frac{\partial^2 V}{\partial v^2}\sigma_v^2 v + \frac{\partial^2 V}{\partial S \partial v}\sigma_v S v \rho + rS\frac{\partial V}{\partial S} - rV = \kappa(v_t - \theta)\frac{\partial V}{\partial v}, \tag{6.44}$$

when applied to the Heston model of equity returns, Equation 6.39.

This pde can be expressed more simply if a change of variable is considered.

Proposition 6.5. *Let us employ a change of variable*

$$x = \ln\left(\frac{e^{r\tau} S_t}{K}\right) = \ln\left(\frac{F}{K}\right), \tag{6.45}$$

where $\tau = T - t$ is the time to maturity and $F = e^{r\tau} S_t$ is the forward price of the stock at maturity and let $c^{\mathrm{SV}}(\tau, x)$ denote the future value to expiration of the

European option price. That is,

$$c^{\text{SV}}(\tau, x) = e^{r\tau} V(t, S_t). \tag{6.46}$$

Then, the Heston valuation pde, Equation 6.44, becomes:

$$-\frac{\partial c^{\text{SV}}}{\partial \tau} + \frac{1}{2} v \frac{\partial^2 c^{\text{SV}}}{\partial x^2} - \frac{1}{2} v \frac{\partial c^{\text{SV}}}{\partial x} + \frac{1}{2} \sigma_v^2 v \frac{\partial^2 c^{\text{SV}}}{\partial v^2} + \rho \sigma_v v \frac{\partial^2 c^{\text{SV}}}{\partial x \partial v} - \kappa(v - \theta) \frac{\partial c^{\text{SV}}}{\partial v} = 0. \tag{6.47}$$

To determine the pricing formula we need to solve Equation 6.47 for c^{SV}. This will be performed in the remaining propositions of this section.

Proposition 6.6. *If we assume that Equation 6.47 has a solution of the form*

$$c^{\text{SV}}(x, v, \tau) = K\big(e^x P_1(x, v, \tau) - P_0(x, v, \tau)\big), \tag{6.48}$$

then the following pde's hold for P_0 and P_1

$$-\frac{\partial P_j}{\partial \tau} + \frac{1}{2} v \frac{\partial^2 P_j}{\partial x^2} - \left(\frac{1}{2} - j\right) v \frac{\partial P_j}{\partial x} + \frac{1}{2} \sigma_v^2 v \frac{\partial^2 P_j}{\partial v^2} + \rho \sigma_v v \frac{\partial^2 P_j}{\partial x \partial v} + (a - b_j v) \frac{\partial P_j}{\partial v} = 0, \tag{6.49}$$

where

$$a = \kappa\theta, \qquad b_j = \kappa - j\rho\sigma_v. \tag{6.50}$$

Also, the initial conditions for the pde's for P_0 and P_1 can be stated as

$$\lim_{\tau \to 0} P_j(x, v, \tau) = \begin{cases} 1 & \text{if } x > 0 \\ 0 & \text{if } x \leq 0 \end{cases} = \theta(x), \tag{6.51}$$

where $\theta(x)$ is the Heaviside-step function, introduced in Section 6.1.

The pde's given by Equation 6.49 for $j = \{0, 1\}$ can be expressed more simply by taking a Fourier transform. This will result in pde's for the Fourier transforms of the P_j which can be found using the following proposition:

Proposition 6.7. *Taking a Fourier transform of Equations 6.49 and the corresponding initial conditions 6.51 allows us to find the following pde's for $\tilde{P}_0$ and $\tilde{P}_1$, the Fourier transforms of P_0 and P_1 respectively*

$$v\left(\alpha \tilde{P}_j - \beta \frac{\partial \tilde{P}_j}{\partial v} + \gamma \frac{\partial^2 \tilde{P}_j}{\partial v^2}\right) + a \frac{\partial \tilde{P}_j}{\partial v} - \frac{\partial \tilde{P}_j}{\partial \tau} = 0, \tag{6.52}$$

where we define the variables α, β and γ by

$$\alpha = -\frac{\omega^2}{2} - \frac{i\omega}{2} + ij\omega, \tag{6.53}$$

$$\beta = \kappa - \rho\sigma_v j - \rho\sigma_v i\omega, \tag{6.54}$$

$$\gamma = \frac{\sigma_v^2}{2} \tag{6.55}$$

and recall that $a = \kappa\theta$.

These pdes for $\tilde{P}_j$, $j = \{0, 1\}$, can be solved by assuming a solution of a certain form involving two new unknown functions of the variables ω and τ. The two new functions then must satisfy a system of ordinary differential equations which can then be solved. These steps will be made clearer in the two propositions which follow:

Proposition 6.8. *If we assume a solution of pde's 6.52 of the form*

$$\begin{aligned}\tilde{P}_j(u, v, \tau) &= \exp(C(\omega, \tau)\theta + D(\omega, \tau)v)\tilde{P}_j(\omega, v, 0) \\ &= \exp(C(\omega, \tau)\theta + D(\omega, \tau)v)\left(\frac{1}{i\omega} + \pi\delta(\omega)\right),\end{aligned} \tag{6.56}$$

then functions C and D must satisfy the following system of differential equations:

$$\begin{aligned}\frac{\partial C}{\partial \tau} &= \kappa D \\ \frac{\partial D}{\partial \tau} &= \alpha - \beta D + \gamma D^2 = \gamma(D - r_+)(D - r_-).\end{aligned} \tag{6.57}$$

Proposition 6.9. *The solution to system of differential equations 6.57 is:*

$$D(\omega, \tau) = r_-\left(\frac{1 - \exp(-d\tau)}{1 - g\exp(-d\tau)}\right) \tag{6.58}$$

$$C(\omega, \tau) = \kappa\left(r_-\tau - \frac{2}{\sigma_v^2}\ln\left(\frac{1 - ge^{-d\tau}}{1 - g}\right)\right) \tag{6.59}$$

where the variable g is defined as $g := \frac{r_-}{r_+}$.

Finally, taking the inverse Fourier transform of Equation 6.52 results in formulae for the terms P_j, $j = \{0, 1\}$. This idea is formalised in Proposition 6.10.

Proposition 6.10. *Taking the inverse Fourier transform of Equation 6.52 gives the following formulae for the functions P_j:*

$$\begin{aligned} P_j(x,v,\tau) &= \frac{1}{2} + \frac{1}{\pi}\int_0^\infty \mathrm{Re}\left(\frac{\phi(\omega)}{i\omega}\right) \\ &= \frac{1}{2} + \frac{1}{\pi}\int_0^\infty \mathrm{Re}\left(\frac{\exp(i\omega x + C_j(\omega,\tau)\theta + D_j(\omega,\tau)v)}{i\omega}\right). \end{aligned} \tag{6.60}$$

These formulae are given in terms of the functions $C_j(\omega,\tau)$ and $D_j(\omega,\tau)$. However, the form of these functions was given in Proposition 6.9. The values of P_0 and P_1 can now be substituted back into Equation 6.48 to give the semi-analytical formula for the value of a European call option under the Heston stochastic volatility model.

Having given an overview of the structure of the proof through this series of propositions, let us now turn our attention to proving each of these in turn. We will begin with Proposition 6.4.

Proof. The Heston model can be recovered from the more general stochastic volatility model given in the last section, Equation 6.20, by setting $\alpha(S,v,t) = \kappa(\theta - v_t)$, $\beta(S,v,t) = 1$ and $\eta = \sigma_v$. With these choices the valuation equation, given by Equation 6.21 for the general stochastic volatility model, now becomes:

$$\frac{\partial V}{\partial t} + \frac{1}{2}\frac{\partial^2 V}{\partial S^2}vS^2 + \frac{1}{2}\frac{\partial^2 V}{\partial v^2}\sigma_v^2 v + \frac{\partial^2 V}{\partial S \partial v}\sigma_v S v\rho + rS\frac{\partial V}{\partial S} - rV = -\big(\kappa(\theta - v_t) - \phi\sigma_v\sqrt{v}\big)\frac{\partial V}{\partial v}. \tag{6.61}$$

Now, as it is our aim to find the price of a European option, the valuation should be performed in a risk-neutral setting, as is standard in financial mathematics literature. At the end of Section 6.2.1, we argued that pricing under the risk-neutral measure was equivalent to setting the market price of volatility risk, ϕ, equal to zero. The valuation equation for the Heston model under the risk-neutral measure can then be expressed as:

$$\frac{\partial V}{\partial t} + \frac{1}{2}\frac{\partial^2 V}{\partial S^2}vS^2 + \frac{1}{2}\frac{\partial^2 V}{\partial v^2}\sigma_v^2 v + \frac{\partial^2 V}{\partial S \partial v}\sigma_v S v\rho + rS\frac{\partial V}{\partial S} - rV = \kappa(v_t - \theta)\frac{\partial V}{\partial v}. \tag{6.62}$$

□

In Proposition 6.5 this pde was given in terms of a set of new variables which were introduced. A proof of this Proposition is given next.

Proof. Firstly, let

$$x = \ln\left(\frac{e^{r\tau}S_t}{K}\right) = \ln\left(\frac{F}{K}\right) \tag{6.63}$$

where $\tau = T - t$ is the time to maturity and $F = e^{r\tau}S_t$ is the forward price of the stock at maturity. Then, let $c^{\mathrm{SV}}(\tau, x)$ denote the future value to expiration of the European option price. That is,

$$c^{\mathrm{SV}}(\tau, x) = e^{r\tau}V(t, S_t). \tag{6.64}$$

With this change of variables, the partial derivative $\frac{\partial V}{\partial t}$ can be expressed in terms of $c^{\mathrm{SV}}(\tau, x)$ as follows:

$$\begin{aligned}
\frac{\partial V}{\partial t} = \frac{\partial}{\partial t}\big(e^{-r\tau}c^{\mathrm{SV}}(\tau, x)\big) &= \frac{\partial}{\partial t}\big(e^{-r\tau}\big)c^{\mathrm{SV}}(\tau, x) + e^{-r\tau}\frac{\partial}{\partial t}\big(c^{\mathrm{SV}}(\tau, x)\big) \\
&= re^{-r\tau}c^{\mathrm{SV}}(\tau, x) + e^{-r\tau}\frac{\partial c^{\mathrm{SV}}}{\partial \tau}\frac{\partial \tau}{\partial t} + e^{-r\tau}\frac{\partial c^{\mathrm{SV}}}{\partial x}\frac{\partial x}{\partial t} \\
&= re^{-r\tau}c^{\mathrm{SV}}(\tau, x) - e^{-r\tau}\frac{\partial c^{\mathrm{SV}}}{\partial \tau} - re^{-r\tau}\frac{\partial c^{\mathrm{SV}}}{\partial x} \\
&= e^{-r\tau}\left(rc^{\mathrm{SV}}(\tau, x) - \frac{\partial c^{\mathrm{SV}}}{\partial \tau} - r\frac{\partial c^{\mathrm{SV}}}{\partial x}\right).
\end{aligned} \tag{6.65}$$

Also, the derivative $\frac{\partial V}{\partial S}$ can be expressed in terms of the changed variables as:

$$\begin{aligned}
\frac{\partial V}{\partial S} = \frac{\partial}{\partial S}\big(e^{-r\tau}c^{\mathrm{SV}}(\tau, x)\big) &= e^{-r\tau}\frac{\partial c^{\mathrm{SV}}}{\partial x}\frac{\partial x}{\partial S} \\
&= e^{-r\tau}\frac{\partial c^{\mathrm{SV}}}{\partial x}\frac{K}{e^{r\tau}S}\cdot\frac{e^{r\tau}}{K} \\
&= \frac{e^{-r\tau}}{S}\frac{\partial c^{\mathrm{SV}}}{\partial x}.
\end{aligned} \tag{6.66}$$

In a similar manner, each of the following partial derivatives can be expressed as in terms of $c^{\mathrm{SV}}(\tau, x)$ as:

$$\frac{\partial^2 V}{\partial S^2} = \frac{e^{-r\tau}}{S^2}\left(-\frac{\partial c^{\mathrm{SV}}}{\partial x} + \frac{\partial^2 c^{\mathrm{SV}}}{\partial x^2}\right), \tag{6.67}$$

$$\frac{\partial^2 V}{\partial S \partial v} = \frac{e^{-r\tau}}{S}\left(\frac{\partial^2 c^{\mathrm{SV}}}{\partial x \partial v}\right), \tag{6.68}$$

$$\frac{\partial V}{\partial v} = e^{-r\tau}\frac{\partial c^{\mathrm{SV}}}{\partial v}, \tag{6.69}$$

$$\frac{\partial^2 V}{\partial v^2} = e^{-r\tau}\frac{\partial^2 c^{\mathrm{SV}}}{\partial v^2}. \tag{6.70}$$

Now, using Equations 6.65-6.70, we can express the Heston risk-neutral valuation pde, Equation 6.44, in terms of the new variable $c^{\mathrm{SV}}(\tau, x)$ as

$$e^{-r\tau}\bigg(rc^{\mathrm{SV}} - \frac{\partial c^{\mathrm{SV}}}{\partial \tau} - r\frac{\partial c^{\mathrm{SV}}}{\partial x} + \frac{1}{2}vS^2\bigg(-\frac{1}{S^2}\frac{\partial c^{\mathrm{SV}}}{\partial x} + \frac{1}{S^2}\frac{\partial^2 c^{\mathrm{SV}}}{\partial x^2}\bigg) + \rho\sigma_v vS\bigg(-\frac{1}{S}\frac{\partial^2 c^{\mathrm{SV}}}{\partial x \partial v}\bigg) + \frac{1}{2}\sigma_v^2 v\frac{\partial^2 c^{\mathrm{SV}}}{\partial v^2} + r\frac{\partial c^{\mathrm{SV}}}{\partial x} - rc^{\mathrm{SV}}\bigg) = e^{-r\tau}\kappa(v-\theta)\frac{\partial c^{\mathrm{SV}}}{\partial v}. \tag{6.71}$$

After some simplification, we arrive at the Heston risk-neutral valuation pde in terms of the new variables x and τ:

$$-\frac{\partial c^{\mathrm{SV}}}{\partial \tau} + \frac{1}{2}v\frac{\partial^2 c^{\mathrm{SV}}}{\partial x^2} - \frac{1}{2}v\frac{\partial c^{\mathrm{SV}}}{\partial x} + \frac{1}{2}\sigma_v^2 v\frac{\partial^2 c^{\mathrm{SV}}}{\partial v^2} + \rho\sigma_v v\frac{\partial^2 c^{\mathrm{SV}}}{\partial x \partial v} - \kappa(v-\theta)\frac{\partial c^{\mathrm{SV}}}{\partial v} = 0. \tag{6.72}$$

□

The solution to pde given in Equation 6.72 is found through the Propostions 6.6-6.10. A proof for each of these propositions will now be given. We will begin by considering Proposition 6.6:

Proof. Let us assume the solution of Equation 6.72 will take the same form as the solution of the equivalent Black-Scholes pde for the value of a European call option. That is, we assume a solution of the form:

$$c^{\mathrm{SV}}(x, v, \tau) = K\big(e^x P_1(x, v, \tau) - P_0(x, v, \tau)\big). \tag{6.73}$$

If the function $c^{\mathrm{SV}}(x, v, \tau)$ does take this form, then the following partial derivatives can be easily found:

$$\frac{\partial c^{\mathrm{SV}}}{\partial \tau} = Ke^x\frac{\partial P_1}{\partial \tau} - K\frac{\partial P_0}{\partial \tau}, \tag{6.74}$$

$$\frac{\partial c^{\mathrm{SV}}}{\partial x} = Ke^x P_1 + Ke^x\frac{\partial P_1}{\partial x} - K\frac{\partial P_0}{\partial x}, \tag{6.75}$$

$$\frac{\partial^2 c^{\mathrm{SV}}}{\partial x^2} = Ke^x P_1 + 2Ke^x\frac{\partial P_1}{\partial x} + Ke^x\frac{\partial^2 P_1}{\partial x^2} - K\frac{\partial^2 P_0}{\partial x^2}, \tag{6.76}$$

$$\frac{\partial c^{\mathrm{SV}}}{\partial v} = Ke^x\frac{\partial P_1}{\partial v} - K\frac{\partial P_0}{\partial v}, \tag{6.77}$$

$$\frac{\partial^2 c^{\mathrm{SV}}}{\partial v^2} = Ke^x\frac{\partial^2 P_1}{\partial v^2} - K\frac{\partial^2 P_0}{\partial v^2}, \tag{6.78}$$

$$\frac{\partial^2 c^{\mathrm{SV}}}{\partial x \partial v} = Ke^x\frac{\partial P_1}{\partial v} + Ke^x\frac{\partial^2 P_1}{\partial x \partial v} - K\frac{\partial^2 P_0}{\partial x \partial v}. \tag{6.79}$$

Substituting these partial derivatives back into Equation 6.72 yields

$$
\begin{aligned}
&-Ke^x\frac{\partial P_1}{\partial \tau}+K\frac{\partial P_0}{\partial \tau}+\frac{1}{2}\left(Ke^xP_1+2Ke^x\frac{\partial P_1}{\partial x}+Ke^x\frac{\partial^2 P_1}{\partial x^2}-K\frac{\partial^2 P_0}{\partial x^2}\right)\\
&\quad-\frac{1}{2}v\left(Ke^xP_1+Ke^x\frac{\partial P_1}{\partial x}-K\frac{\partial P_0}{\partial x}\right)+\frac{1}{2}\sigma_v^2v\left(Ke^x\frac{\partial^2 P_1}{\partial v^2}-K\frac{\partial^2 P_0}{\partial v^2}\right)\\
&\quad+\rho\sigma_vv\left(Ke^x\frac{\partial P_1}{\partial v}+Ke^x\frac{\partial^2 P_1}{\partial x\partial v}-K\frac{\partial^2 P_0}{\partial x\partial v}\right)-\kappa(v-\theta)\left(Ke^x\frac{\partial P_1}{\partial v}-K\frac{\partial P_0}{\partial v}\right)=0,
\end{aligned}
\tag{6.80}
$$

which after collecting terms in P_0 and P_1 becomes:

$$
\begin{aligned}
&Ke^x\left(-\frac{\partial P_1}{\partial \tau}+\frac{1}{2}v\frac{\partial^2 P_1}{\partial x^2}+\frac{1}{2}v\frac{\partial P_1}{\partial x}+\frac{1}{2}\sigma_v^2v\frac{\partial^2 P_1}{\partial v^2}+\rho\sigma_vv\frac{\partial^2 P_1}{\partial x\partial v}+(\kappa\theta-\kappa v+\rho\sigma_vv)\frac{\partial P_1}{\partial v}\right)\\
&-K\left(-\frac{\partial P_0}{\partial \tau}+\frac{1}{2}v\frac{\partial^2 P_0}{\partial x^2}-\frac{1}{2}v\frac{\partial P_0}{\partial x}+\frac{1}{2}\sigma_v^2v\frac{\partial^2 P_0}{\partial v^2}+\rho\sigma_vv\frac{\partial^2 P_0}{\partial x\partial v}+(\kappa\theta-\kappa v)\frac{\partial P_0}{\partial v}\right)=0.
\end{aligned}
\tag{6.81}
$$

For this partial differential equation to hold for all values of τ, x and v, both the expressions inside the brackets above must be equal to zero. This can be expressed succinctly as follows; for both $j=0$ and $j=1$, the following pde must be satisfied:

$$
-\frac{\partial P_j}{\partial \tau}+\frac{1}{2}v\frac{\partial^2 P_j}{\partial x^2}-\left(\frac{1}{2}-j\right)v\frac{\partial P_j}{\partial x}+\frac{1}{2}\sigma_v^2v\frac{\partial^2 P_j}{\partial v^2}+\rho\sigma_vv\frac{\partial^2 P_j}{\partial x\partial v}+(a-b_jv)\frac{\partial P_j}{\partial v}=0, \tag{6.82}
$$

where the terms a and b_j are given by

$$
a = \kappa\theta \tag{6.83}
$$

$$
b_j = \kappa-j\rho\sigma_v. \tag{6.84}
$$

The initial conditions, which are given at maturity for an option, can be constructed as follows; as we approach maturity, the variable $\tau\to 0$ and the variable $x\to \ln(S_T/K)$. But, recalling the first line of Equation 6.15, the payoff function at maturity can be expressed in terms of the variable x as $K\max(e^x-1,0)$. Thus, at maturity, if $x\geq 0$, the option is worth Ke^x-K, whilst if $x<0$ the option is worth zero. Now, as we assumed a solution of the form

$$
c^{\mathrm{SV}}(x,v,\tau)=K\big(e^xP_1(x,v,\tau)-P_0(x,v,\tau)\big), \tag{6.85}
$$

the initial conditions for the functions P_0 and P_1 can be stated as

$$\lim_{\tau \to 0} P_j(x, v, \tau) = \begin{cases} 1 & \text{if } x > 0 \\ 0 & \text{if } x \leq 0 \end{cases} =: \theta(x), \tag{6.86}$$

where $\theta(x)$ is the Heaviside-step function, defined in Section 6.1. □

A proof of Proposition 6.7 is given next:

Proof. Define $\tilde{P}_j(\omega, v, \tau)$ as the Fourier transform of $P_j(\omega, v, \tau)$:

$$\tilde{P}_j(u, v, \tau) = \int_{-\infty}^{\infty} e^{-i\omega x} P_j(x, v, \tau) dx. \tag{6.87}$$

The Fourier transform of the initial conditions are given by

$$\tilde{P}_j(u, v, 0) = \int_{-\infty}^{\infty} e^{-i\omega x} \theta(x) dx = \frac{1}{i\omega} + \pi\delta(\omega), \tag{6.88}$$

where we have used the result for the Fourier transform of the Heaviside-step function given in Equation 6.9.

Taking the Fourier transform of partial differential equations 6.82 and using the rules for taking the Fourier transform of partial derivatives, Equations 6.4-6.6, gives the following partial differential equation for $\tilde{P}_j$, $j = \{0, 1\}$:

$$-\frac{\partial \tilde{P}_j}{\partial \tau} - \frac{1}{2}\omega^2 v \tilde{P}_j - \left(\frac{1}{2} - j\right) i\omega v \tilde{P}_j + \frac{1}{2}\sigma_v^2 v \frac{\partial^2 \tilde{P}_j}{\partial v^2} + \rho\sigma_v i\omega v \frac{\partial \tilde{P}_j}{\partial v} + (a - b_j v)\frac{\partial \tilde{P}_j}{\partial v} = 0. \tag{6.89}$$

If we define the variables α, β and γ by

$$\alpha = -\frac{\omega^2}{2} - \frac{i\omega}{2} + ij\omega, \tag{6.90}$$

$$\beta = \kappa - \rho\sigma_v j - \rho\sigma_v i\omega, \tag{6.91}$$

$$\gamma = \frac{\sigma_v^2}{2}, \tag{6.92}$$

then this partial differential equation can be expressed more succinctly as:

$$v\left(\alpha \tilde{P}_j - \beta \frac{\partial \tilde{P}_j}{\partial v} + \gamma \frac{\partial^2 \tilde{P}_j}{\partial v^2}\right) + a\frac{\partial \tilde{P}_j}{\partial v} - \frac{\partial \tilde{P}_j}{\partial \tau} = 0. \tag{6.93}$$

□

It was then argued in Proposition 6.8 that if a candidate solution was substituted into Equation 6.93, then the two functions introduced within this solution would have to satisfy a given system of ode's. A proof of this fact is given now.

Proof. Assuming a solution of Equation 6.93 of the form,

$$\begin{aligned}\tilde{P}_j(u,v,\tau) &= \exp(C(\omega,\tau)\theta + D(\omega,\tau)v)\tilde{P}_j(\omega,v,0)\\ &= \exp(C(\omega,\tau)\theta + D(\omega,\tau)v)\left(\frac{1}{i\omega} + \pi\delta(\omega)\right),\end{aligned} \tag{6.94}$$

results in the following partial derivatives:

$$\frac{\partial\tilde{P}_j}{\partial\tau} = \left(\theta\frac{\partial C}{\partial\tau} + v\frac{\partial D}{\partial\tau}\right)\tilde{P}_j, \tag{6.95}$$

$$\frac{\partial\tilde{P}_j}{\partial v} = D\tilde{P}_j, \tag{6.96}$$

$$\frac{\partial^2\tilde{P}_j}{\partial v^2} = D^2\tilde{P}_j. \tag{6.97}$$

Substituting these partial derivatives into Equation 6.93 and recalling the fact that $a = \kappa\theta$ will give:

$$v\big(\alpha - \beta D + \gamma D^2\big)\tilde{P}_j + \kappa\theta D\tilde{P}_j - \theta\frac{\partial C}{\partial\tau}\tilde{P}_j - v\frac{\partial D}{\partial\tau}\tilde{P}_j = 0. \tag{6.98}$$

With the differential equation in this form, it is clear that Equation 6.94 will give a solution to the partial differential equation 6.93 if:

$$\begin{aligned}\frac{\partial C}{\partial\tau} &= \kappa D\\ \frac{\partial D}{\partial\tau} &= \alpha - \beta D + \gamma D^2 = \gamma(D - r_+)(D - r_-),\end{aligned} \tag{6.99}$$

where the terms r_+ and r_- are defined as

$$r_\pm = \frac{\beta \pm \sqrt{\beta^2 - 4\alpha\gamma}}{2\gamma} =: \frac{\beta \pm d}{\sigma_v^2}. \tag{6.100}$$

□

A solution to the system of ode's 6.99 is stated in Proposition 6.9. The following proof shows how such a solution can be found.

Proof. Let us begin by finding the solution to the second of the two partial differential equations given in Equation 6.57:

$$\frac{\partial D}{\partial \tau} = \gamma(D - r_+)(D - r_-). \tag{6.101}$$

Since D, r_+ and r_- depend on ω, whilst γ does not, the partial differential equation can be expressed as a separable ordinary differential equation (ode). This equivalent ode can be written as:

$$\frac{1}{(D - r_+)(D - r_-)} dD = \gamma d\tau. \tag{6.102}$$

Integrating both sides of this equation yields

$$\frac{\ln(D - r_-) - \ln(D - r_+)}{r_- - r_+} = \gamma\tau + c_1, \tag{6.103}$$

where c_1 is an arbitrary constant of integration. Rearranging and exponentiating the above equation then gives

$$\begin{aligned} D &= \frac{r_- - r_+ \exp(\gamma(r_- - r_+)\tau)c_2}{1 - \exp(\gamma(r_- - r_+)\tau)c_2} \\ &= r_-\left(\frac{1 - \frac{r_+}{r_-}\exp(-d\tau)c_2}{1 - \exp(-d\tau)c_2}\right), \end{aligned} \tag{6.104}$$

where $c_2 = \exp(c_1(r_- - r_+))$ is just another arbitrary constant. In the second line above, we have used the fact $r_- - r_+ = \frac{\beta - d - \beta - d}{\sigma_v^2} = -\frac{2d}{\sigma_v^2}$ and hence $\gamma(r_- - r_+) = -d$.

The initial condition for this differential equation is $D(\omega, \tau) = 0$ for $\tau = 0$, which implies that the constant c_2 is equal to $\frac{r_-}{r_+}$. The solution of the differential equation for D is then

$$D(\omega, \tau) = r_-\left(\frac{1 - \exp(-d\tau)}{1 - g\exp(-d\tau)}\right), \tag{6.105}$$

where $g := \frac{r_-}{r_+}$.

Let us now proceed to solve the second of these pde's. Using the solution for $D(\omega, \tau)$ just found, the pde for $C(\omega, \tau)$ becomes

$$\frac{\partial C}{\partial \tau} = \kappa D = \kappa r_-\left(\frac{1 - \exp(-d\tau)}{1 - g\exp(-d\tau)}\right) = \frac{\kappa r_-}{1 - g\exp(-d\tau)} - \frac{\kappa r_- \cdot \exp(-d\tau)}{1 - g\exp(-d\tau)}. \tag{6.106}$$

Integrating this equation gives

$$C(\omega,\tau) = \frac{\kappa r_- \ln(e^{d\tau} - g)}{d} - \frac{\kappa r_-(\ln(e^{d\tau} - g) - d\tau)}{dg} + c_3. \tag{6.107}$$

The initial condition for this differential equation is $C(\omega,\tau) = 0$ when $\tau = 0$, which implies that

$$c_3 = \frac{\kappa r_- \ln(1-g)(1-g)}{dg}. \tag{6.108}$$

Now, substituting this expression for c_3 into Equation 6.107 and using the fact that $\ln(e^{d\tau} - g) = d\tau + \ln(1 - ge^{-d\tau})$ gives

$$\begin{aligned}
C(\omega,\tau) &= \frac{\kappa r_-(d\tau + \ln(1-ge^{-d\tau}))}{d} - \frac{\kappa r_-(d\tau + \ln(1-ge^{-d\tau} - d\tau))}{d} \\
&\quad + \frac{\kappa r_- \ln(1-g)}{dg} - \frac{\kappa r_- \ln(1-g)}{d} \\
&= \kappa r_- \tau + \frac{\kappa r_-}{d} \ln\left(\frac{1-ge^{-d\tau}}{1-g}\right) - \frac{\kappa r_-}{dg} \ln\left(\frac{1-ge^{-d\tau}}{1-g}\right) \\
&= \kappa r_- \tau + \frac{\kappa}{d}(r_- - r_+) \ln\left(\frac{1-ge^{-d\tau}}{1-g}\right) \\
&= \kappa\left(r_-\tau - \frac{2}{\sigma_v^2} \ln\left(\frac{1-ge^{-d\tau}}{1-g}\right)\right),
\end{aligned} \tag{6.109}$$

where in the third equality we substituted in $g = \frac{r_-}{r_+}$ and in the last equality we, once again, used the identity $(r_- - r_+) = -\frac{2d}{\sigma_v^2}$. □

Finally, Proposition 6.10 uses the inverse Fourier transform to find formulae for P_0 and P_1. A proof of this proposition follows:

Proof. The $\tilde{P}_j$ are given by substituting the expressions for $C(\omega,\tau)$ and $D(\omega,\tau)$ given in Proposition 6.9 back into Equation 6.56. Recall this equation was given by

$$\tilde{P}_j(\omega, v, \tau) = \exp(C_j(\omega,\tau)\theta + D_j(\omega,\tau)v)\left(\frac{1}{i\omega} + \pi\delta(\omega)\right). \tag{6.110}$$

Taking the inverse Fourier transform gives the functions P_j, $j = \{0, 1\}$, as

$$\begin{aligned} P_j(x, v, \tau) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\omega x} \exp(C_j(\omega, \tau)\theta + D_j(\omega, \tau)v) \left(\frac{1}{i\omega} + \pi\delta(\omega) \right) d\omega \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp(i\omega x + C_j(\omega, \tau)\theta + D_j(\omega, \tau)v) \left(\frac{1}{i\omega} + \pi\delta(\omega) \right) d\omega \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \phi(\omega) \left(\frac{1}{i\omega} + \pi\delta(\omega) \right) d\omega. \end{aligned} \tag{6.111}$$

Using an analogous mathematical argument to that used in deriving Equation 6.16 gives us the final form for the functions P_j as

$$\begin{aligned} P_j(x, v, \tau) &= \frac{1}{2} + \frac{1}{\pi} \int_0^{\infty} \mathrm{Re} \left(\frac{\phi(\omega)}{i\omega} \right) \\ &= \frac{1}{2} + \frac{1}{\pi} \int_0^{\infty} \mathrm{Re} \left(\frac{\exp(i\omega x + C_j(\omega, \tau)\theta + D_j(\omega, \tau)v)}{i\omega} \right). \end{aligned} \tag{6.112}$$

□

6.2.3 Numerical evaluation of the complex integral

In the paper in which the Heston model is introduced, Heston [Hes93], the valuation of European options using Fourier techniques is discussed. In the pricing formula given here, however, the variable $C_j(\tau, \omega)$ is defined as

$$C_j(\omega, \tau) = \kappa \left(r_+ \tau - \frac{2}{\sigma_v^2} \ln \left(\frac{e^{d\tau} - g}{1 - g} \right) \right). \tag{6.113}$$

This function appears equivalent to our definition of $C_j(\tau, \omega)$ given in Proposition 6.3. Indeed, if we multiply the argument of the logarithm in the above definition for $C_j(\tau, \omega)$ by $\frac{e^{-d\tau}}{e^{-d\tau}} = 1$ and use the identity $(r_- - r_+) = -\frac{2d}{\sigma_v \sigma_v^2}$, we end up with the definition of $C_j(\tau, \omega)$ which we derived earlier. As they contain a logarithm whose argument is defined in the complex plane, we cannot just simply state that the two definitions of $C_j(\tau, \omega)$ in Equation 6.113 and Proposition 6.3 are completely equivalent. Gatheral [Gat06] comments that the two definitions coincide "only if the imaginary part of the complex logarithm [in Equation 6.113] is chosen so that $C_j(\tau, \omega)$ is continuous with respect to ω". If one takes the principal value of the logarithm in this equation, then $C_j(\tau, \omega)$ will "jump discontinuously each time the imaginary part of the argument of the logarithm crosses the negative real axis".

The usual approach one takes for this problem of calculating logarithms in the complex plane is to ensure we remain on the same Riemann sheet by keeping note of the winding number associated with Equation 6.113. Unfortunately this causes problems for the standard techniques used to numerically integrate the formulae for the P_j, given in terms of C_j and D_j in Proposition 6.3.

By defining the variable $C_j(\tau, \omega)$ as we have done in Proposition 6.3 this is no longer a problem. With this definition, the real part of the argument of the logarithm is positive whenever the imaginary part is zero. Gatheral states that taking the principal value of the logarithm with this definition of $C_j(\tau, \omega)$ "seems to lead to a continuous integrand over the full range of integration", though he states a proof of this fact "remains elusive". Since the publication of Gatheral's textbook, however, a proof of this has been given. See Lord and Kahl [Lor10] for further details.

An interesting point of note is that a fairly robust and accurate numerical integration technique which uses Heston's definition of $C_j(\tau, \omega)$ has been developed. Kahl and Jäckel [Kah05b] outline an approach which allows one to employ a Gauss-Lobatto quadrature scheme together with a new rotation count algorithm to correct the problem of the discontinuous jumps in the complex logarithm.

6.2.4 Semi-analytical formulae for Heston model with jumps

A semi-analytical pricing formula can also be found under the SVJD model. Recall that this model, introduced in Section 5.2, combines the Heston stochastic volatility model with the Merton jump diffusion model, where the asset price can jump discontinuously at certain random times. The occurrence of these jumps is governed by a Poisson process, whilst the magnitude of each such jump is given by a lognormal random variable. Since the jump process is independent of the Heston model, the characteristic function of the transformed stock price under the SVJD model is just given by the product of the characteristic functions from the Heston model and from the Merton jump diffusion model. In Chapter 6 of Gatheral [Gat06], a semi-analytical formula for the price of an option is given in terms of a general characteristic function of an equity process. Therefore, a semi-analytical pricing formula under the SVJD model can be derived with little further effort, in comparison with the Heston model.

6.3 Semi-analytical insurance liabilities under the Heston model

In this section the semi-analytical formulae which we have just derived will be considered together with formulae developed in the textbook Investment Guarantees by Hardy [Har03]. This will give semi-analytical formulae for the liabilities on some simple unit-linked insurance products under the Heston model. To begin this analysis, a summary of how closed-form solutions can be developed for simple unit-linked insurance liabilities under the Black-Scholes model will be given.

6.3.1 Analytical U-L liabilities under Black-Scholes

The two types of unit-linked insurance product for which analytical liability formulae will be developed are variable annuity contracts of the guarantee minimum maturity benefit (GMMB) and guaranteed minimum death benefit (GMDB) class. For a recap of the different types of variable annuity contracts offered on the market see Section 1.3 of this thesis. The following analysis outlining how analytical formulae can be constructed for the liabilities of these products will follow the treatment given in Chapter 8 of Hardy [Har03]. We begin by considering a GMMB contract.

The GMMB is essentially just a put option on the underlying variable annuity fund level or value. This fund value is often known in North America as the segregated fund. We want to value the liability on this product today, or at time $t = 0$. Let the current fund value be denoted by FV_0 and let the guaranteed payment which will be made to the policyholder at time T be denoted by G. The liability from this simple GMMB product at time T is then just $\max(G - FV_T, 0)$. For a standard Canadian policy with this sort of embedded guarantee, G is typically 75 or 100 percent of the initial single premium from the policyholder. If we set m as the monthly management charge deducted from the fund value, then

$$FV_T = FV_0 \frac{S_T}{S_0}(1-m)^T. \tag{6.114}$$

where S_t is the stock index with which this VA contract is linked. If we set $FV_0 = S_0$, then the liability from this product is given as a discounted expectation under the

risk-neutral probability measure Q as

$$\begin{aligned} L_0 &= e^{-rT}\mathbb{E}_Q[\max(G - FV_T, 0)] \\ &= e^{-rT}\mathbb{E}_Q[\max(G - S_T(1-m)^T, 0)]. \end{aligned} \tag{6.115}$$

We can appeal to the Black-Scholes analytic European put option valuation formula to determine L_0 analytically. Replacing S_0 by $S_0(1-m)^T$ in the standard put option valuation formula, to account for the monthly management charges, gives

$$L_0 = Ge^{-rT}\Phi(-d_2) - S_0(1-m)^T\Phi(-d_1), \tag{6.116}$$

where

$$d_1 = \frac{\log(S_0(1-m)^T/G) + (r + \sigma^2/2)T}{\sigma\sqrt{T}} \tag{6.117}$$

and $d_2 = d_1 - \sigma\sqrt{T}$. Note that this is just the well-known adaption of the Black-Scholes formula to account for dividends, except here the dividend represents the monthly charge. This liability valuation formula does not take into account either mortality or policyholder lapses. The risk that a greater than expected number of policyholders survive until maturity can be hedged against by diversification. By selling a sufficiently large number of policies, the rates of mortality will be known with increasing accuracy. This gives some justification to this deterministic assumption. Often the lapse risk can also be treated as diversifiable, however it is perhaps unrealistic to assume lapse rates are uncorrelated with the fund value level. To capture this dependence would require a model for the policyholder lapsation rate in terms of fund level. For the moment we will follow the derivation of Hardy and assume that policyholder lapse rates can be treated similarly to mortality rates. That is, we assume they are diversifiable and independent of G under the probability measure Q. The analytic liability formula is then ${}_Tp_x^\tau L_0$, where ${}_Tp_x^\tau$ is the probability that the contract will still be in force at maturity.

We now reproduce results given by Hardy for this GMMB contract under the Black-Scholes model. In Table 6.1 we calculate the liability value L_0 for a 50 year old annuitant holding this GMMB product and assume that the mortality and lapses follow those of the double decrement table in Appendix A of Hardy. The annualised

Liability est.'s	Term to maturity, T		
Guarantee	5	10	20
60	0.549	0.604	0.217
80	2.333	1.696	0.473
100	5.866	3.423	0.826
120	11.099	5.725	1.262
Delta est.'s	Term to maturity, T		
Guarantee	5	10	20
60	-0.025	-0.017	-0.004
80	-0.082	-0.040	-0.008
100	-0.163	-0.067	-0.012
120	-0.248	-0.096	-0.016
Gamma est.'s	Term to maturity, T		
Guarantee	5	10	20
60	1.18×10^{-3}	5.55×10^{-4}	9.39×10^{-5}
80	2.88×10^{-3}	1.03×10^{-3}	1.54×10^{-4}
100	4.30×10^{-3}	1.44×10^{-3}	2.09×10^{-4}
120	4.97×10^{-3}	1.74×10^{-3}	2.57×10^{-4}
Vega est.'s	Term to maturity, T		
Guarantee	5	10	20
60	11.84	11.01	3.76
80	28.75	20.54	6.14
100	43.03	28.89	8.37
120	49.73	34.80	10.30
${}_Tp_x^\tau$	0.65520	0.42247	0.15972

Table 6.1: Analytical GMMB liability, delta, gamma and vega values under the Black-Scholes model, reproduced from Hardy [Har03]. Note that this is the corrected table from the errata list for this textbook (which is available to view online).

volatility in this example is $\sigma = 20\%$ and the risk-free rate of interest is $r = 6\%$. The initial policy premium is £100. Note that the values that appear in Table 6.1 do not match those in the corresponding table in Hardy. However, if one consults the official errata list for this textbook, available online, then the updated correct values are consistent with those quoted here in the thesis. Also given in this table are some analytical sensitivities of L_0.

The second class of VA product for which analytical liability valuations can be found is the GMDB. The only difference between the GMDB contract and the GMMB which was just considered is that the maturity date is now contingent on the death of the policyholder, as opposed to on his or her survival. For a GMDB, T is now a random variable representing the future lifetime of the policyholder. Let ${}_Tp_x^\tau$ represent the double decrement survival probability, as in the GMMB formula,

and let $\mu_{x,t}^{(d)}$ denote the force of mortality at time t for a life aged x at time $t = 0$. Discretising the term of the contract into n time units, the liability for this GMDB contract can be expressed as

$$\mathbb{E}_T[L_0(T)] = \int_0^n L_0(t)_t p_x^\tau \mu_{x,t}^{(d)} dt, \tag{6.118}$$

where $\mathbb{E}_T[\cdot]$ denotes an expectation over the random variable T. This integral can be evaluated numerically. One such approximation is given by

$$\mathbb{E}_T[L_0(T)] \approx \sum_{t=1}^{n} L_0(t)_{t-1} p_x^\tau \, {}_1q_{x,t-1}^d, \tag{6.119}$$

for some suitably small timestep. In the above numerical approximation ${}_{t-1}p_x^\tau$ is the probability of survival for $t-1$ time units and ${}_1q_{x,t-1}^d$ is the probability that the policyholder dies in the interval $t-1$ to t, conditional on them surviving for $t-1$ time units.

In Table 6.2, analytical liabilities and sensitivities (using this numerical approximation to the integral) are given for this GMDB contract with monthly timesteps. The policyholder is again assumed to be a 50 year-old male and the annualised volatility and risk-free rate of interest will be the same as for the GMMB example just considered. The decrement rates are those found in Appendix A of Hardy [Har03]. The liability values in Table 6.2 are again not the same as those in the corresponding table in Hardy. However, they do correspond closely to the corrected table given by Hardy in the list of errata for her textbook. The small discrepancy which remains is down to different interpretations of the mortality data given in the appendix of Hardy.

One can see from Tables 6.1 and 6.2 that the liabilities from the GMDB contract are significantly smaller than those from the GMMB. This is expected, since the mortality rates for a 50 year-old male are fairly low. The equations for the liability on a GMDB can be easily adapted to accommodate a more complex death benefit. For example, a common feature with VA contracts is to have a death benefit guarantee which increases at a compound rate. If we were to set this compound rate at a 5% increase per year, then the guarantee G would now be given as a function of T, the month in which the policyholder dies, as $G_T = G_0 \cdot 1.05^{T/12}$. Results of the

Liability est.'s	Term to maturity, T		
Guarantee	5	10	20
60	0.0061	0.0303	0.0939
80	0.0389	0.1179	0.2718
100	0.1379	0.3117	0.5977
120	0.3294	0.6353	1.0905
Delta est.'s	Term to maturity, T		
Guarantee	5	10	20
60	-0.0004	-0.0012	-0.0026
80	-0.0019	-0.0041	-0.0071
100	-0.0060	-0.0100	-0.0147
120	-0.0104	-0.0162	-0.0228
Gamma est.'s	Term to maturity, T		
Guarantee	5	10	20
60	2.28×10^{-5}	5.42×10^{-5}	9.29×10^{-5}
80	1.05×10^{-4}	1.70×10^{-4}	2.37×10^{-4}
100	2.60×10^{-4}	3.54×10^{-4}	4.47×10^{-4}
120	2.01×10^{-4}	3.13×10^{-4}	4.26×10^{-4}
Vega est.'s	Term to maturity, T		
Guarantee	5	10	20
60	0.150	0.617	1.734
80	0.525	1.482	3.404
100	0.919	2.296	4.949
120	0.948	2.581	5.817

Table 6.2: Analytical GMDB liab. values under the Black-Scholes model, reproduced from Hardy [Har03]. Note that there is a corrected table to that which is given in the original textbook in the errata list for this textbook (which is available online).

liability for a GMDB with a compounding increase in the guarantee level under the Black-Scholes model are given by Hardy.

Further to the GMMB and GMDB products, Hardy also derives a formula for the liability of a guaranteed minimum accumulation benefit (GMAB) type of variable annuity contract under the Black-Scholes model. This product would also yield a semi-analytical solution under more sophisticated economic models, in the same manner as will discussed in the next section. However, we will not discuss this type of VA product further in the thesis.

6.3.2 Analytical U-L liabilities under a Heston model

Having given an outline on how to calculate the analytical liabilities for simple GMMB and GMDB variable annuity contracts under the Black-Scholes model, the

aim of this section will be to extend these results to the Heston stochastic volatility and stochastic volatility jump diffusion (SVJD) models, introduced in Sections 5.1 and 5.2. In Sections 6.1 and 6.2 the semi-analytical formula for a European option under the Heston model was derived. As the liability values given in the last section are just simple extensions of the Black-Scholes analytic European put option formula, analogous semi-analytical liability formulae can be readily found. To the knowledge of this author, semi-analytical liability formulae for simple variable annuity contracts under the Heston or SVJD models have not previously been given.

For comparison with the results for the two simple VA contracts under the Black-Scholes model, let us set-up the SVJD model in the following way: Firstly, we set the mean reversion level and initial variance by $\theta = V_0 = 0.04$, so that the variance reverts to the level that was assumed in the Black-Scholes example in the long-run. Also, let $r = 6\%$, as it was for the Black-Scholes model. The mean-reversion speed, κ, will be set at 3.00 and the volatility of the variance process, ϵ, will be set as 0.4. The correlation between the equity and variance process risk-drivers, ρ, will be set at -0.80. For the independent jump process, we will fix $\lambda = 0.1$, $\mu_S = -0.15$ and $\sigma_S = 0.2$. This means that the SVJD set-up is as consistent as possible with the Black-Scholes example from before, except we have allowed the volatility to be stochastic (with reversion level equal to the constant volatility of the Black-Scholes model) and have allowed random discontinuous, lognormally distributed jumps in the equity price process, which will occur on average once every 10 years.

In Table 6.3 the liabilities for the different terms to maturity and guarantee levels under this SVJD model are given. In comparison with Table 6.1, we can see that by allowing the volatility to be stochastic and the stock prices to have discontinuous jumps, the insurance liabilities found for these simple guarantees are quite different. The change in the liabilities that we see from considering the SVJD model over the Black-Scholes model appears to be more significant for guarantees which are far out-of-the-money. The semi-analytical values of the liability delta and gamma sensitivities from considering the SVJD model are also given here. These liability sensitivity values also differ from the corresponding Black-Scholes values. The vega Greek here is the sensitivity with respect to $\sigma_0 = \sqrt{V_0}$. That is, the sensitivity with respect to the initial level of the stochastic volatility. Analytic values can also be found for the sensitivities with respect to other parameters of the Heston model.

Liability est.'s	Term to maturity, T		
Guarantee	5	10	20
60	1.030	0.909	0.298
80	2.998	2.095	0.581
100	6.406	3.823	0.948
120	11.308	6.056	1.389
Delta est.'s	Term to maturity, T		
Guarantee	5	10	20
60	-0.030	-0.018	-0.004
80	-0.076	-0.038	-0.007
100	-0.144	-0.061	-0.011
120	-0.224	-0.088	-0.015
Gamma est.'s	Term to maturity, T		
Guarantee	5	10	20
60	1.04×10^{-3}	4.83×10^{-4}	8.46×10^{-5}
80	2.34×10^{-3}	8.75×10^{-4}	1.35×10^{-4}
100	3.78×10^{-3}	1.27×10^{-3}	1.85×10^{-4}
120	4.91×10^{-3}	1.61×10^{-3}	2.31×10^{-4}
Vega est.'s	Term to maturity, T		
Guarantee	5	10	20
60	0.863	0.364	0.061
80	1.745	0.623	0.094
100	2.570	0.861	0.125
120	3.063	1.047	0.153
${}_Tp_x^\tau$	0.65520	0.42247	0.15972

Table 6.3: Semi-analytical GMMB liability, delta, gamma and vega values under the SVJD model with parameters chosen to aid comparison with the Black-Scholes tests. These SVJD parameters are given in the main text of Section 6.3.2.

The liabilities on the GMDB variable annuity contract under the SVJD model are displayed in Table 6.4. If we compare these values with the corresponding values under the Black-Scholes model, given in Table 6.2, we see that the liabilities under this type of product are also quite strongly dependent on the model choice for the equity returns. The differences in the liabilities resulting from using the SVJD model rather than the Black-Scholes, again, seem most significant for guarantees which are far out-of-the-money.

The liability sensitivities for the GMDB contract under the SVJD model are also given in Table 6.4. Using the SVJD model for the underlying equity index returns also gives significantly different values for these sensitivities. However, here the differences do not appear as significant for far out-of-the-money guarantees, in comparison with the case of the guarantee being in-the-money.

Liability est.'s	Term to maturity, T		
Guarantee	5	10	20
60	0.0149	0.0541	0.1446
80	0.0548	0.1536	0.3427
100	0.1475	0.3397	0.6637
120	0.3253	0.6435	1.1350
Delta est.'s	Term to maturity, T		
Guarantee	5	10	20
60	-0.0005	-0.0015	-0.0029
80	-0.0019	-0.0040	-0.0068
100	-0.0051	-0.0087	-0.0131
120	-0.0098	-0.0151	-0.0212
Gamma est.'s	Term to maturity, T		
Guarantee	5	10	20
60	2.38×10^{-5}	5.13×10^{-5}	8.57×10^{-5}
80	8.21×10^{-5}	1.37×10^{-4}	1.95×10^{-4}
100	2.35×10^{-4}	3.17×10^{-4}	3.99×10^{-4}
120	2.34×10^{-4}	3.40×10^{-4}	4.43×10^{-4}
Vega est.'s	Term to maturity, T		
Guarantee	5	10	20
60	0.0230	0.0445	0.0697
80	0.0670	0.1065	0.1474
100	0.1329	0.1890	0.2444
120	0.1159	0.1837	0.2513

Table 6.4: Semi-analytical GMDB liability, delta, gamma and vega values under the SVJD model with parameters chosen to aid comparison with the Black-Scholes tests. These SVJD parameters are given in the main text of Section 6.3.2.

Overall, comparing the liabilities and sensitivities for both types of guarantee, we see that the semi-analytical values can be significantly different under the consideration of different models for the underlying equity index. Furthermore, this example set the mean-reversion level and initial variance in the SVJD model to be consistent with the constant volatility of the Black-Scholes model. With these SVJD parameters calibrated to market data, the differences in liabilities and sensitivities could be even larger, compared to a market-calibrated Black-Scholes model.

Earlier in the thesis we argued that the SVJD model will give a distribution of equity returns with a much heavier tail than the Black-Scholes model gives, which is much more consistent with the distribution of returns from real-life equity indices. See Duffie and Pan [Duff97] for more discussion on the merits of the SVJD model over Black-Scholes in terms of their ability to realistically fit to market data.

Liability est.'s	Term to maturity, T		
Guarantee	5	10	20
60	0.403	0.640	0.413
80	2.095	2.143	1.000
100	6.269	4.842	1.849
120	13.328	8.673	2.917
Delta est.'s	Term to maturity, T		
Guarantee	5	10	20
60	-0.0207	-0.0223	-0.0093
80	-0.0900	-0.0621	-0.0190
100	-0.2195	-0.1164	-0.0298
120	-0.3681	-0.1726	-0.0405
Gamma est.'s	Term to maturity, T		
Guarantee	5	10	20
60	1.16×10^{-3}	8.80×10^{-4}	2.52×10^{-4}
80	4.09×10^{-3}	1.96×10^{-3}	4.22×10^{-4}
100	7.46×10^{-3}	2.86×10^{-3}	5.47×10^{-4}
120	8.26×10^{-3}	3.20×10^{-3}	6.12×10^{-4}
Vega est.'s	Term to maturity, T		
Guarantee	5	10	20
60	0.347	0.234	0.063
80	1.078	0.484	0.101
100	1.740	0.665	0.127
120	1.699	0.701	0.138
${}_Tp_x^\tau$	0.65520	0.42247	0.15972

Table 6.5: Semi-analytical GMMB liability, delta, gamma and vega values under the SVJD model, with parameters taken from Broadie and Kaya [Bro04] and quoted in the main text of Section 6.3.2.

To finish the analysis of this chapter, let us consider liabilities and sensitivities under the parameters given for the SVJD model by Broadie and Kaya [Bro04] and originally quoted by Duffie, Pan and Singleton [Duff00]. The variance process parameters are $\kappa = 3.99$, $\theta = 0.014$, $\sigma_V = 0.27$ and $V_0 = 0.008836$. The correlation between equity and variance processes is $\rho = -0.79$ and the risk-free rate $r = 3.19\%$. The arrival of the jumps follows an independent Poisson process with parameter $\lambda = 0.11$ and the size of the jumps follow a lognormal distribution with $\mu_S = -0.1389$ and $\sigma_S = 0.15$, resulting in an average decline in stock price from a jump of $\mu = -0.12$.

The liability values and delta, gamma and vega sensitivities for the GMMB variable annuity contract under the SVJD model, calibrated as above, are given in Table 6.5. The corresponding results for the GMDB contract under the market-calibrated SVJD model are given in Table 6.6.

Liability est.'s	Term to maturity, T		
Guarantee	5	10	20
60	0.004	0.026	0.119
80	0.030	0.116	0.371
100	0.124	0.340	0.850
120	0.358	0.772	1.620
Delta est.'s	Term to maturity, T		
Guarantee	5	10	20
60	-0.0003	-0.0012	-0.0037
80	-0.0017	-0.0046	-0.0104
100	-0.0066	-0.0127	-0.0226
120	-0.0137	-0.0232	-0.0371
Gamma est.'s	Term to maturity, T		
Guarantee	5	10	20
60	1.77×10^{-5}	5.82×10^{-5}	1.39×10^{-4}
80	1.02×10^{-4}	2.10×10^{-4}	3.62×10^{-4}
100	4.48×10^{-4}	6.21×10^{-4}	8.29×10^{-4}
120	2.44×10^{-4}	4.37×10^{-4}	6.70×10^{-4}
Vega est.'s	Term to maturity, T		
Guarantee	5	10	20
60	0.006	0.017	0.038
80	0.031	0.058	0.095
100	0.091	0.013	0.179
120	0.042	0.084	0.135

Table 6.6: Semi-analytical GMDB liability, delta, gamma and vega values under the SVJD model, with parameters taken from Broadie and Kaya [Bro04] and quoted in the main text of Section 6.3.2.

6.3.3 Conclusion

In conclusion to this chapter, we remark that for simple unit-linked insurance guarantees the insurance liabilities can be found analytically under both the Black-Scholes and SVJD models. Some tests showed that the liabilities and sensitivities can be quite different, even if we just allow the volatility to follow a mean reverting stochastic process around the constant volatility level assumed in the Black-Scholes model and allow for the rare possibility of a discontinuous jump in the equity index level. If the Black-Scholes model and the SVJD model were individually calibrated to market data, the differences in the sensitivity estimates could be even greater.

It is important to be aware, however, that these analytic solutions can only be found for the simplest of insurance guarantees. Once the various complexities of typical insurance liabilities are taken into account, simple closed-form solutions are generally

not available. In the remainder of Part II of the thesis we will investigate how Monte Carlo simulation estimators can be constructed for the sensitivities of more complex insurance products under more sophisticated market models. We pause to outline a possible practical benefit of these semi-analytical liability formulae.

In Section 1.4 a brief introduction was given to the Monte Carlo variance reduction technique of employing a control variate. The idea of this approach is that if there exists some random quantity with known expected value under the financial model being considered, we can often use this to improve the accuracy of the unknown quantity which is being estimated. The greater the correlation between the quantity with known expected value and the quantity which we are attempting to estimate, the more successful this approach will be in improving the efficiency of the estimator. Now, although these guarantees which yield semi-analytical liability values are simplistic, they are likely to have a significant positive correlation with more realistic, complex insurance liabilities which must be valued and hedged using simulation. Therefore, these semi-analytical solutions, which can be calculated very efficiently, appear to be excellent candidate control variates for improving the efficiency of Monte Carlo estimators of complex insurance liabilities and sensitivities. Testing these potential control variates for the valuation of more complex liabilities, such as those considered in Chapter 4, is a line of promising future research. If these simple guarantees do prove to be successful control variates, the practical benefits in valuation and hedging could be great.

The analysis given in Section 6.3 could be extended to value these simple guarantees under an even more sophisticated model. Scott [Sco97] gives a framework for calculating semi-analytical formulae for the price of stock options under a Heston-CIR model. The Heston-CIR model will be introduced in Section 8.3 of this thesis. He also shows that the option prices still yield semi-analytical valuations if an independent jump process is incorporated into this sophisticated model of equity returns. Following this framework, the discussion given in Section 6.3 could then be used to give semi-analytical valuation formulae for the simple guarantees under a Heston-CIR model featuring random discontinuous jumps for the equity returns process. This increase in model sophistication under which semi-analytical liability valuation is available further increases the likelihood of these simple guarantees acting as successful control variates for more complex insurance guarantees.

Chapter 7

Option sensitivity estimators using Monte Carlo simulation

7.1 Option Price Sensitivity Estimators

The liability from a unit-linked insurance contract takes a similar form to the payoff of an option, as was discussed in Section 3.1. Thus, in attempting to develop efficient Monte Carlo methods for determining the sensitivities of insurance liabilities it would seem sensible to study equivalent estimators for option price sensitivities which have been developed by researchers in recent years. We now summarise three classes of approach for calculating the sensitivities of options by simulation. Firstly, we introduce a simple method commonly adopted by insurers to estimate VA liability Greeks; the "bump and revalue" approach.

7.1.1 Bump and revalue approach

This term refers to the concept of simulating the cost of the option under some base scenarios for key risk-drivers and then again under 'bumped scenarios', that is with the sensitivity parameter increased by some small perturbation, say $\Delta\theta$. This sensitivity parameter could be the current equity index level, say. This is just a forward difference estimate in the sensitivity parameter. If the function $Y(\theta)$ gives the discounted payoff of an option, then the price of this option is then given by $\alpha(\theta) = \mathbb{E}[Y(\theta)]$ with respect to a pricing measure. Now let $Y_1(\theta), \ldots, Y_n(\theta)$ represent the discounted option payoff along simulation paths $1, \ldots, n$, the estimator of the first-order sensitivity would then be given by

$$\hat{\Delta}^{\mathrm{B}} = \frac{\bar{Y}(\theta + \Delta\theta) - \bar{Y}(\theta)}{\Delta\theta}, \tag{7.1}$$

for the chosen perturbation size $\Delta\theta$, where $\bar{Y}(\theta)$ is the average of $Y_1(\theta), \ldots, Y_n(\theta)$.

The expectation of this estimator is given by

$$\mathbb{E}\big[\hat{\Delta}^{\mathrm{B}}\big] = \frac{\alpha(\theta + \Delta\theta) - \alpha(\theta)}{\Delta\theta}. \tag{7.2}$$

The variance will be significantly reduced if the same random number stream is used in calculating the 'bumped' and 'base' option prices, as opposed to using independent random number streams. In practice it is also usually beneficial in terms of accuracy to simulate a downward perturbed path as well as an upward perturbed path and calculate a central difference estimator.

The bump and revalue approach can incur fairly large sampling errors, particularly with discontinuous payoff functions and second-order sensitivity estimates. See Glasserman [Gla03] for some numerical tests of calculating the gamma of a European option with this type of estimator. Alternative approaches for estimating sensitivities are, therefore, required by practitioners looking to construct a hedging strategy for their insurance liabilities. Glasserman introduces two more sophisticated general approaches to determining the sensitivities of option prices.

7.1.2 Pathwise estimator

The first of these approaches is known as the pathwise derivative method, introduced in a financial context by Broadie and Glasserman [Bro97]. The estimator of the first-order sensitivity with respect to θ is given as follows: We can find the derivative of $\alpha(\theta) = \mathbb{E}[Y(\theta)]$ analytically along each simulation path using

$$Y'(\theta) = \lim_{h \to 0} \frac{Y(\theta + h) - Y(\theta)}{h}. \tag{7.3}$$

If the interchanging of differentiation and taking expectations is justified, that is if

$$\mathbb{E}\left[\frac{d}{d\theta} Y(\theta)\right] = \frac{d}{d\theta}\mathbb{E}[Y(\theta)], \tag{7.4}$$

then $\frac{1}{n}\sum_{i=1}^{n} Y_i'(\theta)$ is an unbiased estimator of $\alpha'(\theta)$. That is, the average of these analytic derivatives calculated along each simulation path gives us an unbiased estimate of the sensitivity. This is what we refer to as the pathwise derivative estimator.

Example 7.1. As an example of this method, consider the challenge of estimating the delta of a call option under the Black-Scholes model (that is, the sensitivity of the option price with respect to S_0 the initial underlying asset value). This can, of course, be found analytically, but considering this problem helps illustrate the

method. The discounted payoff of a call option is given by

$$Y = e^{-rT} \max(S_T - K, 0), \tag{7.5}$$

where r, K, T and S_T are the risk-free rate, strike price, time to maturity and equity price at maturity, respectively. Under Geometric Brownian Motion

$$S_T = S_0 e^{(r - \frac{1}{2}\sigma^2)T + \sigma\sqrt{T}Z}, \tag{7.6}$$

where Z is a standard normal variate. Applying the chain rule to differentiate Y with respect to S_0 with all other parameters held fixed gives

$$\frac{dY}{dS_0} = \frac{dY}{dS_T}\frac{dS_T}{dS_0}. \tag{7.7}$$

With the term $\frac{dY}{dS_T}$ the derivative fails to exist at the strike price, however the event $\{S_T = K\}$ has measure zero and hence $\frac{dY}{dS_T} = e^{-rT} I(S_T > K)$ almost surely, where $I(A)$ represents the indicator function of event A. For the second-term, Equation 7.6 gives $\frac{dS_T}{dS_0} = \frac{S_T}{S_0}$. Thus, the pathwise estimator for the delta of the call option is

$$\frac{dY}{dS_0} = e^{-rT} I(S_T > K)\frac{S_T}{S_0}. \tag{7.8}$$

If we wish to find the Black-Scholes gamma, i.e., sensitivity of the delta to the initial asset price, we have to differentiate $W = e^{-rT} I(S_T > K)$ with respect to S_0, which is equivalent to estimating the delta of a digital option. Here W is differentiable with respect to S_0 with probability one, and takes the value zero. However, in this case

$$0 = \mathbb{E}\left[\frac{dW}{dS_0}\right] \neq \frac{d}{dS_0}\mathbb{E}[W] \tag{7.9}$$

and the pathwise estimator is inapplicable. Indeed, the method is generally inapplicable when the payoff is discontinuous or in estimating second-order derivatives. The change in $\mathbb{E}[W]$ with a change in S_0 is explained by the fact that it could cause S_T to become in-the-money. Glasserman [Gla03] gives some technical conditions for when the pathwise method can be applied. However, a less rigorous 'rule of thumb' is that it can be used when the payoff function is continuous in the parameter of interest.

7.1.3 Likelihood ratio method (LRM)

The second more sophisticated approach to estimating option sensitivities introduced by Glasserman [Gla03] is known as the likelihood ratio method (LRM). The LRM approach relies on differentiating probability densities rather than payoff functions. As such it does not require smoothness in the payoff function, as was required in the pathwise derivative method. An overview of this approach for estimating option price sensitivities will now be given.

Suppose we have a discounted payoff Y expressed as a function $f(\mathbf{X})$, where $\mathbf{X}$ is a m-dimensional vector of different asset prices (or alternatively, one asset price at multiple valuation dates). Then, assuming that $\mathbf{X}$ has a probability density g with parameter θ, taking the expected discounted payoff with respect to this density gives

$$\mathbb{E}[Y] = \mathbb{E}[f(\mathbf{X}(\theta))] = \int_{\mathbb{R}^m} f(\mathbf{x}) g_\theta(\mathbf{x}) d\mathbf{x}. \tag{7.10}$$

Now, similar to the pathwise derivative approach, we assume the order of differentiation and integration can be interchanged. Here, however, this is not such a strong assumption, as typically densities are smooth functions, whereas payoff functions are not. This gives

$$\begin{aligned} \alpha'(\theta) = \frac{d}{d\theta}\mathbb{E}[Y] &= \int_{\mathbb{R}^m} f(\mathbf{x}) \frac{d}{d\theta} g_\theta(\mathbf{x}) d\mathbf{x} \\ &= \int_{\mathbb{R}^m} f(\mathbf{x}) \frac{\frac{d}{d\theta} g_\theta(\mathbf{x})}{g_\theta(\mathbf{x})} g_\theta(\mathbf{x}) d\mathbf{x} \\ &= \mathbb{E}\left[f(\mathbf{X}) \frac{\frac{d}{d\theta} g_\theta(\mathbf{X})}{g_\theta(\mathbf{X})} \right] \\ &= \mathbb{E}\left[f(\mathbf{X}) \frac{d}{d\theta} \ln(g_\theta(\mathbf{X})) \right]. \end{aligned} \tag{7.11}$$

Then, $f(\mathbf{X}) \frac{d}{d\theta} \ln(g_\theta(\mathbf{X}))$ gives the likelihood ratio estimator for the sensitivity with respect to the parameter θ, and this estimator is unbiased. The term $\frac{d}{d\theta} \ln(g_\theta(\mathbf{x}))$ is often known in statistics as the "score function" and in this context is referred to as the likelihood ratio weight, since it multiplies the discounted payoff function to give the sensitivity estimator.

The likelihood ratio method will still be applicable and robust in the case of options with discontinuous payoff functions (and in estimating second-order sensitivities) as

this approach looks to differentiate the probability density, rather than the payoff function.

One important point of note is the LRM weight for the sensitivity with respect to θ is independent of the form of the payoff function. This means that the LRM could be applied for an alternative payoff function by simply multiplying this function with the same LRM weight. This generality of the LRM is a feature which makes it appealing as a technique for estimating option sensitivities with different types of payoff in the same simulation run.

Example 7.2. To show how this approach works in practice let us again consider the problem of estimating the delta of a call option under the Black-Scholes model. In this case the discounted payoff function is given by $f(S_T) = e^{-rT}\max(S_T - K, 0)$ and the lognormal density function of S_T is given by

$$g(x) = \frac{1}{x\sigma\sqrt{T}}\phi\left(\frac{\ln\left(\frac{x}{S_0}\right) - (r - \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}}\right), \tag{7.12}$$

where $\phi(\cdot)$ represents the standard normal density function. The score function in this case is:

$$\frac{d}{dS_0}\ln(g(x)) = \frac{\ln\left(\frac{x}{S_0}\right) - (r - \frac{1}{2}\sigma^2)T}{S_0\sigma^2 T}. \tag{7.13}$$

Evaluating this function at S_T and multiplying by the discounted payoff of the option gives an unbiased estimator of the delta

$$e^{-rT}\max(S_T - K, 0)\frac{\ln\left(\frac{S_T}{S_0}\right) - (r - \frac{1}{2}\sigma^2)T}{S_0\sigma^2 T}. \tag{7.14}$$

As we simulate for S_T using the relationship

$$S_T = S_0 e^{(r - \frac{1}{2}\sigma^2)T + \sigma\sqrt{T}Z}, \tag{7.15}$$

this likelihood ratio estimator can be written more simply as

$$e^{-rT}\max(S_T - K, 0) \cdot \frac{Z}{S_0\sigma\sqrt{T}}, \tag{7.16}$$

where the term $Z/(S_0\sigma\sqrt{T})$ is known as the LRM weight.

To obtain the delta sensitivity of a digital option under the Black-Scholes model we

just have to multiply this alternative payoff function with the LRM weight which we have already calculated. This gives

$$e^{-rT} I(S_T > K) \cdot \frac{Z}{S_0 \sigma \sqrt{T}}. \tag{7.17}$$

For reference, the LRM weight for the option gamma, found by an analogous process of differentiating the density $g_\theta(x)$ twice with respect to θ, can be easily shown to be $\frac{Z^2 - Z\sigma\sqrt{T} - 1}{S_0^2 \sigma^2 T}$.

For path-dependent options the LRM weights have a similar structure to the equivalent European option LRM weight for the required sensitivity, except they will only use information along the path up until the first valuation date. The following Example demonstrates this.

Example 7.3. The analysis thus far has studied the LRM for European-type payoffs, however many unit-linked insurance products exhibit some degree of path dependency in their liabilities. With this in mind let us turn our attention to studying how the LRM extends to path-dependent options.

Following the example in Glasserman [Gla03], let us consider estimating the sensitivities of an arithmetic Asian call option with m valuation dates. The payoff on this option at maturity is then given by

$$Y = f(S_{t_1}, \ldots, S_{t_m}) = e^{-rT} \max(\bar{S} - K, 0), \quad \text{where } \bar{S} = \frac{1}{m} \sum_{i=1}^{m} S_{t_i}. \tag{7.18}$$

Using the Markov property of Geometric Brownian motion, the underlying density for the equity asset path can be factorised as

$$g(x_1, \ldots, x_m) = g_1(x_1|S_0) g_2(x_2|x_1) \cdots g_m(x_m|x_{m-1}), \tag{7.19}$$

where $g_j(x_j|x_{j-1})$ is the transition density from time t_{j-1} to t_j, i.e.,

$$g_j(x_j|x_{j-1}) = \frac{1}{x_j \sigma \sqrt{t_j - t_{j-1}}} \phi(\zeta_j(x_j|x_{j-1})), \tag{7.20}$$

$$\zeta_j(x_j|x_{j-1}) = \frac{\log(x_j/x_{j-1}) - (r - \sigma^2/2)(t_j - t_{j-1})}{\sigma \sqrt{t_j - t_{j-1}}}. \tag{7.21}$$

Suppose we wish to get an LRM estimator for the delta of the Asian option. From

the above factorisation it is clear that S_0 is a parameter of the first factor, $g_1(x_1|S_0)$, only. This means we can express the score function, corresponding to the delta sensitivity, as

$$\frac{\partial \log g(S_{t_1}, \ldots, S_{t_m})}{\partial S_0} = \frac{\partial \log g_1(S_{t_1}|S_0)}{\partial S_0} = \frac{\zeta_1(S_{t_1}|S_0)}{S_0\sigma\sqrt{t_1}} = \frac{Z_1}{S_0\sigma\sqrt{t_1}}, \tag{7.22}$$

where $Z_1 = \zeta_1(S_{t_1}|S_0)$ is the Gaussian increment which takes us from time zero to time t_1. Likewise, the LRM estimator of the gamma sensitivity has a score function component which only relies on the equity asset path out to the first valuation date. Thus, the LRM estimator for the delta sensitivity of an Asian option is given by

$$e^{-rT}\max(\bar{S} - K, 0) \cdot \frac{Z_1}{S_0\sigma\sqrt{t_1}}. \tag{7.23}$$

7.1.4 Mixed estimators for second-order sensitivities

In estimating the second-order sensitivities of an option price, for example the gamma sensitivity, one can also combine the pathwise and LRM approaches to create hybrid mixed estimators. This idea is most easily illustrated by means of an example and we will consider the following from Glasserman [Gla03]:

Example 7.4. Let us consider, again, estimating the gamma Greek of a European call option under the Black-Scholes model. By applying the pathwise approach to the LRM delta estimator, we obtain the first of the two mixed gamma estimators:

$$\begin{aligned} \Gamma^{\text{LR-PW}} &= \frac{d}{dS_0}\left(e^{-rT}\max(S_T - K, 0)\frac{Z}{S_0\sigma\sqrt{T}}\right) \\ &= e^{-rT}I(S_T > K)K\frac{Z}{S_0^2\sigma\sqrt{T}}. \end{aligned} \tag{7.24}$$

Alternatively applying the LRM to the pathwise delta estimator gives the PW-LR mixed gamma estimator. The pathwise delta estimator has both functional and distributional dependence on S_0. To capture the distributional dependence, the pathwise estimator is multiplied by the LRM weight. For the functional dependence another pathwise derivative is taken. This gives the second of the two mixed gamma

estimators as:

$$\begin{aligned}\Gamma^{\text{PW-LR}} &= e^{-rT} I(S_T > K) \frac{S_T}{S_0} \cdot \frac{Z}{S_0 \sigma \sqrt{T}} + \frac{d}{dS_0}\left(e^{-rT} I(S_T > K) \frac{S_T}{S_0} \right) \\ &= e^{-rT} I(S_T > K) \frac{S_T}{S_0^2}\left(\frac{Z}{\sigma\sqrt{T}} - 1 \right). \end{aligned} \tag{7.25}$$

These mixed estimators typically give smaller standard errors for the gamma of European options, than the bump and revalue or pure LRM approaches. See Glasserman [Gla03] for a numerical comparison of different estimators for the gamma sensitivity of a European option under the Black-Scholes model.

Essentially, these mixed estimators combine the increased accuracy of the pathwise estimator and the robustness of the LRM estimator to discontinuous payoff functions, to give a highly accurate and reliable approach for estimating second-order sensitivities. In Chapter 8 a mixed estimator will be constructed for the gamma sensitivity of the liability on a stylised unit-linked insurance product. Numerical tests conducted in this chapter will show this is both efficient and robust in estimating this liability sensitivity, in comparison to the bump and revalue and LRM approaches.

7.2 Option sensitivities under the Black-Scholes withdrawals model

In this section the Black-Scholes model will be extended to incorporate periodic fixed withdrawals from the equity asset. This structure is the starting point for modelling GMWB variable annuity products, which we recall entitles the policyholder to periodic withdrawals from some underlying equity-linked fund. As far as this author is aware, this form of model has not been considered elsewhere in the literature. The derivation and analysis of sensitivity estimators within such a framework are, therefore, original to this thesis.

Let us describe the stochastic cashflows of some equity asset in time by S_t. The returns on this asset are governed by the Black-Scholes model, but at each valuation date a fixed withdrawal of w is made from the asset. After a withdrawal is made from the asset, the remaining funds are kept invested in the equity asset until the next valuation date. We note that in this analysis an assumption regarding policyholder behaviour is being made: under a typical insurance contract the policyholder will be

able to choose how much to withdraw at each valuation date, up to some maximum amount. Here it is assumed they will always withdraw the full amount w. However, Milevsky and Salisbury [Mil06] show that, under certain conditions, withdrawing the full amount will be optimal for the policyholder in the sense that it maximises the value of the embedded option. Therefore, in the context of insurance solvency testing, such an assumption regarding policyholder behaviour seems reasonable.

At first this may seem to just be a simple case of the Black-Scholes model with discrete dividends, which is well known to still yield lognormal returns and, hence, give analytical formulae for option prices and Greeks. However, in that instance the dividends must be proportional to the current level of the equity-index, i.e., the dividend will be $0.05S_{t_i}$ at each dividend date t_i for a 5% dividend rate. The case of fixed withdrawals is quite different. Here, no matter what the level of S_t the withdrawal amount is fixed as w. This means that if w is sufficiently large the value of S_t will be likely to diminish fairly rapidly in time and eventually (given a long enough contract term) reach zero. On the other hand, with fixed proportional dividends, the dividend amount would get smaller in proportion with the lower the value of S_t, thus having less of an effect on the underlying dynamics. Indeed, the returns would still be lognormally distributed, which is not the case with the subtraction of fixed withdrawals.

To begin our analysis of this equity asset, let us calculate the expected value of the stock at time T, i.e., we wish to determine $\mathbb{E}[S_T]$. Before we consider the general case of N withdrawal dates, let us go through the case of three withdrawal dates (with one of these dates coinciding with time $t = T$). The diagram given below shows the expected cashflows of this equity asset featuring withdrawals. The asset begins with value S_0, and by the first withdrawal date, $t = \Delta t$, the expected value of this asset will be $S_0 e^{r\Delta T}$. At this point a withdrawal is made resulting in the expected value of the asset being given by $S_0 e^{r\Delta T} - w$. This amount is then expected to accumulate by $e^{r\Delta T}$ until the next periodic withdrawal date, at which point another withdrawal will be deducted from the new accumulated amount. The expected cashflows are shown in full for the case $N = 3$ in Figure 7.1 overleaf.

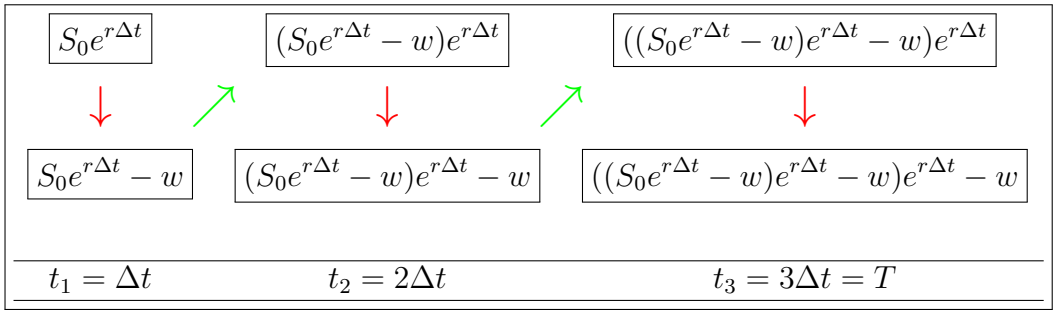

Figure 7.1: Black-Scholes model with constant withdrawals at each timestep: Example of calculating the expected value of S_T, where $T = N\Delta t$ and N = 3. The result of the cashflows in this simple case is $\mathbb{E}[S_T] = ((S_0e^{r\Delta t} - w)e^{r\Delta t} - w)e^{r\Delta t} - w$.

Generalising this to the case of N timesteps and collecting terms, gives us a fairly simple expression for the value of $\mathbb{E}[S_T]$. This can then be simplified further by using the standard formula for the sum of a geometric progression:

$$\begin{aligned}\mathbb{E}\big[S_T\big] &= S_0e^{rN\Delta t} - w\big(1 + e^{r\Delta t} + e^{2r\Delta t} + e^{3r\Delta t} + \cdots + e^{(N-1)r\Delta t}\big) \\ &= S_0e^{rT} - w\left(\frac{1 - e^{rT}}{1 - e^{r\Delta t}}\right).\end{aligned} \tag{7.26}$$

When valuing and determining the sensitivities of a basic European option written on this equity asset in Section 7.3, we shall set the strike as being at the money, that is $K = \mathbb{E}[S_T]$, with $\mathbb{E}[S_T]$ given by the formula above.

In calculating a pathwise derivative estimate for the delta sensitivity of an option written on this equity asset, it is critical that we can find an expression for S_T in terms of the shocks along a particular simulation path. With the basic Black-Scholes model under the risk-neutral measure, the equity price at maturity is given by $S_T = S_0 \exp((r - \sigma^2/2)T + \sigma\sqrt{T}Z)$, where Z is the shock which takes the asset from time zero out to time T along a particular path. In this case things are not so simple, as there are many intermittent timesteps up until maturity at time T, at each of which a withdrawal is subtracted from the equity asset fund level. One can easily simulate under the Black-Scholes model with many intermittent timesteps, indeed this is required for some path-dependent options such as an Asian option. For such path-dependent assets one would simply find the terminal equity price as

$$S_T = S_0 \exp((r - \sigma^2/2)T + \sigma\sqrt{\frac{T}{N}}(Z_{t_1} + \cdots + Z_{t_N}), \tag{7.27}$$

where $Z_{t_1}, \ldots, Z_{t_N}$ are the respective shocks over each of the N intermittent timesteps until maturity. For the remainder of this chapter, Z_n will be used to represent Z_{t_n} for ease of notation. Taking into account the withdrawals from the fund makes the situation somewhat more complicated, though.

With some thought this problem can be overcome. By considering the stochastic cashflows which arise on this equity asset along each simulation path, it is possible to find an expression for S_T in terms of the N shocks which govern the returns of the equity asset in between each of the withdrawal dates. This expression will now be derived, but we first look at the cashflows over the first couple of timesteps for intuition. These cashflows are shown schematically in Figure 7.2, given on the next page.

The equity asset begins with a value of S_0 and by the first withdrawal date, its value will be given by $S_0 e^{(r-\frac{\sigma^2}{2})\Delta t+\sigma\sqrt{\Delta t}Z_1}$, where Z_1 is a standard random normal variate representing the shock out to the first timestep. At this point a withdrawal of w is taken from the equity asset. The remainder of the equity asset after this withdrawal is then kept invested and by the second withdrawal date this will have increased by a factor of $e^{(r-\frac{\sigma^2}{2})\Delta t+\sigma\sqrt{\Delta t}Z_2}$. (This factor may of course be less than one, indicating a depreciation of the equity fund level from the previous timestep). The new fund level will then have a withdrawal of w subtracted.

This process continues throughout all N timesteps until the final withdrawal at time T. The value of the equity asset at time T, inclusive of the final withdrawal, is then the value of S_T which is required in the pathwise approach for estimating the delta sensitivity of an option.

In practice a $\max(\cdot, 0)$ function must also be applied to ensure that the value of the asset cannot go negative, but this is not shown in the diagram of the cashflows for clarity of illustration.

Having gained some intuition about the cashflows under the Black-Scholes withdrawals model, we will now turn our attention to deriving a general formula for the value of S_T. This formula will then be used to find the pathwise estimator for the delta of a European option written on this product.

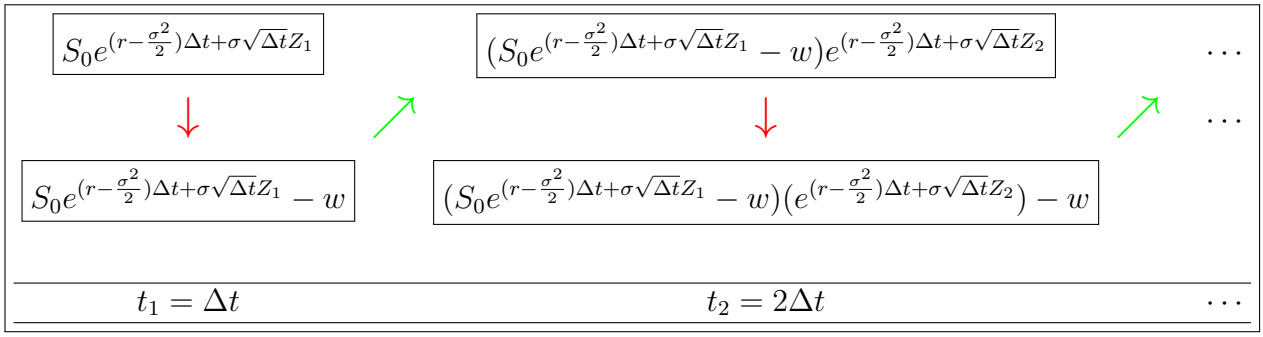

Figure 7.2: Black-Scholes model with constant withdrawals at each timestep: Example of calculating the value of the equity asset at timesteps $t_1 = \Delta t$ and $t_2 = 2\Delta t$ under a simulation path which uses shocks $Z_1, Z_2, \ldots$. The value at future timesteps can be found by extending this approach.

Keeping track of all the exponential terms and the w terms which they multiply, one can find an expression for the equity asset at time T as:

$$\begin{aligned}
S_T = &\left[S_0\left(\exp\left((r-\frac{\sigma^2}{2})\Delta t+\sigma\sqrt{\Delta t}Z_1\right)\right)-w\right)\\
&\times\left(\exp\left((r-\frac{\sigma^2}{2})\Delta t+\sigma\sqrt{\Delta t}Z_2\right)\right)-w\right)\\
&\times\cdots\times\left(\exp\left((r-\frac{\sigma^2}{2})\Delta t+\sigma\sqrt{\Delta t}Z_N\right)\right)\Big]-w\\
=\ &S_0\exp\left((r-\frac{1}{2}\sigma^2)T+\sigma\sqrt{\Delta t}\big(Z_1+Z_2+\cdots+Z_N\big)\right)\\
&-w\Big[1+\exp\left((r-\frac{1}{2}\sigma^2)\Delta t+\sigma\sqrt{\Delta t}Z_N\right)\\
&\qquad+\exp\left((r-\frac{1}{2}\sigma^2)2\Delta t+\sigma\sqrt{\Delta t}\big(Z_N+Z_{N-1}\big)\right)\\
&\qquad+\exp\left((r-\frac{1}{2}\sigma^2)3\Delta t+\sigma\sqrt{\Delta t}\big(Z_N+Z_{N-1}+Z_{N-2}\big)\right)+\cdots\Big]\\
=\ &S_0\exp\left((r-\frac{\sigma^2}{2})T+\sigma\sqrt{\Delta t}\sum_{i=1}^{N}Z_i\right)\\
&-w\Big[1+\sum_{p=1}^{N-1}\left(\exp\left((r-\frac{\sigma^2}{2})p\Delta t+\sigma\sqrt{\Delta t}\sum_{q=1}^{p}Z_{N-q+1}\right)\right)\Big].
\end{aligned}\tag{7.28}$$

Appealing to this expression for S_T, it is easy to see that

$$\frac{dS_T}{dS_0}=\exp\left((r-\frac{1}{2}\sigma^2)T+\sigma\sqrt{\Delta t}\sum_{i=1}^{N}Z_i\right).\tag{7.29}$$

It is worth noting that, unlike for the simple Black-Scholes model, the equity value is not linear in S_0. That is, under the extended Black-Scholes model featuring withdrawals $\frac{dS_T}{dS_0} \neq \frac{S_T}{S_0}$.

Recalling the discounted payoff function for a European call option,

$$Y = e^{-rT} \max(S_T - K, 0). \tag{7.30}$$

we can now apply the pathwise estimator methodology with the equity following the Black-Scholes withdrawals dynamics. This results in the following estimator for the delta of the European call option:

$$\begin{aligned}
\Delta^{\mathrm{PW}} &= \frac{dY}{dS_0} \\
&= \frac{dY}{dS_T} \cdot \frac{dS_T}{dS_0} \\
&= e^{-rT} I(S_T > K) \exp\left((r - \frac{1}{2}\sigma^2)T + \sigma\sqrt{\Delta t} \sum_{i=1}^{N} Z_i \right).
\end{aligned} \tag{7.31}$$

Applying the likelihood ratio method (LRM) to estimate the delta of the European call option is very similar to how the technique applies to a basic European option. The likelihood ratio weight along each simulation path is simply given by

$$\frac{Z_1}{S_0 \sigma \sqrt{\Delta t}}, \tag{7.32}$$

where Z_1 is the shock which takes us from time zero out to the first withdrawal date of the equity asset. The LRM delta estimator is then given by multiplying this weight by the discounted value of the option along the same simulation path, i.e.,

$$\Delta^{\mathrm{LRM}} = e^{-rT} \max(S_T - K, 0) \frac{Z_1}{S_0 \sigma \sqrt{\Delta t}}, \tag{7.33}$$

where S_T is the terminal equity asset price (inclusive of the final withdrawal occurring at time T). A similar LRM estimator for the gamma of the European option can be found for this equity asset by employing the LRM weight for the gamma sensitivity, rather than the one for the delta sensitivity. The expression for this weight can be found as

$$\frac{Z_1^2 - Z_1 \sigma \sqrt{\Delta t} - 1}{S_0^2 \sigma^2 \Delta t}. \tag{7.34}$$

When looking to estimate the gamma sensitivity, one can also appeal to the mixed estimators, which combine the LRM and pathwise methods to produce an estimator which can potentially have much smaller associated standard error. The first of these mixed estimators involves applying the pathwise methodology to the LRM estimator for the delta of the call option. This yields the LR-PW mixed estimator for the gamma of the option. By differentiating the LRM delta estimator with respect to S_0, an expression for the LR-PW gamma estimator is found as follows:

$$\begin{aligned}
\Gamma^{\text{LR-PW}} &= \frac{d}{dS_0}\Delta^{\text{LRM}} \\
&= \frac{d}{dS_0}\left(e^{-rT}\max(S_T-K,0)\frac{Z_1}{S_0\sigma\sqrt{\Delta t}}\right) \\
&= \frac{Z_1}{S_0\sigma\sqrt{\Delta t}}\cdot e^{-rT}\frac{d}{dS_0}\max(S_T-K,0) \\
&\quad +e^{-rT}\max(S_T-K,0)\frac{d}{dS_0}\left(\frac{Z_1}{S_0\sigma\sqrt{\Delta t}}\right) \\
&= \frac{Z_1}{S_0\sigma\sqrt{\Delta t}}\cdot e^{-rT}I(S_T>K)\exp\left((r-\frac{1}{2}\sigma^2)T+\sigma\sqrt{\Delta t}\sum_{i=1}^{N}Z_i\right) \\
&\quad -e^{-rT}\max(S_T-K,0)\cdot\frac{Z_1}{S_0^2\sigma\sqrt{\Delta t}} \\
&= e^{-rT}I(S_T>K)\cdot\frac{Z_1}{S_0^2\sigma\sqrt{\Delta t}}\cdot \\
&\quad \left(K-S_T+S_0\exp\left((r-\frac{1}{2}\sigma^2)T+\sigma\sqrt{\Delta t}\sum_{i=1}^{N}Z_i\right)\right). \qquad (7.35)
\end{aligned}$$

If we apply the LRM to the pathwise estimator for the delta sensitivity, we can find an alternative mixed estimator for the gamma sensitivity of the option. In this case the mixed estimator is denoted the PW-LR mixed estimator for the gamma sensitivity of the European option. Let us proceed to derive an expression for this second mixed estimator.

When defining the PW-LR estimator for a European option under the Black-Scholes model, one may recall that the estimator had both a functional dependence on S_0 and a distributional dependence through the density of S_T. Multiplying the pathwise estimator for delta by the score function captures the distributional dependence, but we still must account for the functional dependence. For the basic Black-Scholes model this functional dependence term was obtained by taking the pathwise estimate

and differentiating it by S_0, i.e.,

$$\frac{d}{dS_0}\left(e^{-rT}I(S_T > K)\frac{S_T}{S_0}\right) = -e^{-rT}I(S_T > K)\frac{S_T}{S_0^2}. \tag{7.36}$$

Under the Black-Scholes withdrawals model, we no longer have $\frac{dS_T}{dS_0} = \frac{S_T}{S_0}$. Indeed, in this case

$$\begin{aligned} S_T = S_0 \exp\left((r - \frac{\sigma^2}{2})T + \sigma\sqrt{\Delta t}\sum_{i=1}^{N} Z_i\right) \\ - w\left[1 + \sum_{p=1}^{N-1}\left(\exp\left((r - \frac{\sigma^2}{2})p\Delta t + \sigma\sqrt{\Delta t}\sum_{q=1}^{p} Z_{N-q+1}\right)\right]. \end{aligned} \tag{7.37}$$

However, with the standard Black-Scholes model we had

$$\frac{dS_T}{dS_0} = \frac{S_T}{S_0} = \exp\left((r - \frac{1}{2}\sigma^2)T + \sigma\sqrt{T}Z\right) \tag{7.38}$$

and we were therefore differentiating this exponential term when finding the functional dependence in the PW-LR estimator. Thus, we wish to differentiate the analogous exponential term in our expression for S_T to find this functional dependence in the case of the Black-Scholes model featuring withdrawals. Let us define

$$A = \exp\left((r - \frac{1}{2}\sigma^2)T + \sigma\sqrt{\Delta t}\sum_{i=1}^{N} Z_i\right). \tag{7.39}$$

Then, by rearranging the expression for S_T in the Black-Scholes withdrawals model, we get

$$A = \frac{S_T}{S_0} + \frac{w}{S_0}\left[1 + \sum_{p=1}^{N-1}\exp\left((r - \frac{1}{2}\sigma^2)p\Delta t + \sigma\sqrt{\Delta t}\sum_{q=1}^{p} Z_{N-q+1}\right)\right]. \tag{7.40}$$

This can then be differentiated to yield

$$\frac{dA}{dS_0} = -\frac{S_T}{S_0^2} - \frac{w}{S_0^2}\left[1 + \sum_{p=1}^{N-1}\left(\exp\left((r - \frac{1}{2}\sigma^2)p\Delta t + \sigma\sqrt{\Delta t}\sum_{q=1}^{p} Z_{N-q+1}\right)\right]. \tag{7.41}$$

The PW-LR mixed estimator for the gamma sensitivity can then be found by adding the distributional dependence term, which is given by multiplying the pathwise delta

estimator by the score function. Using the expression for $\frac{dA}{dS_0}$, the PW-LR mixed estimator for the gamma sensitivity of the European option is then given by

$$\begin{aligned}\Gamma^{\text{PW-LR}} &= B(S_T)\exp\left((r-\frac{1}{2}\sigma^2)T+\sigma\sqrt{\Delta t}\sum_{i=1}^{N}Z_i\right)\frac{Z_1}{S_0\sigma\sqrt{\Delta t}}+B(S_T)\frac{dA}{dS_0}\\ &= B(S_T)\exp\left((r-\frac{1}{2}\sigma^2)T+\sigma\sqrt{\Delta t}\sum_{i=1}^{N}Z_i\right)\frac{Z_1}{S_0\sigma\sqrt{\Delta t}}\\ &- \frac{B(S_T)}{S_0^2}\left[S_T+w\left[1+\sum_{p=1}^{N-1}\left(\exp\left((r-\frac{1}{2}\sigma^2)p\Delta t+\sigma\sqrt{\Delta t}\sum_{q=1}^{p}Z_{N-q+1}\right)\right)\right]\right],\end{aligned} \tag{7.42}$$

where $B(S_T)=e^{-rT}I(S_T>K)$.

7.3 Testing sensitivity estimators

Having developed various estimators for both delta and gamma Greeks of a European option under a Black-Scholes model featuring withdrawals, let us now perform some analysis to see how these estimators perform in practice.

In order to test these various approaches, some basic European options will be considered, with different configurations of volatility, withdrawals dates and withdrawal amounts. These different option parameters used in the tests are stated in Table 7.1. For cases A-H the strike level is set by $K=\mathbb{E}[S_T]$, the formula for which was derived in Equation 7.26. In cases I-L the same settings of case H are considered, however we vary the level of the strike to test the methods in these instances.

In Table 7.2 the call option prices under all the cases considered are given for reference. In Table 7.3 the delta estimates for the options in all the cases are given for the bump and revalue, pathwise and LRM approaches. One can see the bump and revalue and pathwise estimates and standard errors are essentially indistinguishable. This is expected as in the small perturbation limit, the bump and revalue estimator will converge toward the pathwise estimator. In all cases, the pathwise method is preferable to the LRM for the delta sensitivity. This is also the case for European option delta sensitivities under the standard Black-Scholes model.

The gamma estimates from the bump and revalue and LRM are given in Table 7.4. From this table we can see that the LRM estimator generally seems to provide

Case	S_0	K	r	σ	w	T	Δt
A	100	22.87	4%	20%	20	1	0.25
B	100	22.66	4%	20%	10	1	0.125
C	100	63.47	4%	20%	10	1	0.25
D	100	63.37	4%	20%	5	1	0.125
E	100	22.87	4%	30%	20	1	0.25
F	100	22.66	4%	30%	10	1	0.125
G	100	63.47	4%	30%	10	1	0.25
H	100	63.37	4%	30%	5	1	0.125
I	100	45	4%	30%	5	1	0.125
J	100	55	4%	30%	5	1	0.125
K	100	75	4%	30%	5	1	0.125
L	100	85	4%	30%	5	1	0.125

Table 7.1: **Test Cases for Estimators:** European call option under the Black-Scholes model featuring withdrawals. Settings for Cases A-H. In all cases $1,000,000$ paths have been employed and the perturbation in the bumped paths is 0.05% of S_0. For cases A-H, $K = \mathbb{E}[S_T]$. Cases I-L take Case H settings, but vary the strike.

Case:	A	B	C	D	E	F	G	H	I	J	K	L
Price:	5.88	5.52	6.84	6.66	8.80	8.27	10.24	9.97	20.29	14.00	5.99	3.77
St.Err:	0.01	0.01	0.01	0.01	0.02	0.01	0.02	0.02	0.02	0.02	0.01	0.01

Table 7.2: **Price Estimates:** European call option under the Black-Scholes model featuring withdrawals. Results for each of cases A-L.

estimates for this gamma sensitivity which have smaller standard error than the bump and revalue estimator. We recall that the bump and revalue method is not expected to perform well in estimating gamma, since the pathwise estimator (to which the bump and revalue method converges in the small perturbation limit) is not available for second-order sensitivities. This is also why there is no pathwise gamma sensitivity estimator quoted in Table 7.4.

We can, however, combine the pathwise and LRM approaches to create the mixed estimators for this second-order sensitivity. The resultant mixed estimators and their standard errors are quoted in Table 7.5, together with a recap of the LRM estimator results for comparison. From these results, it appears that both the mixed estimators provide gamma estimates with significantly smaller standard error, than the LRM (or bump and revalue) estimator. The two mixed estimators also seem to provide estimates which are fairly consistent with each other for each of the cases considered.

Case	B&R Est.	St. Err.	PW Est.	St. Err.	LRM Est.	St. Err.
A	0.5322	5.9×10^{-4}	0.5322	5.9×10^{-4}	0.5320	1.6×10^{-3}
B	0.5314	5.9×10^{-4}	0.5314	5.9×10^{-4}	0.5338	1.9×10^{-3}
C	0.5365	5.9×10^{-4}	0.5365	5.9×10^{-4}	0.5364	1.7×10^{-3}
D	0.5363	5.9×10^{-4}	0.5362	5.9×10^{-4}	0.5393	2.1×10^{-3}
E	0.5483	6.4×10^{-4}	0.5483	6.4×10^{-4}	0.5480	1.7×10^{-3}
F	0.5470	6.5×10^{-4}	0.5470	6.5×10^{-4}	0.5498	1.9×10^{-3}
G	0.5548	6.5×10^{-4}	0.5548	6.5×10^{-4}	0.5546	1.8×10^{-3}
H	0.5544	6.5×10^{-4}	0.5544	6.5×10^{-4}	0.5578	2.3×10^{-3}
I	0.8245	5.4×10^{-4}	0.8245	5.4×10^{-4}	0.8283	3.2×10^{-3}
J	0.6822	6.2×10^{-4}	0.6822	6.2×10^{-4}	0.6862	2.7×10^{-3}
K	0.3906	6.3×10^{-4}	0.3906	6.3×10^{-4}	0.3928	1.8×10^{-3}
L	0.2751	5.9×10^{-4}	0.2751	5.9×10^{-4}	0.2774	1.5×10^{-3}

Table 7.3: **Delta Estimates:** European call option under the Black-Scholes model featuring withdrawals. Results for each of cases A-L.

Case	B&R Est.	St. Err.	LRM Est.	St. Err.
A	0.02801	6.1×10^{-4}	0.02691	2.8×10^{-4}
B	0.02942	6.3×10^{-4}	0.02934	4.5×10^{-4}
C	0.02238	5.5×10^{-4}	0.02307	3.0×10^{-4}
D	0.02310	5.5×10^{-4}	0.02460	4.9×10^{-4}
E	0.01802	4.9×10^{-4}	0.01788	2.0×10^{-4}
F	0.01908	5.0×10^{-4}	0.01954	3.2×10^{-4}
G	0.01511	4.5×10^{-4}	0.01530	2.1×10^{-4}
H	0.01612	4.7×10^{-4}	0.01637	3.5×10^{-4}
I	0.01077	3.4×10^{-4}	0.01135	4.7×10^{-4}
J	0.01489	4.2×10^{-4}	0.01504	4.0×10^{-4}
K	0.01500	4.7×10^{-4}	0.01564	2.8×10^{-4}
L	0.01334	4.7×10^{-4}	0.01348	2.4×10^{-4}

Table 7.4: **Gamma Estimates:** European call option under the Black-Scholes model featuring withdrawals. Results for each of cases A-L.

Overall, these tests show that all the Monte Carlo techniques for estimating option sensitivities considered earlier in the report can be extended to the Black-Scholes withdrawals model. Furthermore, the same advantages and disadvantages of each approach under the basic Black-Scholes model seem to apply when we incorporate withdrawals. That is, for the delta sensitivity the pathwise approach should be the preferred technique, since it gives the lowest variance in estimating this Greek and only requires a base simulation run and no further perturbed runs. For the gamma sensitivities, the mixed estimators provide the lowest standard errors and also only require a base simulation run.

Case	LRM Est.	St. Err.	PW-LR Est.	St. Err.	LR-PW Est.	St. Err.
A	0.02691	2.8×10^{-4}	0.02696	7.5×10^{-5}	0.02696	6.5×10^{-5}
B	0.02942	4.5×10^{-4}	0.02882	1.1×10^{-4}	0.02880	9.6×10^{-5}
C	0.02307	3.0×10^{-4}	0.02310	7.7×10^{-5}	0.02310	6.6×10^{-5}
D	0.02460	4.9×10^{-4}	0.02391	1.1×10^{-4}	0.02388	9.6×10^{-5}
E	0.01788	2.0×10^{-4}	0.01793	5.4×10^{-5}	0.01793	4.3×10^{-5}
F	0.01954	3.2×10^{-4}	0.01917	7.8×10^{-5}	0.01914	6.4×10^{-5}
G	0.01530	2.1×10^{-4}	0.01533	5.6×10^{-5}	0.01533	4.3×10^{-5}
H	0.01637	3.5×10^{-4}	0.01589	8.0×10^{-5}	0.01586	6.3×10^{-5}
I	0.01135	4.7×10^{-4}	0.01055	9.2×10^{-5}	0.01052	6.6×10^{-5}
J	0.01504	4.0×10^{-4}	0.01444	8.6×10^{-5}	0.01440	6.5×10^{-5}
K	0.01564	2.8×10^{-4}	0.01528	7.1×10^{-5}	0.01525	6.0×10^{-5}
L	0.01348	2.4×10^{-4}	0.01318	6.4×10^{-5}	0.01315	5.5×10^{-5}

Table 7.5: **Gamma mixed Estimates:** European call option under the Black-Scholes model featuring withdrawals. Results for each of cases A-L.

7.4 Liability sensitivities under the Black-Scholes withdrawals model

In the previous section, the problem of estimating the sensitivities of a European option under the Black-Scholes model featuring withdrawals was considered. This is, however, still quite different to the challenge of estimating the sensitivities of the liabilities on variable annuity contracts. With a European option the payoff consists of a maximum function, involving the terminal asset price and some chosen strike (which will typically not be too close to zero). On the other hand, a liability will be modelled by a maximum function which will only be non-zero when the asset price is equal or close to zero. In this section we will investigate how the estimators developed can be extended to liability sensitivities.

To determine the form of payoff representing an insurance liability, we must first understand what constitutes a liability to an insurer. Let us imagine that the underlying equity index is now some fund owned by a policyholder. Furthermore, let us imagine some insurance company has guaranteed that the policyholder will receive the agreed withdrawals for the lifetime of the contract. What will happen if, due to poor equity performance and the policyholder withdrawals, the fund runs out of money, but the term of the contract has not been reached yet? Under the terms of the contract, the policyholder is still entitled to receive the remaining withdrawals. In such circumstances, the insurer then becomes responsible to meet all the future withdrawals due to the policyholder out of its own reserves. The money required to

do this is what we refer to as a liability to the insurer. Thus, having described what an insurance liability is, let us now model this mathematically.

Before a formula for the liability can be constructed, the notation for the cashflows on the equity fund must be amended slightly. Let us now think of S_t as the stochastic cashflows on the policyholders fund, which is simply the result of their premium invested wholly in some equity index. Again some fixed withdrawals will be deducted from this fund at specific, periodic valuation dates. We define S_t^- to be the fund level at valuation time t before a withdrawal has been made and S_t^+ after the required withdrawal has been deducted and paid to the policyholder. This can be expressed mathematically as:

$$\begin{cases} S_t^- &= S_{t-1}^+ \exp\left((r - \frac{1}{2}\sigma^2)\Delta t + \sigma\sqrt{\Delta t} Z_t\right) \\ S_t^+ &= S_t^- - w. \end{cases} \tag{7.43}$$

The liability cashflow at any given valuation date can then be expressed as $L_t = \max(w - S_t^-, 0)$ and the overall liability for the insurer resulting from this contract is then given by the sum of the discounted liability cashflows at all the valuation dates, i.e.,

$$L = \sum_{t=1}^{N} e^{-rt\Delta t} \max(w - S_t^-, 0). \tag{7.44}$$

We now wish to estimate the first- and second-order sensitivities of this liability, L, with respect to the initial fund level, S_0, using the various approaches employed in Section 7.2. Extending the bump and revalue approach will present us with no new challenges, so we will begin by developing a pathwise estimator for the delta sensitivity of the liability L.

Applying the pathwise methodology to the liability, requires us to estimate $\frac{dL}{dS_0}$ along each simulation path. The derivative operator can be taken inside the summation and thus our problem becomes one of estimating $\frac{dL_t}{dS_0}$ at each valuation date along each path within the simulation.

This derivative can be expressed as

$$\frac{dL_t}{dS_0} = \frac{dL_t}{dS_t^-} \cdot \frac{dS_t^-}{dS_0} = -I(w > S_t^-) \cdot \frac{dS_t^-}{dS_0}. \tag{7.45}$$

To calculate $\frac{dS_t^-}{dS_0}$, we use the formulae for the dynamics of the fund given in Equation 7.43. This calculation proceeds recursively, as follows:

$$\begin{aligned}
\frac{dS_t^-}{dS_0} &= \frac{dS_t^-}{dS_{t-1}^+} \cdot \frac{dS_{t-1}^+}{dS_0} \\
&= \frac{dS_t^-}{dS_{t-1}^+} \cdot \frac{dS_{t-1}^+}{dS_{t-1}^-} \cdot \frac{dS_{t-1}^-}{dS_0} \\
&= \frac{dS_t^-}{dS_{t-1}^+} \cdot \frac{dS_{t-1}^+}{dS_{t-1}^-} \cdot \frac{dS_{t-1}^-}{dS_{t-2}^+} \cdot \frac{dS_{t-2}^+}{dS_{t-2}^-} \cdot \cdots\cdots \frac{dS_2^-}{dS_1^+} \cdot \frac{dS_1^+}{dS_1^-} \cdot \frac{dS_1^-}{dS_0} \\
&= \exp\left((r - \frac{1}{2}\sigma^2)\Delta t + \sigma\sqrt{\Delta t} Z_t\right) \exp\left((r - \frac{1}{2}\sigma^2)\Delta t + \sigma\sqrt{\Delta t} Z_{t-1}\right) \cdots \\
&= \exp\left((r - \frac{1}{2}\sigma^2)t\Delta t + \sigma\sqrt{\Delta t}\sum_{i=1}^{t} Z_i\right), \qquad (7.46)
\end{aligned}$$

where in the penultimate line above we have used the fact that $dS_n^+/dS_n^- = 1$ for $n = 1, \ldots, t-1$. Taking this expression for the derivative, the pathwise estimator for the delta sensitivity of the liability can be expressed as

$$\Delta^{\text{PW}} = -\sum_{t=1}^{N}\left[e^{-rt\Delta t} \cdot I(w > S_t^-) \cdot \exp\left((r - \frac{1}{2}\sigma^2)t\Delta t + \sigma\sqrt{\Delta t}\sum_{i=1}^{t} Z_i\right)\right], \quad (7.47)$$

where N is the number of timesteps (or valuation dates in the contract) and Z_i is the shock taking the asset from S_i^+ to S_{i+1}^- along the particular simulation path.

A LRM estimator can also be found for the delta sensitivity of the liability. The generality of this approach means such an estimator is constructed using exactly the same LRM weight as was used in the estimator for the European option under the Black-Scholes withdrawals model, which was given in Section 7.2. Recall, the formula for this weight was given by $\frac{Z_1}{S_0\sigma\sqrt{\Delta t}}$, where Z_1 is the shock which takes us from time zero out to the first withdrawal date of the equity asset. In this instance, however, the LRM estimator will be given by multiplying this weight with the expression for the liability, given in Equation 7.44. In full, the LRM estimator for the delta sensitivity of this fund's liability is given by

$$\Delta^{\text{LRM}} = \left(\sum_{t=1}^{N} e^{-rt\Delta t}\max(w - S_t^-, 0)\right) \cdot \frac{Z_1}{S_0\sigma\sqrt{\Delta t}}. \qquad (7.48)$$

An LRM estimator for the gamma sensitivity of the liability can be found by multiplying the liability, L, with the weight corresponding to the gamma Greek. The

formula for this weight is identical to the European option case. This was given in Equation 7.34.

In the European option tests performed in Section 7.3, the best estimators for the gamma sensitivity were found to be the mixed estimators. Therefore, we will now construct these estimators for the case of the gamma sensitivity of the liability in a similar manner. Firstly, the LR-PW mixed estimator will be developed.

To derive an expression for this estimator, we apply the pathwise methodology to the LRM estimator for the delta sensitivity of the liability. This proceeds as follows:

$$\begin{aligned}
\Gamma^{\text{LR-PW}} &= \frac{d}{dS_0}\Delta^{\text{LRM}} \\
&= \frac{d}{dS_0}\left[\left(\sum_{t=1}^{N} e^{-rt\Delta t}\max(w - S_t^-, 0)\right)\cdot\frac{Z_1}{S_0\sigma\sqrt{\Delta t}}\right] \\
&= -\left(\frac{Z_1}{S_0\sigma\sqrt{\Delta t}}\right)\sum_{t=1}^{N}\left[e^{-rt\Delta t}\cdot I(w > S_t^-)\cdot\exp\left((r-\frac{1}{2}\sigma^2)t\Delta t\right.\right. \\
&\quad \left.\left. + \sigma\sqrt{\Delta t}\sum_{i=1}^{t} Z_i\right)\right] - \left(\frac{Z_1}{S_0^2\sigma\sqrt{\Delta t}}\right)\left(\sum_{t=1}^{N} e^{-rt\Delta t}\max(w - S_t^-, 0)\right) \\
&= \frac{Z_1}{S_0\sigma\sqrt{\Delta t}}\cdot\Delta^{\text{PW}} - \frac{Z_1}{S_0^2\sigma\sqrt{\Delta t}}\cdot L \\
&= \frac{Z_1}{S_0\sigma\sqrt{\Delta t}}\cdot\left(\Delta^{\text{PW}} - \frac{L}{S_0}\right). \qquad (7.49)
\end{aligned}$$

If the pathwise estimator for the delta sensitivity of the liability is already being calculated, then it is fairly simple to combine this with the LRM weight to obtain the LR-PW estimator for the gamma sensitivity of the liability along each simulation path.

The PW-LR mixed estimator can also be found for the liability gamma. With this estimator, both the functional and distributional dependence on S_0 must be accounted for in calculating the derivative of the delta. This makes the calculation somewhat more complicated, however a PW-LR estimator can still be found in the following manner:

Recall, the pathwise estimator for the delta sensitivity of the liability is given by

$$\Delta^{\text{PW}} = -\sum_{t=1}^{N}\left[e^{-rt\Delta t}\cdot I(w > S_t^-)\cdot\exp\left((r-\frac{1}{2}\sigma^2)t\Delta t + \sigma\sqrt{\Delta t}\sum_{i=1}^{t} Z_i\right)\right]. \quad (7.50)$$

Now, the distributional dependence for the PW-LR mixed gamma estimator is captured by multiplying the pathwise estimator by the LRM weight for delta. The functional dependence then also needs to be accounted for, i.e.,

$$\Gamma^{\text{PW-LR}} = \Delta^{\text{PW}} \cdot \frac{Z_1}{S_0 \sigma \sqrt{\Delta t}} + \text{Functional Dependence.} \tag{7.51}$$

In order to derive an expression for this functional dependence, let us follow the approach used for finding the functional dependence for the European option gamma PW-LR mixed estimator, given in Section 7.2. The functional dependence will originate through the exponential terms within the summation of Equation 7.50, similarly to the exponential term A in the European option analysis. In order to proceed, let us define

$$A_k = \exp\left((r - \frac{1}{2}\sigma^2)k\Delta t + \sigma\sqrt{\Delta t}\sum_{i=1}^{k} Z_i\right). \tag{7.52}$$

Then, following the derivation of the European option gamma PW-LR estimator, the derivative of A_k can be expressed as

$$\begin{cases} \frac{dA_k}{dS_0} &= -\frac{S_k^+}{S_0^2} - \frac{w}{S_0^2}\left(1 + \sum_{p=2}^{k}\prod_{q=k}^{p} e^{(r-\frac{1}{2}\sigma^2)\Delta t + \sigma\sqrt{\Delta t}Z_q}\right), \quad \text{for k} = 2,\ldots, \text{N}. \\ \frac{dA_k}{dS_0} &= -\frac{S_1^+}{S_0^2} - \frac{w}{S_0^2}, \quad \text{for k} = 1. \end{cases} \tag{7.53}$$

Taking into account this functional dependence, the gamma PW-LR mixed estimator can finally be expressed as

$$\Gamma^{\text{PW-LR}} = \Delta^{\text{PW}} \cdot \frac{Z_1}{S_0 \sigma \sqrt{\Delta t}} - \sum_{t=1}^{N}\left[e^{-rt\Delta t} \cdot I(w > S_t^-) \cdot \frac{dA_t}{dS_0}\right], \tag{7.54}$$

where the derivatives of A_t are found recursively, as described above. This is slightly more difficult to implement than the LR-PW mixed estimator, which can be calculated without a great deal of further effort, if the pathwise estimator for the delta sensitivity has already been found.

Case	S_0	r	σ	w	T	Δt
A	100	4%	30%	25	1	0.25
B	100	4%	30%	12.5	1	0.125
C	100	4%	30%	6.25	1	0.0625
D	100	4%	30%	35	1	0.25
E	100	4%	30%	17.5	1	0.125
F	100	4%	30%	8.75	1	0.0625
G	100	4%	30%	20	1	0.25
H	100	4%	30%	10	1	0.125
I	100	4%	30%	5	1	0.0625

Table 7.6: **Test Cases for Estimators:** Liability under the Black-Scholes model featuring withdrawals. Settings for Cases A-I. In all cases $1,000,000$ paths have been employed and the perturbation in the bumped paths is 0.05% of S_0.

Case:	A	B	C	D	E	F	G	H	I
Liability:	6.88	6.39	6.13	37.18	37.29	37.38	1.02	0.80	0.70
St. Err:	0.009	0.009	0.008	0.017	0.016	0.015	0.003	0.003	0.002

Table 7.7: **Liability Estimates:** Liability under the Black-Scholes model featuring withdrawals. Results for each of cases A-I given in Table 7.6.

7.5 Testing sensitivity estimators: Liability case

Having developed the various estimators for the delta and gamma sensitivities of the liability in the Black-Scholes withdrawals model, let us now run some tests to compare their respective efficiencies. These tests will, again, consist of a number of different parameter configurations which we label as cases A-I. These are given in Table 7.6. The cases A-C have been chosen such that it is expected that the fund will be close to zero at the maturity of the contract. Cases D-F have been chosen, such that the fund is likely to reach zero (resulting in a liability to the insurer) earlier on in the contract, than in cases A-C, respectively. Cases G-I are set-up in the opposite manner, i.e., the fund is less likely to reach zero during the lifetime of the contract, than in cases A-C, respectively.

The expected value of the liabilities to the insurer for this product under cases A-I are given in Table 7.7. The values found for cases A-C are around the same level. This is also the situation for cases D-F and cases G-I, respectively. This is not surprising, since in the configuration of these cases, whenever the number of withdrawal times is doubled, the withdrawal amount is halved. The results in this table also show that

Case	B&R Est.	St. Err.	PW Est.	St. Err.	LRM Est.	St. Err.
A	-0.4074	4.1×10^{-4}	-0.4074	4.1×10^{-4}	-0.4075	9.9×10^{-4}
B	-0.4107	4.1×10^{-4}	-0.4107	4.1×10^{-4}	-0.4114	1.2×10^{-3}
C	-0.4125	4.1×10^{-4}	-0.4125	4.1×10^{-4}	-0.4114	1.5×10^{-3}
D	-0.9244	3.1×10^{-4}	-0.9244	3.1×10^{-4}	-0.9239	2.8×10^{-3}
E	-0.9444	2.9×10^{-4}	-0.9444	2.9×10^{-4}	-0.9473	3.9×10^{-3}
F	-0.9548	2.8×10^{-4}	-0.9548	2.8×10^{-4}	-0.9543	5.4×10^{-3}
G	-0.0948	2.3×10^{-4}	-0.0948	2.3×10^{-4}	-0.0946	3.9×10^{-4}
H	-0.0825	2.2×10^{-4}	-0.0825	2.2×10^{-4}	-0.0824	4.0×10^{-4}
I	-0.0759	2.1×10^{-4}	-0.0759	2.1×10^{-4}	-0.0754	4.4×10^{-4}

Table 7.8: **Delta Estimates:** Liability under the Black-Scholes model featuring withdrawals. Results for each of cases A-I given in Table 7.6.

Case	B&R Est.	St. Err.	LRM Est.	St. Err.
A	0.01901	5.0×10^{-4}	0.01893	1.3×10^{-4}
B	0.01943	5.0×10^{-4}	0.02068	2.0×10^{-4}
C	0.02088	5.2×10^{-4}	0.02121	3.4×10^{-4}
D	0.00781	4.1×10^{-4}	0.00715	3.2×10^{-4}
E	0.00622	3.8×10^{-4}	0.00661	5.7×10^{-4}
F	0.00562	3.7×10^{-4}	0.00468	1.1×10^{-3}
G	0.00781	2.8×10^{-4}	0.00803	5.6×10^{-5}
H	0.00789	2.8×10^{-4}	0.00784	7.6×10^{-5}
I	0.00754	2.7×10^{-4}	0.00756	1.1×10^{-4}

Table 7.9: **Gamma Estimates:** Liability under the Black-Scholes model featuring withdrawals. Results for each of cases A-I given in Table 7.6.

cases D-F produce larger liabilities than cases A-C, respectively. The opposite fact is true of cases G-I. This makes sense, since larger sized withdrawals occurring at the same number of timesteps should obviously produce larger expected liabilities.

Let us now turn our attention to how the liability delta sensitivity estimators perform. The bump and revalue, pathwise and LRM estimators and corresponding standard errors for this sensitivity are reported in Table 7.8. Similarly to the European option analysis in Section 7.3, the bump and revalue and pathwise estimators are consistent both in value and in the size of standard errors for each of the cases considered. Again, this is due to the bump and revalue method converging to the pathwise method in the small perturbation limit. The LRM estimates for cases A-I are also fairly consistent with the corresponding pathwise estimates, however the standard errors are consistently larger in the LRM. This is further confirmation that the pathwise method is generally more efficient than the LRM.

Case	LRM Est.	St. Err.	PW-LR Est.	St. Err.	LR-PW Est.	St. Err.
A	0.01893	1.3×10^{-4}	0.01897	3.6×10^{-5}	0.01897	4.2×10^{-5}
B	0.02068	2.0×10^{-4}	0.02062	5.3×10^{-5}	0.02062	6.1×10^{-5}
C	0.02121	3.4×10^{-4}	0.02162	7.6×10^{-5}	0.02161	8.7×10^{-5}
D	0.00715	3.2×10^{-4}	0.00717	5.8×10^{-5}	0.00716	8.4×10^{-5}
E	0.00661	5.7×10^{-4}	0.00621	8.9×10^{-5}	0.00624	1.2×10^{-4}
F	0.00468	1.1×10^{-3}	0.00552	1.3×10^{-4}	0.00552	1.8×10^{-4}
G	0.00803	5.6×10^{-5}	0.00807	2.3×10^{-5}	0.00806	2.5×10^{-5}
H	0.00784	7.6×10^{-5}	0.00784	2.8×10^{-5}	0.00784	3.0×10^{-5}
I	0.00756	1.1×10^{-4}	0.00760	3.5×10^{-5}	0.00760	3.7×10^{-5}

Table 7.10: **Gamma mixed Estimates:** Liability under the Black-Scholes model featuring withdrawals. Results for each of cases A-I given in Table 7.6.

In Table 7.9, the bump and revalue and LRM estimates and standard errors are given for the liability gamma sensitivities for each of the test cases. Again, the two methods seem to provide estimates which are in fairly close agreement for each of the cases A-I. Generally, the LRM provides estimates with lower variance than the bump and revalue approach, a fact which was also found in the European option analysis. This can be attributed to the fact that the pathwise approach is inapplicable for second-order sensitivities, thus the similar bump and revalue method is unlikely to give efficient estimates for gamma.

In the European option analysis, the mixed gamma sensitivity estimators outperformed both the bump and revalue and the LRM approaches, so it is likely that this will remain the case in estimating the gamma sensitivity of the liability. The estimates for this sensitivity under both the PW-LR and LR-PW mixed estimators are given in Table 7.10, together with a recap of the LRM estimates for comparison. The results in this table show that the mixed estimators do, indeed, outperform the gamma estimators of Table 7.9 for all the cases considered. For each of the cases A-I, the gamma estimates for the PW-LR and LR-PW methods are very close to one another, and the LRM approach also gives broadly similar estimates to those given by the two mixed estimators.

In conclusion, these tests show that the various Monte Carlo estimators for option Greeks can be adapted to estimate the sensitivities of the liability on the Black-Scholes withdrawals equity-linked product. Furthermore, the same advantages and disadvantages found for each of these option sensitivity estimators under the basic Black-Scholes model seem to apply in this extended setting. Again, the pathwise

approach performs best for the delta sensitivity and the mixed estimators are most efficient for the gamma sensitivities. Furthermore, these estimators only require the base simulation run and no further perturbed simulation path runs.

Overall, this chapter has shown how Monte Carlo estimators can be constructed for the sensitivities of the liability on an equity-linked product which offers periodic withdrawals to the policyholder. Such analysis will serve as an excellent guide in developing similar estimators for the sensitivity of the liability on a realistic unit-linked life insurance product. This is the purpose of Chapter 8 of the thesis.

Chapter 8

VA sensitivities under the Heston and Heston-CIR models

8.1 Introduction

As previously discussed in the thesis, one approach insurers can take in managing the risks they face in selling unit-linked products is to construct a hedging strategy. Setting up such a scheme involves calculating the key market risk sensitivities (or Greeks) on which the liability depends. The complex, path-dependent nature of many products means these sensitivities must be estimated by Monte Carlo simulation. Standard practice amongst insurers is to measure such sensitivities using a "bump and revalue" method. However, as well as requiring multiple valuations, this approach can be unreliable for higher order Greeks, for example in estimating the gamma sensitivity.

In this chapter we extend the standard Monte Carlo approaches for estimating option price sensitivities to the problem of estimating the sensitivities of the liability of a stylised unit-linked insurance product. This will be similar to the analysis given in Chapter 7, however here these approaches are adapted to be compatible with an economic model which is more realistic than the Black-Scholes framework. This will incorporate both stochastic volatility and stochastic interest rates. The estimators developed can also be easily generalised to work with the addition of equity jumps.

A review of the literature which is relevant to the analysis which will be conducted in this chapter is given in Section 1.1. To begin the process of developing the sophisticated liability sensitivity estimators, we will now discuss how the likelihood ratio method (LRM) can be adapted to work under the Heston and SVJD models which were introduced earlier in the thesis.

8.2 Conditional likelihood ratio method (CLRM)

In order to use the likelihood ratio method or construct the mixed second-order sensitivity estimators, described in Section 7.1.3 and 7.1.4, one must be able to derive the probability density function for the underlying asset returns. With the

Black-Scholes model, this density is easily found in closed-form, but it is well known that this model does not give a realistic description of true market dynamics (see for example Duffie and Pan [Duff97]). There is, however, an elegant approach which allows us to incorporate stochastic volatility and discontinuous jumps into the equity returns process, whilst still being able to utilise the tractability of the Black-Scholes model.

The conditional likelihood ratio method (CLRM) was introduced by Broadie and Kaya [Bro04] for the SVJD equity model, which was described in Section 5.2. An original innovation of this thesis is to extend this approach to also incorporate stochastic interest rates through the inclusion of the Cox-Ingersoll-Ross (CIR) model. However, to introduce and test the CLRM, let us first describe this technique in the context of the SVJD model.

Consider the problem of estimating the sensitivities of a European option written on an equity asset whose dynamics are modelled by a Heston model. An expression for S_T can be found in terms of the initial asset price S_0 as

$$S_T = S_0 \exp(Y_T) \exp\left(rT - \frac{1}{2}(1-\rho^2) \int_0^T V_t dt + \sqrt{1-\rho^2} \int_0^T \sqrt{V_t} dW_t^{Ind} \right), \quad (8.1)$$

where $Y_T = -\frac{1}{2}\rho^2 \int_0^T V_t dt + \rho \int_0^T \sqrt{V_t} dW_t^V$. The idea behind how to derive such an expression for S_T will be discussed in the context of the more sophisticated Heston-CIR model in Section 8.3. Similar logic can be applied here to give the above formula.

It can be seen from this expression that conditional on the variance process V_t, for $t \in [0, T)$, S_T is given by a lognormal distribution. With this insight, we can now appeal to any theoretical results which have been developed for the Black-Scholes model (for each fixed realisation of the variance process). In particular, the likelihood ratio weights (which are easily found for the Black-Scholes model) can be applied to the Heston model if one first conditions on the variance process. Also, the price of a European put option under the Heston SV model can now be expressed as

$$p^{\text{BS}}\left(S_0 \exp(Y_T), K, T, r, \sqrt{\frac{(1-\rho^2)}{T} \int_0^T V_t dt} \right), \quad (8.2)$$

where $p^{\text{BS}}(S_0, K, \sigma, r, T)$ is the Black-Scholes formula for the price of a European put

option in terms of the initial asset price S_0, strike price K, (constant) volatility of the underlying asset σ, risk-free interest rate r and time to maturity T. The various formulae for the Greeks under the Black-Scholes model can also be used to determine the conditional option sensitivities under the Heston model, with $S_0 \exp(Y_T)$ and $\frac{(1-\rho^2)}{T}\int_0^T V_t dt$ replacing S_0 and σ.

Lewis [Lew02] and [Lew01] noted that the problem of pricing options and finding sensitivities reduces to an expectation over the risk-adjusted volatility process alone. This idea is made more concrete by employing the Mixing Theorem of Romano and Touzi [Rom97], however this theorem is essentially a formalisation of the result we have derived above. Based on this insight Lewis developed the Monte Carlo mixing method, where one simulates a large number of volatility process paths and along each path values the option or calculates the required sensitivity analytically. An estimate for the price of the option or the Greek required is then given by averaging the values calculated over all the simulated volatility process paths.

The likelihood ratio weights for the Heston model can also now be determined given a fixed realistation of the underlying variance process:

By defining

$$\xi_T = \exp(Y_T) = \exp\left(-\frac{1}{2}\rho^2\int_0^T V_t dt + \rho\int_0^T \sqrt{V_t}dW_t^V\right), \tag{8.3}$$

$$\bar{\sigma}_T = \sqrt{\frac{(1-\rho^2)}{T}\int_0^T V_t dt}, \tag{8.4}$$

$$d_T = \frac{\ln\left(\frac{S_T}{S_0\xi_T}\right) - (r - \frac{1}{2}\bar{\sigma}_T^2)T}{\bar{\sigma}_T\sqrt{T}}, \tag{8.5}$$

the likelihood ratio weights for the delta, gamma and rho of a European option (with time T until maturity) can be expressed, respectively, as

$$\Delta_T^{\mathrm{W}} = \frac{d_T}{S_0\bar{\sigma}_T\sqrt{T}}, \tag{8.6}$$

$$\Gamma_T^{\mathrm{W}} = \frac{d_T^2 - d_T\bar{\sigma}_T\sqrt{T} - 1}{S_0^2\bar{\sigma}_T^2 T}, \tag{8.7}$$

$$\mathcal{P}_T^{\mathrm{W}} = -T + \frac{d\sqrt{T}}{\bar{\sigma}_T}. \tag{8.8}$$

These are just the Black-Scholes likelihood ratio weights from Section 7.1.3 with $S_0\xi_T$ replacing S_0 and $\bar{\sigma}_T$ replacing σ in the Black-Scholes model.

We approximate the first integral in Equations 8.3-8.5 with a discrete sum as

$$\int_0^T V_t dt \approx \sum_{i=0}^{T-\Delta t} V_i \Delta t \tag{8.9}$$

and the following relationship is used to determine the second integral

$$\int_0^T \sqrt{V_t} dW_t^V = \frac{1}{\sigma_v}\Big(\kappa\theta T - \kappa \int_0^T V_t dt + V_T - V_0\Big). \tag{8.10}$$

The above identity is also used in deriving the Andersen discretisation scheme (which is employed in the following analysis). Of course, by approximating the first integral using a discrete sum we introduce possible further bias into our resultant estimate of the option sensitivities. However, by increasing the number of timesteps this bias can be reduced. Broadie and Kaya [Bro06] proposed a method of determining these integrals exactly by using Fourier inversion techniques to invert the characteristic function of the variance distribution, as mentioned in Section 5.3.3.

Having derived the conditional likelihood ratio weights, we will now give a precise overview of how these can be used to estimate the sensitivities of a European option under the Heston model through Monte Carlo simulation. This will be given in the form of a series of high-level steps. A visual representation of the conditional simulation used in the CLRM is given in Figure 8.1.

Conditional Likelihood Ratio Method Overview:

1. Simulate a realisation of the variance process in the Heston model using one of the approaches introduced in Sections 5.3. All the analysis in this thesis simulating the Heston, SVJD and Heston-CIR models will use the Andersen method, which was described in Section 5.3.2.

2. Simulate a realisation of the equity asset return process, given the variance process generated in Step (1).

3. Calculate the likelihood ratio weights, given the realisations of the variance and asset price processes.

4. Calculate the discounted payoff of the option, given the realisations of the variance and asset price processes.

5. Estimate the sensitivities along this path by multiplying the values that were output in Steps (3) and (4).

6. Go to step (2) and generate another realisation of the asset process (but using the same variance process realisation). Steps (2)-(5) should be repeated until the required number of asset paths (per variance path) has been reached, say N paths. Once this number has been reached, proceed to step (7).

7. Go to step (1) and generate a new realisation of the variance process. Steps (1)-(7) should be repeated until the required number of variance process paths has been performed. Denote this number M. Now move to step (8).

8. In total we should now have MN estimates of each sensitivity required. Taking the average of these estimates will give the conditional likelihood ratio estimator for the required sensitivities of the option.

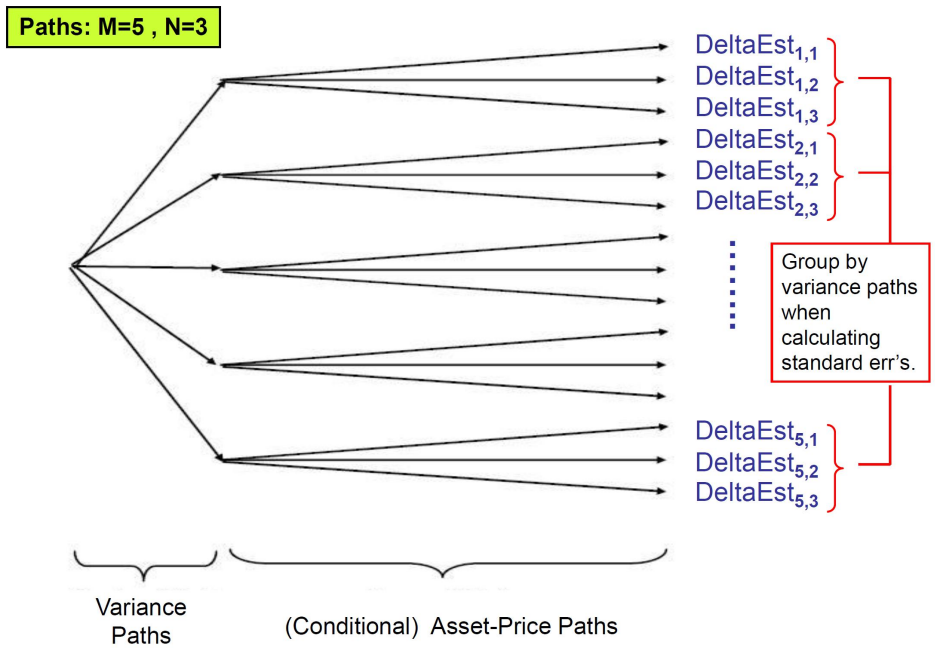

Figure 8.1: A visual representation of the conditional simulation set-up with 5 variance paths and 3 asset price paths. These asset paths are conditional on the realisation of the adjacent variance path. The graphic also shows the groupings of independent asset paths, used in calculating the standard errors of estimators.

The likelihood ratio method can also incorporate discontinuous jumps in the equity model, without much extra effort. Under the SVJD model we now condition both on a realisation of the Heston variance process and the observed number of jumps. The only difference between the likelihood ratio weights for the SVJD and Heston models is in the calculation of ξ_T and $\bar{\sigma}_T$ terms, which was performed by Equations 8.3 and 8.5 for the Heston model. For the SVJD model, we replace these by

$$\xi_T = \exp\left(n_j(\mu_S + \sigma_S^2/2) - \lambda\bar{\mu}T - \frac{1}{2}\rho^2\int_0^T V_t dt + \rho\int_0^T \sqrt{V_t}dW_t^V\right), \tag{8.11}$$

$$\bar{\sigma}_T = \sqrt{\frac{n_j\sigma_S^2 + (1-\rho^2)\int_0^T V_t dt}{T}}, \tag{8.12}$$

where n_j, μ_S, σ_S, $\bar{\mu}$ and λ are as denoted in Section 5.2.

With ξ_T and $\bar{\sigma}_T$ now modified to take into account the addition of the jumps in the equity returns, the conditional asset price at time T can be expressed as

$$S_T = S_0\xi_T \exp\left((r - \frac{1}{2}\bar{\sigma}_T)T + \bar{\sigma}_T\sqrt{T}Z\right), \tag{8.13}$$

where Z is a standard normal variate. Thus, lognormally distributed equity returns can be obtained under the SVJD model by first conditioning on the variance process *and* the number of the lognormally distributed asset price jumps. The conditional likelihood ratio weights are then just given by the same formulae as for the Heston model (Equations 8.6-8.8), but with the above expressions for the ξ_T and $\bar{\sigma}_T$ used therein.

An important point worth mentioning is the way in which the standard error of the CLRM estimator is calculated. The MN pathwise sensitivity estimates are not independent, due to the fact that there are M groups of N estimates which are all based upon the same variance process realisation. Therefore, it is not valid to simply take the standard error as $\sigma_{\text{all}}/\sqrt{MN}$, where σ_{all} is the standard deviation of all the MN estimates together. Instead, we group the MN paths into M subgroups each of size N, where each estimate in a subgroup is based upon the same realisation of the variance process. The average of each subgroup is then taken, resulting in M 'averaged' estimates. These 'averaged' estimates will now all be based on independent realisations of the variance process. Thus, if we let σ^* denote the standard deviation of the M 'averaged' estimates, the conditional likelihood ratio

estimator standard error is then given by $\sigma^*/\sqrt{M}$. This approach for calculating the standard error of a CLRM estimator is illustrated graphically in Figure 8.1.

In the Broadie and Kaya paper, the formula for the likelihood ratio weights are not derived, therefore this thesis will demonstrate how they can be found. However, rather than give this derivation in the context of the SVJD model, we will derive the weights for the more sophisticated Heston-CIR model in the next section. The same logic which is followed there can be used to derive the weights for the SVJD model. Detailing this calculation is a welcome addition to the literature, where a clear reference to a paper showing how to derive such formulae is lacking. Furthermore, the derivation of the likelihood ratio weights under the more sophisticated Heston-CIR economic model is a wholly original contribution of this thesis.

To compare the CLRM with the simple bump and revalue approach, some tests will be performed determining the delta and gamma sensitivities on a European put option, where the underlying asset follows the Heston model. The estimators will be determined using two different simulation approaches. Firstly, the standard bump and revalue approach will be used to give estimates of the value of the option and the sensitivities by generating upward and downward perturbed simulation paths, as well as the base simulation run. Next, the CLRM estimators will be found by performing a new simulation involving just a base run. In order to have a fair comparison of the two approaches, the number of timesteps was kept at 50/year in all the simulations and the number of paths was chosen so both the bump and revalue and CLRM simulation approaches took approximately the same time to run. This was set as 40,000 paths for the bump and revalue simulations and 10,000 variance paths with 10 asset price paths per variance path for the CLRM. In all the analysis which follows the option considered is a European put option with $S_0 = 100$, $r = 4\%$, $T = 6$ and asset process-volatility process correlation $\rho = -0.7$.

The 'analytical' values for the option price and sensitivities will also be quoted in the results given. These 'analytical' (often denoted semi-analytical) values are found numerically to high accuracy using Fourier inversion techniques and numerical integration. A full introduction to the semi-analytical valuation of European option prices under the Heston model was given in Chapter 6 of this thesis.

In Table 8.1 the estimates for the European option prices under both simulation ap-

S_0	K	κ	σ_V	θ	V_0	B&R	St.Err.	CLRM	St.Err.	Analyt. Val.
100	90	2	0.3	0.04	0.04	6.09	0.08	6.02	0.11	6.12
100	100	2	0.3	0.04	0.04	8.29	0.10	8.14	0.14	8.30
100	110	2	0.3	0.04	0.04	10.84	0.11	10.84	0.17	10.84
100	90	1	0.6	0.04	0.04	5.99	0.10	6.08	0.15	6.07
100	100	1	0.6	0.04	0.04	7.74	0.12	7.63	0.18	7.83
100	110	1	0.6	0.04	0.04	9.78	0.13	9.83	0.20	9.88
100	90	4	0.1	0.04	0.04	5.81	0.07	5.74	0.10	5.83
100	100	4	0.1	0.04	0.04	8.11	0.08	8.04	0.12	8.12
100	110	4	0.1	0.04	0.04	10.79	0.10	10.57	0.15	10.81

Table 8.1: **Price Estimates**: Heston model. Two different simulation set-ups – standard MC for the bump and revalue; conditional simulation for the CLRM.

proaches are given for five different model parameter settings. The estimated prices under the bump and revalue and CLRM are similar, with neither being consistently closer to the analytical value for all parameter settings. The bump and revalue approach, however, does have consistently smaller standard errors, by a factor of around 1.5. This is expected since the CLRM requires a conditional simulation framework which will not give as many independent equity price paths, as will be simulated in the bump and revalue method. However, efficient valuation of the current liability is not of upmost concern. It is efficient estimators for the sensitivities of this liability which are needed.

In Table 8.2 the estimates for the delta sensitivities of the option are given for both simulation approaches, together with the analytical value obtained by the Fourier inversion techniques outlined in Chapter 6. These results seem to show that the bump and revalue approach consistently gives estimates which are closer to the analytical value than the CLRM. Furthermore, the standard errors associated with these estimates are around 0.75 times the size of the standard errors associated with the CLRM. This trend is consistent with the previous literature on this topic. See Glasserman [Gla03], for example. Generally, the pathwise approach is more efficient than the likelihood ratio method when it is applicable (and in such circumstances the bump and revalue method will give results which are very similar to the pathwise estimator if it is employed with a small perturbation).

The problem with the pathwise approach is that it is inapplicable when the payoff function is discontinuous, or in estimating second-order sensitivities. This was discussed in more detail in Section 7.1.2. With the similarity between the pathwise and

S_0	K	κ	σ_V	θ	V_0	B&R	St.Err.	CLRM	St.Err.	Analyt. Val.
100	90	2	0.3	0.04	0.04	-0.1203	0.0013	-0.1183	0.0016	-0.1187
100	100	2	0.3	0.04	0.04	-0.1549	0.0014	-0.1543	0.0019	-0.1532
100	110	2	0.3	0.04	0.04	-0.1916	0.0016	-0.1910	0.0021	-0.1906
100	90	1	0.6	0.04	0.04	-0.0857	0.0011	-0.0895	0.0015	-0.0855
100	100	1	0.6	0.04	0.04	-0.1113	0.0013	-0.1098	0.0017	-0.1117
100	110	1	0.6	0.04	0.04	-0.1421	0.0015	-0.1465	0.0020	-0.1423
100	90	4	0.1	0.04	0.04	-0.1302	0.0013	-0.1292	0.0017	-0.1301
100	100	4	0.1	0.04	0.04	-0.1687	0.0015	-0.1673	0.0019	-0.1683
100	110	4	0.1	0.04	0.04	-0.2077	0.0016	-0.2048	0.0021	-0.2086

Table 8.2: **Delta Estimates**: Heston model.

the bump and revalue approaches, it might be expected that the bump and revalue approach will also give poor results in such circumstances. With this in mind, let us investigate the success of the bump and revalue and CLRM in estimating the gamma sensitivity of the European put option under the Heston model dynamics.

In Table 8.3 the estimates for the gamma sensitivities are given for the bump and revalue and CLRM approaches. It is immediately clear from these results that our intuition was correct and the bump and revalue approach gives very large standard errors in the option price gamma estimates for all the different parameter settings. On the other hand the CLRM gives significantly smaller standard errors in estimating this sensitivity. The very large standard errors from the bump and revalue approach essentially makes this method inapplicable for estimating gamma.

Overall, these tests show that the bump and revalue approach can give good results for option prices and sensitivities when the payoff function is continuous. However, in instances when this payoff function is discontinuous, and in estimating second-order sensitivities, this method breaks down and becomes impractical. With the complex features of many unit-linked insurance products, such as ratchets, policyholder bonuses and the possibility of policyholder lapse, the payoff functions associated with the insurance liabilities of these are likely to be discontinuous in nature. Under such discontinuities, the basic bump and revalue approach will become very inefficient in estimating second-order sensitivities. The performance of the CLRM, on the other hand, does not deteriorate in the presence of discontinuities. Therefore, this approach could be of great benefit to practitioners looking to gamma hedge their unit-linked insurance liabilities. In the next section this CLRM technique will be extended to work under both stochastic interest rates and stochastic volatility.

S_0	K	κ	σ_V	θ	B&R	St.Err.	CLRM	St.Err.	Analytical
100	90	2	0.3	0.04	0.002403	0.00044	0.002981	5.61×10^{-5}	0.002956
100	100	2	0.3	0.04	0.003953	0.00058	0.003638	6.49×10^{-5}	0.003608
100	110	2	0.3	0.04	0.003800	0.00060	0.004265	7.63×10^{-5}	0.004241
100	90	1	0.6	0.04	0.002392	0.00041	0.002338	5.04×10^{-5}	0.002163
100	100	1	0.6	0.04	0.002704	0.00049	0.002865	5.87×10^{-5}	0.002836
100	110	1	0.6	0.04	0.003404	0.00061	0.003987	7.70×10^{-5}	0.003613
100	90	4	0.1	0.04	0.004064	0.00057	0.003310	5.94×10^{-5}	0.003305
100	100	4	0.1	0.04	0.004040	0.00062	0.003901	6.92×10^{-5}	0.003943
100	110	4	0.1	0.04	0.005276	0.00072	0.004473	7.93×10^{-5}	0.004521

Table 8.3: **Gamma Estimates**: Heston model. $V_0 = 0.04$ in all the tests above.

8.3 CLRM for the Heston-CIR model

The conditional likelihood ratio method (CLRM) has been introduced for the Heston and SVJD models of equity index returns. This followed the paper of Broadie and Kaya [Bro04]. In this section, this method will be extended to also incorporate stochastic interest rates through the Cox-Ingersoll-Ross (CIR) model. We call this the Heston-CIR model. This could be further extended to also include discontinuous jumps in the equity process, in a manner similar to the previous section. However, this has been omitted here for ease of illustration of the method. In full, the system of stochastic differential equations which govern our asset price dynamics in the Heston-CIR model is given by:

$$dS_t = r_t S_t dt + \sqrt{V_t} S_t dW_t^S, \tag{8.14}$$

$$dV_t = \kappa_V(\theta_V - V_t)dt + \sigma_V \sqrt{V_t} dW_t^V, \tag{8.15}$$

$$dr_t = \kappa_r(\theta_r - r_t)dt + \sigma_r \sqrt{r_t} dW_t^r. \tag{8.16}$$

The variance and interest rates both follow exactly the same form of stochastic process (but calibrated with different parameters). The Brownian motions for the asset, volatility and interest rate processes are correlated as follows:

$$\text{corr}(W_t^S, W_t^V) = \rho_{S,V}, \tag{8.17}$$

$$\text{corr}(W_t^S, W_t^r) = \rho_{S,r}, \tag{8.18}$$

$$\text{corr}(W_t^V, W_t^r) = \rho_{V,r}. \tag{8.19}$$

This model has rarely appeared in the financial mathematics literature. Indeed, when the following analysis was first performed, this author was not aware of the model appearing in any other article. Later, it came to his attention that Grzelak and Oosterlee [Grz10] had constructed an affine approximation to the Heston-CIR model. This article complements the analysis in this thesis, since an affine approximation for this model gives an efficient means for calibrating the Heston-CIR model to market data. This would obviously be a key concern to practitioners looking to use this model in estimating insurance liability sensitivities.

In order to construct this correlation structure among the three normally distributed risk-drivers in this model we employ a Cholesky decomposition. Let (Z_1, Z_2, Z_3) be independent standard normal variates. Then standardised increments (Z_V, Z_r, Z_S) for the variance, interest rate and equity processes are constructed by setting

$$(Z_v, Z_r, Z_S)' = A \cdot (Z_1, Z_2, Z_3). \tag{8.20}$$

Here A is the lower-triangular matrix satisfying $A \cdot A' = \rho$, where ρ is the correlation matrix constructed from $\rho_{S,V}$, $\rho_{S,r}$ and $\rho_{V,r}$. The correlation matrix ρ is assumed to be positive-definite. For completeness, the matrix A is given by:

$$A = \begin{pmatrix} 1 & 0 & 0 \\ \rho_{V,r} & \sqrt{1-\rho_{V,r}^2} & 0 \\ \rho_{S,V} & \frac{\rho_{S,r}-\rho_{S,V}\rho_{V,r}}{\sqrt{1-\rho_{V,r}^2}} & \sqrt{1-\rho_{S,V}^2-\frac{(\rho_{S,r}-\rho_{S,V}\rho_{V,r})^2}{1-\rho_{V,r}^2}} \end{pmatrix} \tag{8.21}$$

The random variable Z_r will now have correlation $\rho_{V,r}$ with Z_V. Also, the random normal variate Z_S will have correlation $\rho_{S,V}$ with Z_V and correlation $\rho_{S,r}$ with Z_r. This completes our Cholesky decomposition for the correlation structure of the three risk-drivers out to the first timestep. Now by conditioning on a realisation of the variance and interest rate processes, the asset returns become lognormal. One can then use the same form of LRM weights as for the Black-Scholes models, but with the addition of a couple of extra factors to account for the conditional information. Let us explore this idea further. Using the Cholesky decomposition the asset price

dynamics can now be expressed as

$$\begin{aligned}\frac{dS_t}{S_t} &= r_t dt + \sqrt{V_t} dW_t^S \\ &= r_t dt + \sqrt{V_t}\bigg(a_{31} dW_t^{I_1} + a_{32} dW_t^{I_2} + a_{33} dW_t^{I_3}\bigg),\end{aligned} \tag{8.22}$$

where $W_t^{I_1}$, $W_t^{I_2}$ and $W_t^{I_3}$ are three independent Brownian motion increments, corresponding to Z_1, Z_2 and Z_3 in our discretisation and a_{ij} is the element in row i and column j of matrix A. This construction for dS_t expresses the asset price dynamics in terms of all the risk-drivers in the system of SDEs, which together dictate the behaviour of the asset returns.

In theory an interest rate process with a larger number of risk-factors could be used. One would employ a Cholesky decomposition to correlate all the risk-drivers as required and express the asset price dynamics, dS_t, in terms of all these other risk-drivers. One would however do this using a numerical Cholesky decomposition program to find the coefficients.

To proceed towards a conditional lognormal representation of the asset returns, we need an expression for the stochastic process representing the logarithm of the returns, X_t. This expression for dX_t is given as

$$dX_t = r_t dt + dY_t - \frac{V_t}{2} a_{33}^2 dt + \sqrt{V_t} a_{33}\ dW_t^{I_3}, \tag{8.23}$$

with

$$dY_t = -\frac{1}{2} V_t a_{31}^2 dt - \frac{1}{2} V_t a_{32}^2 dt + \sqrt{V_t} a_{31} dW_t^{I_1} + \sqrt{V_t} a_{32} dW_t^{I_2}. \tag{8.24}$$

A short proof of the above result will now be given.

Proof. Recall, that by employing the Cholesky decomposition the asset price dynamics under the Heston-CIR model were expressed in Section 8.3 as

$$\begin{aligned}\frac{dS_t}{S_t} &= r_t dt + \sqrt{V_t} dW_t^S \\ &= r_t dt + \sqrt{V_t}\bigg(a_{31} dW_t^{I_1} + a_{32} dW_t^{I_2} + a_{33} dW_t^{I_3}\bigg),\end{aligned} \tag{8.25}$$

where $W_t^{I_1}$, $W_t^{I_2}$ and $W_t^{I_3}$ are three independent Brownian motions.

Using this formula the expression for the stochastic process representing the logarithm of the returns, X_t, can be found as follows. Let $g(t,x)$ be such that $X_t = g(t,S_t) = \ln S_t$. Then,

$$\frac{\partial g}{\partial t} = 0, \qquad \frac{\partial g}{\partial x} = \frac{1}{x} \quad \text{and} \quad \frac{\partial^2 g}{\partial x^2} = -\frac{1}{x^2}. \tag{8.26}$$

By employing Itô's formula, the stochastic process for the logarithm of the asset returns, X_t, under the Heston-CIR model, is then given by:

$$\begin{aligned}
dX_t &= \frac{\partial g}{\partial t}(t,X_t)dt + \frac{\partial g}{\partial x}(t,X_t)dS_t + \frac{1}{2}\frac{\partial^2 g}{\partial x^2}(t,X_t)(dS_t)^2 \\
&= \frac{dS_t}{S_t} + \frac{1}{2}\left(-\frac{1}{S_t^2}\right)(dS_t)^2 \\
&= r_t dt + \sqrt{V_t}\left(a_{31}dW_t^{I_1} + a_{32}dW_t^{I_2} + a_{33}dW_t^{I_3}\right) \\
&\quad -\frac{1}{2}\frac{1}{S_t^2}\left(S_t r_t dt + S_t\sqrt{V_t}\left(a_{31}dW_t^{I_1} + a_{32}dW_t^{I_2} + a_{33}dW_t^{I_3}\right)\right)^2 \\
&= r_t dt + \sqrt{V_t}a_{31}dW_t^{I_1} + \sqrt{V_t}a_{32}dW_t^{I_2} + \sqrt{V_t}a_{33}dW_t^{I_3} \\
&\quad -\frac{1}{2}V_t a_{31}^2 dt - \frac{1}{2}V_t a_{32}^2 dt - \frac{1}{2}V_t a_{33}^2 dt \\
&= r_t dt + dY_t - \frac{V_t}{2}a_{33}^2 dt + \sqrt{V_t}a_{33}dW_t^{I_3} \\
&\quad \text{with} \\
&\quad dY_t = -\frac{1}{2}V_t a_{31}^2 dt - \frac{1}{2}V_t a_{32}^2 dt + \sqrt{V_t}a_{31}dW_t^{I_1} + \sqrt{V_t}a_{32}dW_t^{I_2}.
\end{aligned} \tag{8.27}$$

□

With this result for dX_t, an expression for S_T in terms of the initial asset price S_0 can then be found as

$$S_T = S_0 \exp(Y_T) \exp\left(\frac{1}{T}\int_0^T r_t dt - \frac{1}{2}a_{33}^2\frac{1}{T}\int_0^T V_t dt + a_{33}\frac{1}{T}\int_0^T \sqrt{V_t}dW_t^{I_3}\right), \tag{8.28}$$

where Y_T is given by

$$-\frac{1}{2}a_{31}^2\frac{1}{T}\int_0^T V_t dt - \frac{1}{2}a_{32}^2\frac{1}{T}\int_0^T V_t dt + a_{31}\frac{1}{T}\int_0^T \sqrt{V_t}dW_t^{I_1} + a_{32}\frac{1}{T}\int_0^T \sqrt{V_t}dW_t^{I_2}. \tag{8.29}$$

By defining

$$\xi_T = \exp(Y_T), \tag{8.30}$$

$$\bar{\sigma}_T = \sqrt{\frac{a_{33}^2}{T}\int_0^T V_t dt}, \tag{8.31}$$

$$\bar{r}_T = \frac{1}{T}\int_0^T r_t dt, \tag{8.32}$$

the price of a European put option under the Heston-CIR model is given by

$$p^{\text{BS}}(S_0\xi_T, K, \bar{\sigma}_T, \bar{r}_T, T), \tag{8.33}$$

where p^{BS} is the Black-Scholes analytical formula for the price of a European put option in terms of the model parameters. Also, the CLRM weights can be determined using these adjusted expressions for S_0, σ and r with the Black-Scholes score function, for example Equation 7.13 for the delta sensitivity. These weights can then be used to obtain the CLRM estimators of the sensitivities of options under this sophisticated stochastic model for the underlying equity asset and risk-free interest rate. Naturally, the simulation must, again, be performed conditionally, this time on realisations of the variance and interest rate processes.

Following the same logic in this derivation of the likelihood ratio weights for the Heston-CIR model, one can derive the equivalent weights under the Heston and SVJD models. These weights are usually just quoted without derivation in the literature, so this description hopefully fills this small gap in the literature and makes the technique more transparent for insurance practitioners.

8.4 Variable annuity liability sensitivities

8.4.1 Stylised variable annuity product

The methods for estimating option price sensitivities will now be applied to the problem of estimating the sensitivity of the liabilities on a stylised variable annuity product. The idea behind this stylised example product is that it should be simple enough both in terms of tractability and ease of exposition, yet retain some of the key features which make these products so popular in many markets. It should also

have liabilities which are path-dependent, as is typical of many VA products on the market. This product will be very similar to the Guaranteed Minimum Withdrawal (GMWB) introduced in Sections 4.1 and 4.2. More detail regarding this VA product and its associated cashflows can be found there, however we note two differences behind the analysis given there and that which will follow; firstly, in Section 4.1 it was assumed that the rate of policyholder lapse was given in terms of a 'dynamic' function of the fund level, guarantee base and interest-rates. In the analysis which follows, we will just assume this rate remains constant at 4%. However, with some further research a dynamic model of lapsation, should be compatible with the estimators which will be derived throughout this chapter. Secondly, the underlying VA fund, which is initially funded by the policyholder premium, is now assumed to be invested wholly in a single equity index with dynamics which are governed by the Heston-CIR model, given in Equations 8.14-8.16. Possible extensions of this analysis include investment in a mixture of equities and bonds, as was assumed in Chapter 4, or in a portfolio of two different equity indices. These alternative investment assumptions are another line of further research.

To model the cashflows on this policy mathematically let us define the fund value, guarantee base and policyholder income level at year t after annuitisation by F_t, G_t and I_t, respectively. Also let R_t denote the return from the equity index from year $t-1$ to year t minus the management fees, i.e., $R_t = S_t/S_{t-1} - \eta$, where η is the fund management charge, quoted as an annual percentage. Then we can track the level of fund value throughout the lifetime of the policy using the following equation:

$$F_t = \max\big((F_{t-1} - I_{t-1})(1 + R_t), 0\big). \tag{8.34}$$

This expresses the fund value at year t in terms of the fund value and income level at year $t-1$. Naturally then, we require the starting fund and income levels to initiate this recursion. The initial fund value is just given by the policyholder premium at annuitisation, P (in units of the initial equity level), multiplied by the initial equity index level: $F_0 = P \cdot S_0$. The guarantee base at the end of year t after annuitisation can be expressed as

$$G_t = I(t \leq \alpha) \min\bigg(\max(G_{t-1}, F_t), 1.15 \times G_{t-1} \bigg) + I(t > \alpha) G_{t-1}. \tag{8.35}$$

The income level the policyholder withdraws from the policy fund value at the end of year t is then given by $I_t = (w - \mu)G_t$. Here, w is a fixed parameter dictating the proportion of the guarantee base which is withdrawn by the policyholder at each annual re-balancing date and μ is the guarantee charge, taken as a percentage of the guarantee base each year. In our case we have $w = 4\%$ and $\mu = 1\%$.

The liability, or guarantee shortfall, the insurer faces from issuing this VA contract on the market, measured at annuitisation, can then be expressed as

$$L = \mathbb{E}\left[\sum_{t=1}^{T} D_t p_t^{surv} \max(I_t - F_t, 0)\right], \tag{8.36}$$

where D_t is the factor to ensure the liability at each yearly withdrawal date is discounted back to annuitisation, p_t^{surv} is the probability of the policy remaining in force until year t after annuitisation (encompassing both the possibility of policyholder mortality and lapsation) and T is the maximum contract term. Clearly, the insurer only faces a liability under the GMWB contract when the policyholder income cannot be met by the VA fund level. This is captured by the max function in the above formula for L, which sets the summand to zero at the rebalancing/withdrawal dates where $F_t \geq I_t$. More detailed discussion of these cashflows was given in Section 4.2.

8.4.2 Pathwise VA liability estimator

Having outlined the relevant cashflows on this stylised VA product, we will now develop a pathwise methodology for the sensitivities of the liability on this product. This approach, proposed for a simple VA product by Hobbs et al. [Hob09], is just the natural extension of the pathwise approach for option sensitivities to the case of the VA liability, L. The liability is analogous to a series of European options of increasing maturity, thus the same limitations of the pathwise approach for European option price sensitivities will apply here. Assuming that interchanging the order of differentiation and the taking of the expectation is justified, the derivative of the liability with respect to the initial equity asset price can be expressed as

$$\Delta_{\mathrm{PW}} = \mathbb{E}\left[\sum_{t=1}^{T} D_t p_t^{surv} \frac{\partial}{\partial S_0} \max(I_t - F_t, 0)\right]. \tag{8.37}$$

The derivative only acts on the third factor of each term in the sum, and we can express the derivative of this term as

$$\frac{\partial}{\partial S_0}\max(I_t - F_t, 0) = I(I_t > F_t)\cdot\Big(\frac{\partial I_t}{\partial S_0} - \frac{\partial F_t}{\partial S_0}\Big). \tag{8.38}$$

The problem of estimating the delta sensitivity of the VA liability is now one of estimating the derivative of the fund value, F_t, and income level, I_t, for each year, t, after annuitisation. Appealing to the structure of the product's cashflows, these derivatives must be calculated recursively. Using Equation 8.34 we can express the derivative of the year t fund value with respect to S_0 as

$$\frac{\partial F_t}{\partial S_0} = \max\bigg(\Big(\frac{\partial F_{t-1}}{\partial S_0} - \frac{\partial I_{t-1}}{\partial S_0}\Big)(1+R_t), 0\bigg). \tag{8.39}$$

Similarly, using Equation 8.35, the derivative of the year t income level with respect to S_0 can be expressed using

$$\begin{aligned}\frac{\partial G_t}{\partial S_0} &= I\big((F_t \geq G_{t-1}) \cap (F_t \leq 1.15 \times G_{t-1}) \cap (t \leq \alpha)\big)\frac{\partial F_t}{\partial S_0} \\ &\quad + I\big((F_t \geq G_{t-1}) \cap (F_t > 1.15 \times G_{t-1}) \cap (t \leq \alpha)\big) \times 1.15 \times \frac{\partial G_{t-1}}{\partial S_0} \\ &\quad + I\big((F_t < G_{t-1}) \cap (t \leq \alpha)\big)\frac{\partial G_{t-1}}{\partial S_0} + I(t > \alpha)\frac{\partial G_{t-1}}{\partial S_0},\end{aligned} \tag{8.40}$$

where $A \cap B$ is the intersection of events A and B, and the derivative of the income level is then given by $\frac{\partial I_t}{\partial S_0} = (w-\mu)\cdot\frac{\partial G_t}{\partial S_0}$.

These recursions progress forward annually through the lifetime of the policy, with the initial conditions at time zero (annuitisation) of $\frac{\partial F_0}{\partial S_0} = P$, the policyholder premium, and $\frac{\partial I_0}{\partial S_0} = 0$, since the first withdrawal occurs at the end of year one. Substituting the values output from these recursions into Equation 8.37, via Equation 8.38, gives the pathwise delta estimator.

8.4.3 CLRM VA liability estimator

The second alternative approach for estimating option Greeks proposed earlier was the (conditional) likelihood ratio method (CLRM). We now outline how the CLRM can be used in estimating the VA liability sensitivities.

Firstly, we condition on a realisation of the Heston variance and CIR interest rate processes, in the same way as was performed in Section 8.3. The only difference here is that we only condition out to the first valuation date at the end of year one, since the VA contract liabilities are path-dependent (recall Example 2 in Section 7.1.3).

Therefore, we proceed by calculating

$$\bar{\sigma}_1 = \sqrt{\frac{a_{33}^2}{1}\int_0^1 V_t dt}, \quad \bar{\xi}_1 = \exp\left(\bar{Y}_1\right) \text{ and } \bar{r}_1 = \int_0^1 r_t dt, \tag{8.41}$$

where $\bar{Y}_1$ is given by

$$-\frac{1}{2}a_{31}^2\int_0^1 V_t dt - \frac{1}{2}a_{32}^2\int_0^1 V_t dt + a_{31}\int_0^1 \sqrt{V_t}dW_t^{I_1} + a_{32}\int_0^1 \sqrt{V_t}dW_t^{I_2}. \tag{8.42}$$

The integrals in the above terms can be approximated by simple numerical quadrature and using the relationship in Equation 8.10 (and the analogous relationship for the interest-rate CIR process). Then the (implied) shock out to year one is then given by:

$$Z^\star = \frac{\log(S_1/\bar{\xi}_1 S_0) - (\bar{r}_1 - \bar{\sigma}_1^2/2)\times 1}{\bar{\sigma}_1\sqrt{1}}. \tag{8.43}$$

Using this implied shock, the delta sensitivity can then be determined using

$$\Delta^{\text{LRM}} = \left(\frac{Z^\star}{S_0\bar{\sigma}_1\sqrt{1}}\right)\times\text{Liability} = \left(\frac{Z^\star}{S_0\bar{\sigma}_1\sqrt{1}}\right)\times\sum_{t=1}^{T} D_t p_t^{surv}\max(I_t - F_t, 0), \tag{8.44}$$

and similarly for the gamma sensitivity. The square-root of one is shown in the formula to help make the approach clear; it comes from the fact the first cashflow occurs at the end of year one. Thus, this term need not equal one if the product paid out semi-annual withdrawals, say, in which case this term would be $\sqrt{0.5}$.

8.4.4 VA liability gamma mixed estimator

In Chapter 7 it was found that the most efficient methods for calculating the gamma sensitivities were given by constructing the mixed estimators, which combined the respective advantages of the pathwise and LRM techniques. We will now apply the pathwise approach to the CLRM estimator for the delta of the liability in order to

derive a mixed gamma estimator for the stylised VA product:

$$\begin{aligned}
\Gamma^{\text{LR-PW}} &= \frac{\partial}{\partial S_0}\Delta^{\text{LRM}} \\
&= \frac{\partial}{\partial S_0}\left(\left(\frac{Z^\star}{S_0\bar{\sigma}_1\sqrt{1}}\right)\sum_{t=1}^{T} D_t p_t^{surv}\max(I_t - F_t, 0)\right) \\
&= \left(\frac{Z^\star}{S_0\bar{\sigma}_1\sqrt{1}}\right)\sum_{t=1}^{T} D_t p_t^{surv} I(I_t > F_t)\cdot\left(\frac{\partial I_t}{\partial S_0} - \frac{\partial F_t}{\partial S_0}\right) \\
&\quad -\left(\frac{Z^\star}{S_0^2\bar{\sigma}_1\sqrt{1}}\right)\sum_{t=1}^{T} D_t p_t^{surv}\max(I_t - F_t, 0). \qquad (8.45)
\end{aligned}$$

All the terms in this formula are already obtained in calculating the pathwise estimator of the liability delta (except the LRM weight which is easily found).

8.5 Comparison of VA liability estimators

Having derived the sophisticated delta and gamma estimators for the stylised VA product, we will now investigate the relative accuracy of these approaches. Five test cases will be considered, labelled A-E in Table 8.4, which give different parameter settings for the Heston and CIR processes and different correlations between the normal shocks driving the variance, interest rate and equity processes. In all these cases the contract has term $T = 30$ years and the ratchet term is the first 10 years of the product lifetime. The income drawn each year by the policyholder is 4% of the guarantee base and the initial policyholder premium is £10,000. For these tests two separate simulation set-ups were constructed; one for the bump and revalue and another for the pathwise and CLRM estimates.

In the bump and revalue set-up 36,000 paths were simulated. The perturbation size was set at 0.5%. There is a trade-off to be made here as the smaller the perturbation chosen the less bias in the estimate, however reducing the perturbation size will increase the variance of the estimator – particularly for the gamma sensitivity. In the CLRM/pathwise set-up 10,000 variance/interest rate outer paths were used, with 10 equity paths per outer realisation. In both frameworks 20 timesteps per year were used to discretise the Heston and CIR processes. With this set-up the bump and revalue and CLRM/pathwise approaches took approximately the same amount of time to run, providing a fair basis for the comparison of these approaches.

Case	κ_{V}	θ_{V}	ϵ_{V}	κ_{IR}	θ_{IR}	ϵ_{IR}	$\rho_{\text{S-V}}$	$\rho_{\text{S-IR}}$	$\rho_{\text{V-IR}}$
A	2	0.04	0.15	0.4	0.04	0.1	-0.7	-0.3	0.2
B	1	0.04	0.3	0.4	0.04	0.1	-0.7	-0.3	0.2
C	2	0.04	0.15	0.2	0.04	0.2	-0.7	-0.3	0.2
D	1	0.04	0.3	0.2	0.04	0.2	-0.7	-0.3	0.2
E	1	0.04	0.3	0.2	0.04	0.2	-0.9	-0.3	0.2

Table 8.4: Model settings considered in tests. Heston SV parameters denoted by V subscript, CIR parameters by IR subscript. $\mathrm{V}_0 = \theta_{\mathrm{V}}$, $\mathrm{r}_0 = \theta_{\mathrm{IR}}$.

	Sim. Set-up 1 (B&R)		Sim. Set-up 2 (PW/LRM)	
	Liab.(£)	St.Err	Liab.(£)	St.Err
A	106.99	0.69	106.24	1.17
B	127.00	0.99	125.36	1.74
C	159.21	1.03	157.56	1.82
D	171.62	1.29	168.74	2.29
E	177.42	1.55	174.80	2.85

Table 8.5: **VA liability estimates**: VA example product.

In table 8.5 the estimates of the liability value under the bump and revalue and CLRM/pathwise approaches are given. This table shows that under both simulation set-ups the values of liability are generally consistent. The bump and revalue approach does give estimates with smaller standard errors though. This is due to the conditional simulation framework required by the CLRM approach, which is less efficient than a standard simulation set-up in estimating the product liabilities, due to the reduced number of independent valuation paths. However, it is estimating the sensitivity of the liabilities that is the difficult challenge facing insurers.

Therefore, let us look at how the approaches perform in estimating the delta sensitivity of the liability in Cases A-E. In Table 8.6 and Figure 8.2 estimates of the delta sensitivities using each of the approaches discussed are given. Note that for the bump and revalue estimate, a central difference, using the 'bumped up' and 'bumped down' simulation paths, rather than a forward difference is used. This can help minimise levels of bias in this estimator. The gamma bump and revalue used the simulation analogue of a second-central difference. These results show that the bump and revalue and pathwise approaches give similar estimates and standard errors for the delta in each of the given cases. This is expected, as the pathwise estimator is essentially the small perturbation limit of the bump and and revalue approach. The reason the estimators are not converging to a common value is that

	B&R	St.Err	PW	St.Err	CLRM	St.Err
A	-0.00769	0.00013	-0.00752	0.00014	-0.00772	0.00034
B	-0.00375	0.00014	-0.00384	0.00018	-0.00381	0.00042
C	-0.00799	0.00016	-0.00796	0.00019	-0.00795	0.00045
D	-0.00359	0.00017	-0.00364	0.00022	-0.00355	0.00053
E	-0.00209	0.00019	-0.00200	0.00030	-0.00166	0.00074

Table 8.6: **Delta estimates**: VA example product.

	B&R	St.Err	CLRM	St.Err	PW-LR	St.Err
A	4.83	0.56	5.33	0.47	4.62	0.10
B	3.29	0.54	2.80	0.82	3.36	0.10
C	4.55	0.79	6.26	0.71	5.53	0.13
D	3.71	0.71	2.71	1.14	4.02	0.13
E	2.22	0.75	1.02	3.54	3.15	0.22

Table 8.7: **Gamma estimates**: VA example product. In units of 1×10^{-6}.

they are simulated under the different set-ups: the pathwise results were estimated under the conditional simulation set-up alongside the CLRM estimator, whereas the bump and revalue is from the base run of the bump and revalue set-up. Clearly the CLRM delta estimator is not as efficient as either the bump and revalue or pathwise equivalents. This can be explained by the fact that this method does not make use of the specific form of the payoff function, unlike the pathwise approach. This, however, means the LRM estimator can still be used for options where this payoff is discontinuous (and also for estimating second-order sensitivities). The pathwise estimator appears to be as efficient as the bump and revalue, but of course does not require a perturbed simulation run and is unbiased.

In Figure 8.2, the point estimates of the delta sensitivity under one approach do not always quite lie within the 95% confidence interval from a different approach. The reason for this is that these delta estimators are all biased to some degree. The first source of bias is in the discretisation of the variance and equity processes. Although the Andersen discretisation scheme should give smaller levels of bias than other simpler schemes, it will not eliminate it completely. If one wishes to have delta estimators with no bias whatsoever, employing the pathwise and CLRM estimators together with the exact Heston discretisation scheme of Broadie and Kaya [Bro06] will achieve this. The Broadie and Kaya sampling scheme will be much less efficient than that of Andersen. In addition to this discretisation bias, the bump and revalue

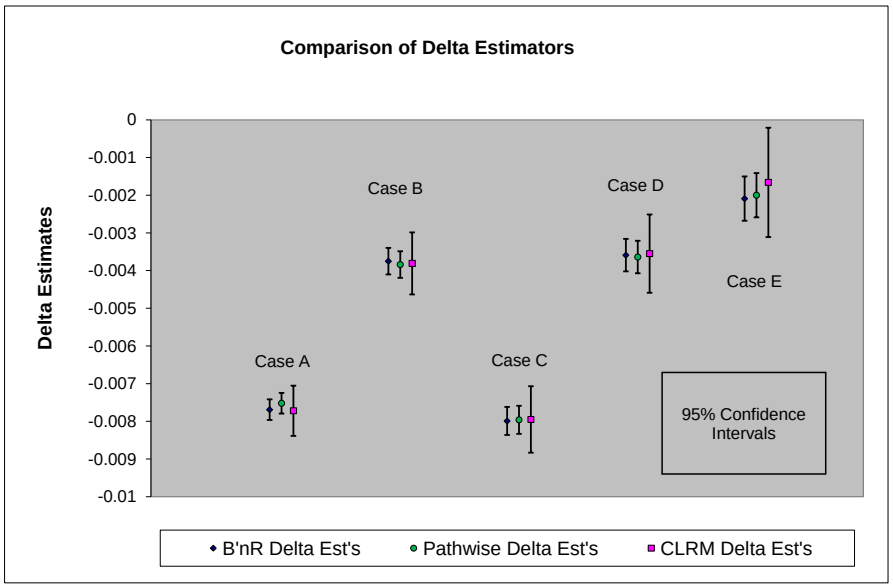

Figure 8.2: Comparison of VA delta estimators constructed.

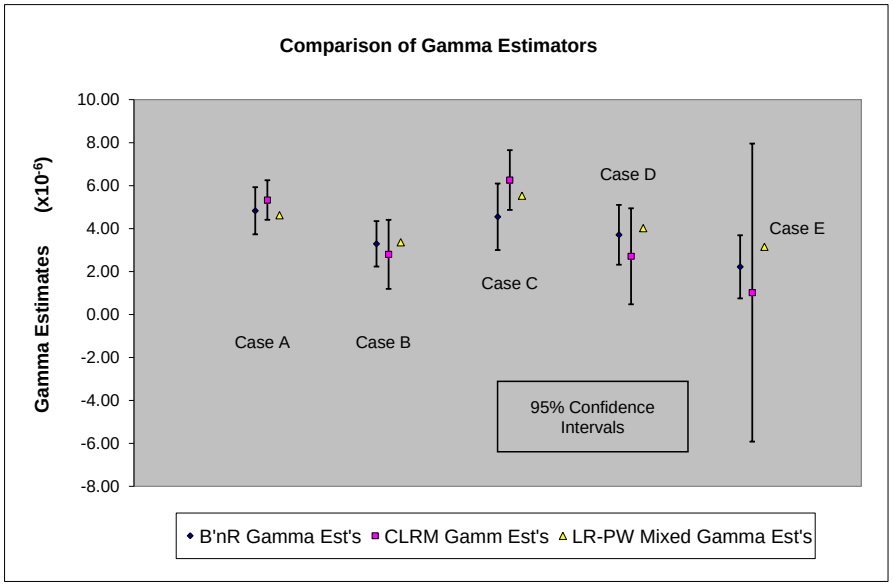

Figure 8.3: Comparison of VA gamma estimators constructed.

has a further source of bias from using a perturbation. This bias can be minimised by considering a smaller perturbation size, however if one uses a very small perturbation the variance of the sensitivity estimators can become very large. Therefore, avoiding this bias-variance trade-off is one advantage of the pathwise and CLRM approaches.

In Table 8.7 and Figure 8.3 the different estimators for the gamma sensitivity of the VA liability are compared. These results show that both the bump and revalue

and CLRM approaches give estimates with fairly large standard errors for cases A-E, with the CLRM being particularly poor for Case E due to the efficiency of this method deteriorating for high levels of $\rho_{\text{S-V}}$. The mixed PW-LR estimator gives estimates which are consistent with these two approaches, however they yield a much smaller standard error for each of the cases considered.

Thus, by constructing a mixed estimator for the gamma sensitivity, we have managed to gain a significant amount of reduction in variance, compared to the bump and revalue approach. Furthermore, this method is unbiased, hence we do not encounter the difficulty of choice of the perturbation step-size, where there is a trade-off between trying to reduce bias, whilst also trying to minimise the standard error of the estimates. The results found are consistent with the analysis of the mixed estimators being applied to European options given in Glasserman [Gla03].

This significant improvement in efficiency in estimating second-order sensitivities could be of great benefit to practitioners in managing a hedging strategy for their VA books. The simple bump and revalue approach cannot give both accurate and unbiased estimates of the gamma sensitivity. As a result, insurers adopting this approach will find it difficult to judge how frequently their hedge portfolio should be rebalanced, due to underlying market movements. With the accurate and unbiased gamma estimator developed in this thesis, practitioners will have a much greater appreciation of the convexity of their unit-linked insurance liabilities. This knowledge will improve the success of such hedging strategies and help mitigate some of the exposure inherent in selling these type of products.

8.6 Extension to VA liability vega sensitivities

In this section an idea of how vega sensitivity estimators might be constructed for the stylised variable annuity product will be given. This discussion will not fully derive such estimators, which is an area of further research, but will hopefully provide the interested reader with an overview of some relevant theory from the literature.

To begin with, we consider a vega pathwise estimator. In deriving the VA delta and gamma pathwise estimators, the stochastic process for the variance was of little concern; the stochastic volatility only had an indirect influence on these Greeks, through the equity index process. On the other hand, if we wish to estimate a

pathwise estimator with respect to the initial level of the variance, the specific form of the stochastic volatility model is going to be of much greater importance. Indeed, one quantity which will required in order to find the vega sensitivity of the liability is the term $\frac{dV_t}{dV_0}$. To illustrate how such a derivative can be estimated for the Heston stochastic volatility model we summarise the discussion given in Example 7.2.5 of Glasserman [Gla03]. Recall, the Heston model stochastic process for the variance

$$dV_t = \kappa_V(\theta_V - V_t)dt + \sigma_V\sqrt{V_t}dW_t^V. \tag{8.46}$$

By Proposition 5.2, we know V_t is distributed as a scaled non-central chi-squared random variable, with a non-centrality parameter which is proportional to V_0, i.e.,

$$V_t \sim c_1\chi_d'^2(c_2V_0), \tag{8.47}$$

where the form of c_1, c_2 and d is given in the aforementioned proposition. Glasserman shows that for $d > 1$, V_t can be simulated using

$$V_t = c_1\left(\left(Z + \sqrt{c_2V_0}\right)^2 + \chi_{d-1}^2\right) \tag{8.48}$$

with Z a standard normal random variable and χ_{d-1}^2 an ordinary chi-squared random variable with $d-1$ degrees of freedom and independent of Z. It then follows that

$$\frac{dV_t}{dV_0} = c_1c_2\left(1 + \frac{Z}{\sqrt{c_2V_0}}\right). \tag{8.49}$$

One can then simulate a variance path $V_1, V_2, \ldots$, through the recursion

$$\frac{dV_{n+1}}{dV_0} = c_1c_2\left(1 + \frac{Z_n}{\sqrt{c_2V_n}}\right)\frac{dV_n}{dV_0}, \tag{8.50}$$

where Z_n is the standard random normal shock used to generate V_{n+1} from V_n. Deriving $\frac{dV_t}{dV_0}$ for the case of $d \leq 1$ is not so straightforward, however for all the Heston model calibrations considered in this chapter and in Chapter 6, the degrees of freedom satisfy $d > 1$. It should be possible to combine the above recursion for $\frac{dV_t}{dV_0}$ with the variable annuity cashflows, to obtain a pathwise sensitivity for the liability with respect to V_0. In finding such sensitivities, one quantity which will be required is $\frac{dS_t}{dV_0}$. Computing such a value is fairly straight forward, however, given the values

of $\frac{dV_t}{dV_0}$. The following idea extends Example 7.2.3 of Glasserman [Gla03].

Discretising the equity asset process under the Heston-CIR model, S_n can be obtained in terms of S_{n-1} as follows

$$S_n = S_{n-1} \exp\left(\int_{t_{n-1}}^{t_n} r_u du - \frac{1}{2} \int_{t_{n-1}}^{t_n} V_u du + \int_{t_{n-1}}^{t_n} \sqrt{V_u} dW_u^S \right). \tag{8.51}$$

Appealing to the numerical quadrature formula for such integrals, given by Equation 5.25, the above expression can be approximated by

$$S_n = S_{n-1} \exp\left((\gamma_1 r_{n-1} + \gamma_2 r_n)\Delta t - \frac{1}{2}(\gamma_1 V_{n-1} + \gamma_2 V_n)\Delta t + (\gamma_1 \sqrt{V_{n-1}} + \gamma_2 \sqrt{V_n})\sqrt{\Delta t} Z_n \right), \tag{8.52}$$

where Z_n is the random shock taking the equity asset from time-point $n-1$ to n. Setting $\gamma_1 = \gamma_2 = 1/2$ yields a central quadrature approximation to these integrals. Differentiating Equation 8.52 using the chain rule, results in the following recursion for the derivative of S_n with respect to v_0:

$$\begin{aligned} \frac{dS_n}{dV_0} = \frac{dS_{n-1}}{dV_0} \cdot \frac{S_n}{S_{n-1}} + S_n \Bigg(& -\frac{1}{2}\gamma_1 \frac{dV_{n-1}}{dV_0} - \frac{1}{2}\gamma_2 \frac{dV_n}{dV_0} \\ & + \left(\frac{\gamma_1}{2\sqrt{V_{n-1}}} \cdot \frac{dV_{n-1}}{dV_0} + \frac{\gamma_2}{2\sqrt{V_n}} \cdot \frac{dV_n}{dV_0} \right) \sqrt{\Delta t} \cdot Z_n \Bigg), \end{aligned} \tag{8.53}$$

with the initial condition $dS_0/dV_0 = 0$. We have already considered how one could find the values of dV_n/dV_0, thus this recursion can be used to approximate the sensitivities of the equity value with respect to V_0. This expression could then be adapted to be consistent with the definition of the vega used earlier in the thesis, which was the sensitivity with respect to $\sigma_0 = \sqrt{V_0}$.

Now, let us discuss how the likelihood ratio method could be adapted to estimate a vega sensitivity. In order to employ this method, we require an explicit expression for the score function of the density of the equity returns. Under complex models, such as the Heston or Heston-CIR model, explicit expressions for the marginal or transition densities are not available. In order to construct likelihood ratio estimators for the delta and gamma sensitivities earlier in this chapter, the technique was employed conditionally, given a realisation of the variance and interest-rate processes. This allowed us to appeal to the simple form of the probability densities associated

with geometric Brownian motion. Such an approach cannot be used in estimating vega sensitivity, since this is a derivative with respect to one of the parameters of the Heston model. So, does this mean that the likelihood ratio method cannot be used to estimate vega under a Heston model?

When working with complex models, we typically have to simulate approximations to these models, using some sort of discretisation method, such as an Euler discretisation or the Andersen method. Glasserman argues that even though it may be impossible to construct a likelihood ratio estimator under some complex model, one may still be able to develop such an estimator for the approximating process. To illustrate this point Glasserman considers an Euler discretisation of the Heston model. Under this structure he is able to construct the score function corresponding to the sensitivity with respect to the σ_V parameter of the Heston model. The derivation will not be given here, but the interested reader can refer to Section 7.3.4 of Glasserman [Gla03] for more information. A similar approach could be possible in approximating the sensitivity of the VA liability with respect to V_0 or σ_0.

It should be noted that if one is only interested in determining the vega sensitivity of the liability, then a successful estimator should be given by simply applying the bump and revalue approach. In the analysis given in the previous section, it was found that the bump and revalue and pathwise estimators gave similar estimates and standard errors for the delta sensitivity. It was only in estimating second-order sensitivities, such as gamma, where the bump and revalue approach became problematic. Of course, for second-order sensitivities involving the vega sensitivity, the pathwise and likelihood ratio method frameworks could be of great benefit.

8.7 Conclusion

With the increasing popularity of VA products and the new Solvency II regulatory framework in Europe, employing effective hedging strategies for unit-linked insurance liabilities is a challenge currently facing insurers. The recent financial crisis has demonstrated that under turbulent market conditions a hedging portfolio can require more frequent rebalancing. The standard bump and revalue approach for estimating the Greeks used in such a strategy has some shortcomings, particularly for second-order sensitivities.

In this chapter some more advanced estimators for VA Greeks have been developed which are unbiased and do not require additional perturbed simulation runs. The mixed estimator developed for the gamma sensitivity also offers far greater efficiency in comparison to the bump and revalue method. This gain in efficiency will increase further as the number of Greeks required for a hedging strategy grows. Furthermore, the bias-variance trade-off in the choice of the perturbation size is avoided.

One interesting line of further research is to adapt the sensitivity calculations given in this chapter to be compatible with the automatic (or algorithmic) differentiation (AD) method of computation. The idea behind this approach is to use adjoint methods to re-arrange the algebraic operations which are used to determine the sensitivity. Using this alternative sequence of the calculations can offer large reductions in the amount of processing required, particularly in situations requiring the calculation of the sensitivities of a small number of outputs with respect to a large number of model input parameters. Thus, if an insurer wishes to determine the sensitivity of a liability to many risk-factors, adopting the AD approach in calculating the estimators given in this chapter could lead to great improvements in the efficiency of such computations. The AD approach has become increasingly popular in the field of financial mathematics in recent years. See Homescu [Hom11], and the articles cited within, for an extensive discussion of the application of the technique in a financial context.

Chapter 9

Conclusions of thesis

With the impending introduction of the Solvency II framework, we are at the dawn of a new age in insurance regulation in Europe. As such, many insurers face the challenge of effectively managing the risk arising from selling unit-linked products in the market. Therefore, novel techniques which can help such businesses plan for and react to adverse changes in market conditions will be of enormous practical benefit to analysts working in this sector. In this thesis, two main approaches which can help in this challenge were investigated.

The first of these approaches was the least-squares Monte Carlo (LSMC) simulation method. This technique, originally developed in the context of pricing American options, has been proposed as a means of approximating the distribution of complex insurance liabilities projected forward one year or more into the future. The natural Monte Carlo simulation approach for this task led to a nested simulation computational framework, which remains inefficient and biased, despite some recent articles in the literature investigating how to optimally allocate the computational budget in such a set-up. The LSMC approach approximates the liability at the projection year by regressing the reduced number of valuation paths simulated on some influential explanatory variables. This greatly reduces the computational effort required to obtain accurate estimates of the projected liability distribution.

In Chapter 3 of the thesis, some of the issues regarding the efficient configuration of the LSMC method were investigated. This analysis found that the stepwise AIC algorithm gave a reasonably good approach for selecting the regression model and one which is robust against over-fitting. It was also shown that sampling the fitting scenarios from the real-world distribution gave a good fit in the centre of the projected value distribution, but a poorer fit in the upper tail. On the other hand, sampling these from an alternative scheme, such as quasi-random sampling, gave a much better fit in the upper tail, with only a slightly poorer fit in the centre of the distribution compared with real-world sampling. Finally, evidence was given which suggested that for maximum efficiency in implementing the algorithm, only one

antithetic pair of inner valuation scenarios should be simulated, with the remaining computational budget being used to generate as large a number of outer fitting scenarios as possible.

In Chapter 4 these findings were put into practice when the LSMC method was used to estimate percentiles of the projected liability distribution for a realistic variable annuity life insurance product. In this analysis, it was found that the technique is successful in approximating the projected liability distribution for a fairly complex product. Furthermore, when it is possible to sample the fitting scenarios from a quasi-random scheme, this offers a superior estimate to percentiles in the upper tail of this distribution. An example of a situation where quasi-random sampling was not possible was also investigated. In this case, it was found that real-world sampling of the fitting points still gave a fairly accurate fit to the upper tail of this distribution.

In Chapter 6 semi-analytical formulae were derived for some simple unit-linked insurance guarantees under the Heston model. Such formulae could prove very useful in practice. The liabilities on these simple guarantees for which semi-analytical valuation is possible, are likely to be fairly highly correlated to the liabilities on more complex guarantees. Therefore, these simple insurance guarantees could potentially be successful control variates in increasing the accuracy of the LSMC estimates of the realistic projected liabilities considered in Chapter 4. Also, the semi-analytical valuation of the sensitivity of these simple guarantees could act as control variates for the sensitivity estimators for more complex liabilities. Investigating these candidate control variates is an interesting area for further research.

The second of the two main approaches investigated in this thesis was the construction of a hedging strategy for managing exposure to unit-linked products. Setting up such a strategy requires the calculation of market risk sensitivities (or 'Greeks') and for complex, path-dependent liabilities these sensitivities typically must be estimated using Monte Carlo simulation. Standard practice amongst many insurers is to measure such sensitivities using a 'bump and revalue' method. This is just the simulation analogue of a finite difference approximation, using standard and perturbed simulation paths. As well as requiring multiple valuations, this approach can be unreliable for higher order Greeks.

More sophisticated approaches have been developed in the literature for estimating option price sensitivities. In Chapter 7 the pathwise and likelihood ratio estimators and the mixed estimator, which combines these two approaches for second-order sensitivities, were extended for a Black-Scholes model featuring periodic fixed withdrawals. This variation of the Black-Scholes model was introduced in this thesis and begins to capture some of the features of unit-linked insurance products which provide a regular income stream to the annuitant. These estimators were then constructed for a stylised variable annuity product in Chapter 8. Firstly, the likelihood ratio method was extended to a Heston-CIR model, incorporating stochastic volatility and interest rates. This allowed the likelihood ratio estimators to be constructed for the stylised VA liability sensitivities under a more sophisticated economic model. The pathwise estimators were then developed for the VA product under this model and finally both estimators were combined to construct a mixed estimator for the second-order gamma sensitivity. It was found that these more advanced estimators for the VA Greeks are unbiased and avoid the need for additional perturbed simulation runs. Also, the mixed estimator developed for the VA liability gamma sensitivity gave much smaller standard errors in comparison to the bump and revalue method. This is an important finding, since it is in estimating the second-order sensitivities that this standard approach is particularly poor. A further advantage of these new estimators is that the bias-variance trade-off which must be made in the choice of the perturbation size in the bump and revalue approach is avoided. Also, the gain in efficiency increases as the total number of Greeks required for a hedging strategy grows.

To conclude this thesis, one final area of further research which could combine the analyses given in Parts I and II will be outlined. The basic idea is to determine the sensitivities of an insurance liability projected at a number of regular future time-points. This would proceed by applying the LSMC method at each time with the sensitivities, determined using the estimators developed in this thesis, acting as the response variables. Projecting these Greeks to many regular future time-points would allow us to determine the value of the actively managed hedge portfolio at each of these times. This could then be used to adjust the liability cashflows by offsetting them against the value of the hedging portfolio. Then, we could approximate a solvency capital requirement which takes into account a particular hedging strategy

being used by an insurer. Naturally, the greater the number of time-points at which the Greeks are projected, the more frequently the hedging strategy is assumed to be re-balanced by the insurer. Maintaining the computational efficiency of such a framework as the hedge re-balancing frequency is increased is one technical challenge which would need to be overcome for this approach to be practical.

With the development of novel techniques, insurance companies are in a much better position to understand the risks inherent in selling unit-linked insurance products and to manage such risks effectively. The research presented in this thesis will help insurers continue to meet this challenge and remain financially secure over the coming years.

Bibliography

[And07] Andersen, L. (2007), Efficient Simulation of the Heston Stochastic Volatility Model, Working paper.
Available at SSRN: `http://ssrn.com/abstract=946405`.

[Bae05] Baek, S., Karaman, F. and Ahn, H. (2005), Variable selection for heteroscedastic data through variance estimation, *Communications in Statistics - Simulation and Computation*, 34 (3), pp. 567-583.

[Bat96] Bates, D. (1996), Jumps and stochastic volatility: Exchange rate processes implicit in Deutsche mark options, *The Review of Financial Studies*, 9 (1), pp. 69-107.

[Bau10] Bauer, D., Bergmann, D. and Reuss, A. (2010), Solvency II and nested simulations – a least-squares Monte Carlo approach, Working paper.

[Bau11] Bauer, D., Bergmann, D. and Reuss, A. (2011), On the calculation of the solvency capital requirement based on nested simulations, Willis Research Network publication. Submitted to the *ASTIN Bulletin*.

[Bax96] Baxter, M. and Rennie, A. (1996), *Financial Calculus: An Introduction to Derivative Pricing*, Cambridge University Press.

[Bin04] Bingham, N. H. and Kiesel, R. (2004), *Risk-Neutral Valuation: Pricing and Hedging of Financial Derivatives*, Springer Finance.

[Bro11a] Broadie, M., Du, Y., Moallemi, C.C. (2011a), Efficient risk estimation via nested sequential simulation, *Management Science*, 57, pp. 1172-1194.

[Bro11b] Broadie, M., Du, Y. and Moallemi, C.C. (2011b), Risk estimation via regression, Working paper, Columbia University.

[Bro11c] Broadie, M., Du, Y. and Moallemi, C.C. (2011c), Risk estimation via weighted regression, *Proceedings of the 2011 Winter Simulation Conference*, pp. 3859-3870.

[Bro96] Broadie, M. and Glasserman, P. (1996), Estimating security price derivatives using simulation, *Journal of Economic Dynamics and Control*, 21, pp. 1323-1352.

[Bro97] Broadie, M. and Glasserman, P. (1997), Pricing American-style securities using simulation, *Journal of Economic Dynamics and Control*, 21, pp. 1323-1352.

[Bro04] Broadie, M. and Kaya, O. (2004), Exact Simulation of Option Greeks under Stochastic Volatility and Jump Diffusion Models, *Proceedings of the 2004 Winter Simulation Conference*, pp. 1607-1615.

[Bro06] Broadie, M. and Kaya, O. (2006), Exact Simulation of Stochastic Volatility and Other Affine Jump Diffusion Processes, *Operations Research*, 54 (2), pp. 217-231.

[Cai04] Cairns, A. (2004), *Interest rate Models: An Introduction*, Princeton University Press.

[Car99] Carr, P. and Madan, D. (1999), Option valuation using the fast Fourier transform, *Journal of Computational Finance*, 2, pp. 61-73.

[Car] Carr, P., et al. , Option pricing using integral transforms, Unpublished work for seminar. Available at:
`http://www.math.nyu.edu/research/carrp/papers/pdf`

[Car96] Carriere, J. (1996), Valuation of early-exercise price of options using simulations and nonparamteric regression, *Insurance: Mathematics and Economics*, 19, pp. 19-30.

[Cer05] Cerrato, M. and Cheung, K.K (2005), An empirical analysis on the convergence of Monte Carlo least squares estimators

[Clé02] Clément, E., Lamberton, D. and Protter, P. (2002). An analysis of a least squares regression method for American option pricing, *Finance and Stochastics*, 6 (4), pp. 449-471.

[Cox85] Cox, J.C., Ingersoll, J.E. and Ross, S.A. (1985), A Theory of the Term Structure of Interest Rates. *Econometrica*, 53, pp. 385-407.

[Duff97] Duffie, D. and Pan, J. (1997), An Overview of Value at Risk, *Journal of Derivatives*, 4 (3), pp. 7-49.

[Duff00] Duffie, D., Pan, J. and Singleton, K. (2000), Transform Analysis and Asset Pricing for Affine Jump Diffusions, *Econometrica*, 68, pp. 1343-1376.

[EMB10] EMB Solvency II document (2010), Solvency II: Understanding the directive. Available at `http://www.emb.com/EMBDOTCOM/UK/UK/ERM/Solvency%20II%20Brochure_FINAL-low%20res.pdf`

[EUSD] European Commission website: Background to the Solvency II Project. Available at: `http://ec.europa.eu/internal_market/insurance/solvency/background_en.htm`

[Gat06] Gatheral, J. (2006), *The Volatility Surface: A Practitioner's Guide*, Wiley Finance, John Wiley & Sons.

[Gla03] Glasserman, P. (2003), *Monte Carlo Methods in Financial Engineering*. Springer-Verlag, New York.

[Gla09] Glasserman, P. and Kim, K. (2009), Gamma Expansion of the Heston Stochastic Volatility Model, Working paper.

[Gla00] Glasserman, P. and Zhao, X. (2000), Arbitrage-free discretization of lognormal forward LIBOR and swap rate models, *Finance and Stochastics*, 4, 35-68.

[Gly87] Glynn, P.W. (1987), Likelihood ratio gradient estimation: an overview, *Proceedings of the Winter Simulation Conference*, pp. 366-374.

[Gor00] Gordy, M. B. and Juneja, S. (2010), Nested simulation in portfolio risk measurement, *Management Science*, 56 (10), pp. 1833-1848.

[Grz10] Grzelak, L. A. and Oosterlee, C. W. (2010), On the Heston model with stochastic interest rates, Munich Personal RePEc Archive.

[Hal09] Halley, W., Malham, S. and Wiese, A. (2009), Positive and Implied Stochastic Volatility Simulation, *Journal of Computational Finance*, To appear (received 2009).

[Har03] Hardy M.R. (2003), *Investment Guarantees: Modeling and Risk Management for Equity-Linked Life Insurance*, Wiley (New York).

[Hes93] Heston, S. (1993), A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options, *The Review of Financial Studies*, 6 (2), pp. 327-343.

[Hig04] Higham, D.J (2004), *An Introduction to Financial Option valuation: Mathematics, Stochastics and Computation*. Cambridge University Press.

[Ho83] Ho, Y.C. and Cao, X.R. (1983) Optimization and perturbation analysis of queuing networks, *Journal of Optimization Theory and Applications*, 40, pp. 559-582.

[Hob09] Hobbs, C., Krishnaraj, B., Lin, Y. and Musselman, J. (2009), Calculation of variable annuity market sensitivities using a pathwise methodology, *Life & Pensions*, September 2009.

[Hom11] Homescu, C. (2011), Adjoints and Automatic (Algorithmic) Differentiation in Computational Finance.
Available at SSRN: `http://ssrn.com/abstract=1828503`.

[Huy08] Huynh, H.T., Lai, V.S. and Soumaré, I. (2008), *Stochastic Simulation and Applications in Finance with MatLab Programs*, Wiley Finance, John Wiley & Sons.

[Jam11] James, J.F. (2011), *A student's guide to Fourier transforms: With applications in physics and engineering*, Third edition, Cambridge University Press (Cambridge).

[Joh95] Johnson, N., Kotz, S. and Balakrishnan, N. (1995), *Continuous Univariate Distributions*, Vol. 2, Wiley Interscience.

[Jun09] Juneja, S. and Ramprasath, L. (2009), Nested simulation for estimating portfolio losses within a time horizon, *Proceedings of the 2009 Winter Simulation Conference*, pp. 434-443.

[Kah05b] Kahl, C., Jäckel, P. (2005), Not-so-complex logarithms in the Heston model, *Wilmott Magazine*, September, 2005.

[Kah05a] Kahl, C., Jäckel, P. (2005), Fast strong approximation Monte-Carlo schemes for stochastic volatility models, Working Paper, ABN AMRO and University of Wuppertal.

[Kem90] Kemna, A. G. Z. and Vorst, A. C. F. (1990), A pricing method for options based on average asset values, *Journal of Banking and Finance*, 14, pp. 113-129.

[Kou11] Koursaris, A. (2011), A primer in replicating portfolios, Barrie and Hibbert Insights, July 2011.

[Lan07] Lan, H., Nelson, B.L and Staum, J. (2007), Two-level simulation for risk management, *Proceedings of the 2007 INFORMS Simulation Society Research Workshop*, pp. 102-107.

[Lan10a] Lan, H., Nelson, B.L and Staum, J. (2010a), A confidence interval procedure for expected shortfall risk measurement via two-level simulation, *Operations Research*, 58, pp. 1481-1490.

[Led10] Ledlie, M.C., et al. (2010), Variable Annuities, *British Actuarial Journal*, 14 (2), pp. 409-430.

[Lew01] Lewis, A. (2001), *Option Valuation Under Stochastic Volatility: with Mathematica Code*, Finance Press.

[Lew02] Lewis, A., The Mixing Approach to Stochastic Volatility and Jump Models, *Wilmott Magazine*, March 2002.

[Liu09] Liu, M. and Nelson, B.L (2009), Estimating expected shortfall with stochastic kriging, *Proceedings of the 2009 Winter Simulation Conference*, pp. 1249-1260.

[Liu10b] Liu, M., Nelson, B.L. and Staum, J. (2010b), An efficient simulation procedure for point estimation of expected shortfall, *Proceedings of the 2010 Winter Simulation Conference*, pp. 2821-2831.

[Lon01] Longstaff, F. and Schwartz, E. (2001), Valuing American options by simulation: A simple least-squares approach, *Review of Financial Studies*, 14, pp. 537-548.

[Lor08] Lord, R., Koekkoek, R. and Van Dijk, D. J. C. (2008), A Comparison of Biased Simulation Schemes for Stochastic Volatility Models. Tinbergen Institute Discussion Paper No. 06-046/4.
Available at SSRN: `http://ssrn.com/abstract=903116`

[Lor10] Lord, R. and Kahl, C. (2010), Complex Logarithms in Heston-Like Models, Forthcoming in *Mathematical Finance*.
Available at SSRN: `http://ssrn.com/abstract=1105998`

[Mat04] Matsuda, K. (2004), Introduction to Option Pricing with Fourier Transform: Option Pricing with Exponential Lévy Models, Part of a Ph.D. thesis which was filed on May 2006, Graduate School and University Center of the City University of New York.

[Mer76] Merton, R.C. (1976), Option pricing when underlying stock returns are discontinuous, *Journal of Financial Economics*, 3, pp. 125-144.

[McN11] McNeil, A.J. (2011), Introduction to the colloquium (slides), SFRA risk colloquium: Solvency II Overcoming the obstacles to success, September 2011. Available at: `http://www.sfra.ac.uk`

[McN05] McNeil, A.J., Frey, R. and Embrechts, P. (2005), *Quantitative Risk Management: Concepts, Techniques and Tools*, Princeton University Press.

[Mil06] Milevsky, M. and Salisbury, T.S. (2006), Financial valuation of guaranteed minimum withdrawal benefits, Insurance: Mathematics and Economics, 38, pp. 21-38.

[Oec07] Oechslin, J., et al. (2007), Replicating embedded options, *Life & Pensions*, February 2007.

[Pat49] Patnaik, P. (1949), The non-central χ^2- and F-distributions and their applications, *Biometrika*, 46, pp. 202-232.

[Pel11] Pelsser, A. (2011), Pricing in incomplete markets, Netspar panel papers, Panel paper 25.

[Rei89] Reimann, M. and Weiss, A. (1989), Sensitivity analysis for simulations via likelihood ratios, *Operations Research*, 37, pp. 830-844.

[Rom97] Romano, M. and Touzi, N. (1997), Contingent Claims and Market Completeness in a Stochastic Volatility Model, *Mathematical Finance*, 7 (4), pp. 399-412.

[Rud10] Rudolph, M. (2004), Actuaries look to avert potential annuity-related ALM crisis, *Risk*, October 2004.

[Sco96] Scott, L. O. (1996), Simulating a multi-factor term structure model over relatively long discrete time periods. *Proceedings of the IAFE First Annual Computational Finance Conference, Graduate School of Business, Stanford University.*

[Sco97] Scott, L. O. (1997), Pricing stock options in a jump-diffusion model with stochastic volatility and interest rates: Applications of Fourier invesrion methods, *Mathematical Finance*, 7, pp. 413-426.

[Ste03] Stentoft, L. (2004), Convergence of the least squares Monte Carlo approach to American option valuation, *Management Science*, 50 (9), pp. 1193-1203.

[Sur88] Suri, R. and Zazanis, M. (1988), Perturbation analysis gives strongly consistent sensitivity estimates for the M/G/1 queue, *Management Science*, 34, pp. 39-64.

[Tsi99] Tsitsiklis, J. and Van Roy, B. (1999), Optimal stopping of Markov processes: Hilbert space theory, approximation algorithms, and an application to pricing high-dimensional financial derivatives, *IEEE Transactions on Automatic Control*, 44, pp. 1840-1851.

[vHaa08] Van Haastrecht, A. and Pelsser, A. (2008), Efficient, Almost Exact Simulation of the Heston Stochastic Volatility Model, Working paper. Available at SSRN: http://ssrn.com/abstract=1131137

[VWP07] Variable Annuity Working Party (2007), Update for professional seminar presentation slides (2007), Faculty of Actuaries 28/09/07.

[Ven02] Venables, W.N. and Ripley, B.D. (2002), *Modern Applied Statistics with S*, Fourth Edition, Springer.

[Whi00] White, H. (2000), Asymptotic Theory for Econometricians: Revised Edition, Academic Press.

[Wil97] Willard, G.A. (1997), Calculating prices and sensitivities for path-independent derivative securities in multifactor models, *The Journal of Derivatives*, 5 (1), pp. 45-61.

[Wil00] Wilmott, P. (2000), *Paul Wilmott on quantitative finance*, John Wiley & Sons.

[Wüt10] Wüthrich, M.V., Bühlmann, H. and Furrer, H. (2010), *Market-consistent actuarial valuation*, Second revised and enlarged edition, Springer.

[Zan09] Zanger, D. (2009), Convergence of a least squares Monte Carlo algorithm for bounded approximating sets, *Applied Mathematical Finance*, 16 (2), pp. 123-150.

[Zhu08] Zhu, J. (2008), A Simple and Exact Simulation Approach to Heston Model, Working paper.
Available at SSRN: http://ssrn.com/abstract=1153950