Downloaded from orbit.dtu.dk on: Dec 20, 2017

Technical University of Denmark

Topics in Financial Engineering

Madsen, Claus; Pisinger, David; Poulsen, Rolf

Publication date:
2013

Document Version
Publisher's PDF, also known as Version of record

Link back to DTU Orbit

Citation (APA):
Madsen, C., Pisinger, D., & Poulsen, R. (2013). Topics in Financial Engineering. Kgs. Lyngby: Technical University of Denmark (DTU).

DTU Library
Technical Information Center of Denmark

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Topics in Financial Engineering

Claus Anderskov Madsen

January 2013

Preface

This thesis was prepared at DTU Management in partial fulfillment of the requirements for acquiring the Ph.D. degree.

The thesis deals with 3 different aspects of financial mathematical modeling problem, each addressing a separate problem.

1. Paper 1: A detailed analysis of the Quadratic time-dependent one-factor Interest-Rate model, both the numerical and analytical properties. A Monte-Carlo pricing method is designed and a discrete-time pricing procedure – along the lines of Hull and White (1993) – are being specified.
2. Paper 2: Design a framework for the capturing of the Risk embedded in Danish Mortgage Backed Bonds, that is both feasible from a calculation point of view and robust during extreme market movements. Test is performed during the Financial Crisis of 2008.
3. Paper 3: An empirical study on the use of stochastic programming (SP) models on mortgage choice and refinancing for Danish household. The focus is to perform a historical ex-ante test of the advice generated by our model framework for the period 1995-2010, and compare the results with the traditional rules of thumb for refinancing.

The thesis consists of a summary report and a collection of research papers. The last of these papers have at this point already been accepted for publication in an international journal within the area of Finance and Operations Research.

Papers included in the thesis

1. Paper 1: Claus Madsen (2012) “The normal class of arbitrage-free spot-rate models”, Technical Document at DTU Management, 2012
2. Paper 2: Claus Madsen (2012) “Empirical Yield-Curve Dynamics, Scenario Símulation and Risk-Measures”, Technical Document at DTU Management, 2013
3. Paper 3: Claus Madsen, Kourosh M. Rasmussen and Rolf Poulsen (2012) “Can Danish households benefit from Stochastic Programming models? – An empirical study of mortgage refinancing in Denmark”, accepted for publication in a special issue of the Journal of Computational Management Science

Acknowledgments

Initially, I wish to thank Associate Professor Kourosh Rasmussen and Professor Rolf Poulsen from University of Copenhagen who co-authored the third paper in this thesis.

Secondly Professor Rolf Poulsen from University of Copenhagen has been of a big help in the process of finishing the first 2 papers of this thesis.

Lastly, the support from my supervisor Professor David Pisinger has been of great importance in finishing this thesis.

I dedicate this thesis to my wife Lene and to my four children Nicolai, Mads, Ella and Anne Sofie.

Part I

Summary Report

Chapter 1: Introduction – a brief history of Mathematical Finance

The “time value of money” and uncertainty are the central elements that influence the value of financial instruments. When only the time aspect of finance is considered, the tools of calculus and differential equations are adequate. When only the uncertainty is considered, the tools of probability theory illuminate the possible outcomes. When time and uncertainty are considered together we begin the study of mathematical finance.

Finance is the study of economic agents’ behavior in allocating financial resources and managing risks across alternative financial instruments and in time in an uncertain environment. Known examples of financial instruments are bank accounts, loans, stocks, government bonds and corporate bonds.

Economic agents are units who buy and sell financial resources in a market, from individuals to banks, businesses and hedge funds etc. Each agent has many choices of where to buy, sell, invest and consume assets, each with advantages and disadvantages. Each agent must distribute his resources among the many possible investments with a particular goal in mind.

Mathematical finance is often characterized as the study of the more sophisticated financial instruments called derivatives. A derivative is a financial agreement between two parties that depends on something that occurs in the future, for example the price or performance of an underlying asset. The underlying asset could for example be a stock, a bond or a currency.

Derivatives have become one of the financial world’s most important risk-management tools.

Finance is about shifting and distributing risk and derivatives are a very efficient vehicle for that purpose. Two such instruments are futures and options[1]. Both futures and options are called “derivatives” and are types of instruments that have been around for quite some time. In 1865, futures on products such as grain, copper, and pork bellies were sold on the Chicago Board of Trade and options were a feature of the “tulip mania” in seventeenth century Holland. These instruments are called derivatives not because they involve a rate of change, but because their value is derived from some underlying asset.

Derivatives come in many types. For example, there are futures, agreements to trade something at a set price at a given dates (plus some “running netting” - margin); options, the right but not the obligation to buy or sell at a given price; forwards, like futures but traded

---

directly between two parties instead of on exchanges; and swaps, exchanging flows of income from different investments to manage different risk exposure. For example, one party in a deal may want the potential of rising income from a loan with a floating interest rate, while the other might prefer the predictable payments ensured by a fixed interest rate. This elementary swap is known as a “plain vanilla swap”.

More complex swaps mix the performance of multiple income streams with varieties of risk (see O’Harrow (2010)). Another more complex swap is a creditdefault swap in which a seller receives a regular fee from the buyer in exchange for agreeing to cover losses arising from defaults on the underlying loans. These swaps are somewhat like insurance. These more complex swaps are the source of controversy since many people believe that they are responsible for the collapse or near-collapse of several large financial firms in late 2008 (for example Lehmann Brothers and AIG)

Derivatives can be based on pretty much anything as long as two parties are willing to trade risks and can agree on a price. Businesses use derivatives to shift risks to other firms, chiefly banks. About 95% of the world’s 500 biggest companies use derivatives. Derivatives with standardized terms are traded in markets called exchanges. Derivatives tailored for specific purposes or risks are bought and sold “over the counter” from big banks. The “over the counter” market dwarfs the exchange trading. The Bank for International Settlements put the amount of outstanding of over-the-counter (OTC) derivatives at $647.76 billions as of 1. December 2011.

Mathematical models in modern finance contain deep and one might even say beautiful applications of differential equations and probability theory. In spite of their complexity, mathematical models of modern financial instruments have had a direct and significant influence on finance practice.

Early history

The origins of much of the mathematics in financial models traces back to Louis Bachelier’s 1900 dissertation on the theory of speculation in the Paris markets. Completed at the Sorbonne in 1900, this work marks the twin births of both the continuous time mathematics of stochastic processes and the continuous time economics of option pricing. While analyzing option pricing, Bachelier provided two different derivations of the partial differential equation for the probability density for the Wiener process or Brownian motion.

In one of the derivations, he works out what is now called the Chapman-Kolmogorov convolution probability integral. Along the way, Bachelier derived the method of reflection to solve for the probability function of a diffusion process with an absorbing barrier. Not a bad at all for a thesis on which the first reader, Henri Poincaré, gave less than a top mark! After Bachelier, option pricing theory lay dormant in the economics literature for over half a century until economists and mathematicians renewed study of it in the late 1960s. Jarrow and Protter (2004) speculate that this may have been because the Paris mathematical elite scorned economics as an application of mathematics.

Bachelier’s work was 5 years before Albert Einstein’s 1905 discovery of the same equations for his famous mathematical theory of Brownian motion. The editor of Annalen der Physik received Einstein’s paper on Brownian motion on May 11, 1905. The paper appeared later that year. Einstein proposed a model for the motion of small particles with diameters on the order of 0.001 mm suspended in a liquid. He predicted that the particles would undergo microscopically observable and statistically predictable motion. The English botanist Robert Brown had already reported such motion in 1827 while observing pollen grains in water with a microscope. The physical motion is now called Brownian motion in honor of Brown’s description. Einstein calculated a diffusion constant to govern the rate of motion of suspended particles. The paper was Einstein’s attempt to convince physicists of the molecular and atomic nature of matter. Surprisingly, even in 1905 the scientific community did not completely accept the atomic theory of matter. In 1908, the experimental physicist Jean-Baptiste Perrin conducted a series of experiments that empirically verified Einstein’s theory. Perrin thereby determined the physical constant known as Avogadro’s number for which he won the Nobel prize in 1926. Nevertheless, Einstein’s theory was very difficult to rigorously justify mathematically. In a series of papers from 1918 to 1923, the mathematician Norbert Wiener constructed a mathematical model of Brownian motion. Wiener and others proved many surprising facts about his mathematical model of Brownian motion, research that continues today. In recognition of his work, his mathematical construction is often called the Wiener process (see Jarrow and Protter (2004)).

Growth of Mathematical Finance

The first influential work of mathematical finance is the theory of portfolio optimization by Harry Markowitz on using mean-variance estimates. Markowitz assumes mean/expected return and variance to be known of portfolios to judge investment strategies. Using a linear regression strategy to understand and quantify the risk (i.e. variance) and return (i.e. mean) of an entire portfolio of stocks and bonds, an optimization strategy was used to choose a portfolio with largest mean return subject to acceptable levels of variance in the return. Simultaneously, William Sharpe (1964) developed the mathematics of determining the correlation between each stock and the market (CAPM). Put more precisely/accurately the CAPM is an equilibrium argument that links a stocks correlation with the market to its expected return. For their pioneering work, Markowitz and Sharpe, along with Merton Miller, shared the 1990 Nobel Memorial Prize in Economic Sciences, for the first time ever awarded for a work in finance.

What we today call modern mathematical finance theory begins in the 1960s. In 1965 the economist Paul Samuelson published two papers that argue that stock prices fluctuate randomly. One explained the Samuelson and Fama efficient markets hypothesis that in a well-functioning and informed capital market, asset-price dynamics are described by a model in which the best estimate of an asset’s future price is the current price (possibly adjusted for a fair expected rate of return).

Under this hypothesis, attempts to use past price data or publicly available forecasts about economic fundamentals to predict security prices are doomed to failure. In the other paper with mathematician Henry McKean, Samuelson shows that a good model for stock price movements is geometric Brownian motion. Samuelson noted that Bachelier’s model failed to ensure that stock prices would always be positive, whereas geometric Brownian motion avoids this error.

The next major revolution in mathematical finance in terms of practice was the 1973 Black-Scholes model for option pricing. The two economists Fischer Black and Myron Scholes (and simultaneously, and somewhat independently, the economist Robert Merton) deduced an equation that provided the first strictly quantitative model for calculating the prices of options. The key variable is the volatility of the underlying asset. These equations standardized the pricing of derivatives in quantitative terms. The formal press release from the Royal Swedish Academy of Sciences announcing the 1997 Nobel Prize in Economics states that the honor was given "*for a new method to determine the value of derivatives. Robert C. Merton and Myron S. Scholes have, in collaboration with the late Fischer Black developed a pioneering formula for the valuation of stock options. Their methodology has paved the way for economic valuations in many areas. It has also generated new types of financial instruments and facilitated more efficient risk management in society*."

For a strongly differing view, see Haug & Taleb (2011). They write the following: *"…the development of scholastic finance appears to be an epiphenomenon rather than a cause of option trading. Once again, lecturing birds how to fly does not allow one to take subsequent credit.*

This is why we call the equation Bachelier-Thorp. We were using it all along and gave it the wrong name, after the wrong method and with attribution to the wrong persons. It does not mean that dynamic hedging is out of the question; it is just not a central part of the pricing paradigm.."

I do not agree completely with Haug and Taleb, though they do have some interesting points, but to claim that the Black-Scholes formula is not used is like going to the extreme!

The Chicago Board Options Exchange (CBOE) began publicly trading options in the United States in April 1973, a month before the official publication of the Black-Scholes model. By 1975, traders on the CBOE were using the model to both price and hedge their options positions.

The basic insight underlying the Black-Scholes model is that a dynamic portfolio trading strategy in the stock can replicate the returns from an option on that stock. This is called "hedging an option" and it is the most important idea underlying the Black-Scholes-Merton approach.

The story of the development of the Black-Scholes-Merton option pricing model is that Black started working on this problem by himself in the late 1960s. His idea was to apply the capital asset pricing model (from Sharpe (1964)) to value the option in a continuous time setting. Using this idea, the option value satisfies a partial differential equation. Black could not find the solution to the equation. He then teamed up with Myron Scholes who had been thinking about similar problems. Together, they solved the partial differential equation using a combination of economic intuition and earlier pricing formulas.

At this time, Myron Scholes was at MIT, so was Robert Merton, who was applying his mathematical skills to various problems in finance. Merton (1973) was the first to call the solution the Black-Scholes option pricing formula. Merton's derivation used the continuous time construction of a perfectly hedged portfolio involving the stock and the call option together with the notion that no arbitrage opportunities exist.

Ironically Merton & Scholes were partners in Long Term Credit Management, the Greenwich, Connecticut-based hedge fund that almost single-handedly triggered a worldwide credit crisis in the late 1990's. One could say, sometimes even the brilliant misjudge risk.

In the 1970s and early 1980s mathematicians Harrison, Kreps and Pliska derived a more abstract formulation for the valuation of a derivative using the concept of martingales and showed that this provides greater generality.

A sometimes overlooked major contribution in mathematical finance is the introduction of duration. Hicks published Value and Capital in 1939, 1 year after Macaulay's book appeared. Hicks defined and used "an elasticity [of a capital value] with respect to a discount ratio [i.e., factor]" that is equivalent to Macaulay's duration. Hicks called his measure "average period." Hicks used his measure to make concrete the intuitive notion that, when interest rates fall, producers will substitute money (or the capital it can buy) for other means of production and that the average period of production plans increases.

Macaulay (1938) wanted a measure of time; Hicks, an elasticity. They derived the same measure. Hicks noted that, although elasticities are ordinarily "pure" numbers, this particular one had dimension time, and he explained why. Grove (1966) appears to have been the first to cite both Macaulay and Hicks. Fisher (1966) at that time unaware of Hicks's interest elasticity, showed that Macaulay's measure had the properties of an elasticity. Not all writers who use duration or duration-like measures cite any authority.

In 1945, Samuelson then unaware of Macaulay's work, analyzed the effect of interest rate changes on institutions such as universities, insurance companies, and banks. He developed a measure-the "weighted average time period of payments"-essentially equivalent to duration, and proved, that if the duration of an institution's assets is larger (smaller) than that of its liabilities, then the institution will lose (profit) when interest rates rise and profit (lose) when interest rates fall.

In 1952, Redington, in a paper not well known by (many) economists, defined the mean term of an asset stream and of a liability stream. He proved that the profits of an insurance company were immune (could not be reduced but might be increased) to any small (infinitesimal) change in interest rates provided that the mean term of assets equaled the mean term of liabilities. Wallas (1960) provide an expanded exposition of Redington's brilliant contribution.

Hopewell and Kaufman (1973) were the first to note the relationship between bond-price volatility and duration – whereas this relationship is used widely in connection with stochastic modeling of the yield-curve.

A few very important later contributions in mathematical finance is Vasicek (1977). The Vasicek model is a mathematical model describing the evolution of interest rates. It is a type of "one-factor model" (more precisely, one factor short rate model) as it describes interest rate movements as driven by only one source of market risk. The model can be used in the valuation of interest rate derivatives, and has also been adapted for credit markets – though less irrelevant, as it is normal.

Vasicek's model was the first one to capture mean reversion, an essential characteristic of the interest rate process that sets it apart from other financial prices. Thus, as opposed to stock prices for instance, interest rates cannot rise indefinitely. This is because at very high levels they would hamper economic activity, prompting a decrease in interest rates. Similarly, interest rates cannot (given recent time, this view might have to be relaxed) decrease (significantly) below 0 (zero). As a result, interest rates move in a limited range, showing a tendency to revert to a long run value.

The drift in Vasiceks model is expressed as: $dr(t) = a(b-r(t))dt$. The drift factor represents the expected instantaneous change in the interest rate at time t. The parameter b represents the long run equilibrium value towards which the interest rate reverts. Indeed, in the absence of shocks, the interest remains constant when $r(t) = b$, if it starts at b. The parameter a, governing the speed of adjustment, needs to be positive to ensure stability around the long term value. For example, when r(t) is below b, the drift term becomes positive for positive a, generating a tendency for the interest rate to move upwards (toward equilibrium).

The main disadvantage is that, under Vasicek's model, it is theoretically possible for the interest rate to become negative, an (normally) undesirable feature. This shortcoming was fixed in notable the Cox–Ingersoll–Ross (1985) model, among many others. The Vasicek model was further extended in the Hull–White (1990) model. The Vasicek model is also a canonical example of the affine term structure model, along with the Cox–Ingersoll–Ross model. Extension to multiple factors was originally and very generally explained in the classic work of Langetieg (1980) (later extended and formalized in Duffie and Kan (1996)).

A major contribution in the area of yield-curve modeling is the HJM framework which originates from the work of David Heath, Robert A. Jarrow and Andrew Morton in the late 1980s. Notable the Ph.d dissertation of Andrew Morton from (1988) which was the forerunner for their ground breaking article in 1992. It however has its critics, Paul Wilmott[2] (2009) describing it as "...actually just a big rug for [mistakes] to be swept under". What Wilmott refers to is the following, put in his own words: "…*Developed in the late 1980s, the formula looks horrifically complicated to the layman. But to a mathematician it's elegant, simple—and dangerous. Behind its simplicity lie hidden mistakes, unobservable variables like volatility and correlation that can provide a false sense of security. "With models, you want to see where things go wrong," says Wilmott. "You want to see the dirt….*"

The key to the HJM techniques is the recognition that the drifts of the no-arbitrage evolution of certain variables can be expressed as functions of their volatilities and the correlations among themselves. In other words, no drift estimation is needed. Models developed according to the HJM framework are different from the so called short-rate models in the sense that HJM-type models capture the full dynamics of the entire forward rate curve, while the short-rate models only capture the dynamics of a point on the curve (the short rate).

---

[2] Paul Wilmott is a 52 year-old Oxford-trained mathematician and arguably the most influential quant today. The Financial Times calls him a "cult derivatives lecturer." Nassim Taleb, the mathematician and author of the bestseller The Black Swan, calls him the smartest quant in the world. "*He's the only one who truly understands what's going on ... the only quant who uses his own head and has any sense of ethics,*" says Taleb.

However, models developed according to the general HJM framework are often non-Markovian and can even have infinite dimensions. A number of researchers have made great contributions to tackle this problem. They show that if the volatility structure of the forward rates satisfy certain conditions, then an HJM model can be expressed entirely by a finite state Markovian system, making it computationally feasible, see for example Ritchken and Sankarasubramanian (1995) and Bhar and Chiarella (1997).

An extension of HJM model - but specified in the same spirit - is the work of Brace, Gatarek and Musiela (BGM) (1997) – this extension is normally termed the LIBOR market model (or sometimes just the The Market Model). The BGM model is a financial model of interest rates. It is today used for the pricing of foremost interest rate derivatives, especially exotic derivatives like for example Bermudan swaptions, and ratchet caps and floors. The quantities that are modeled, rather than the short rate or instantaneous forward rates (like in the Heath-Jarrow-Morton framework) are a set of forward rates (also called forward LIBORs), which have the advantage of being directly observable in the market, and whose volatilities are naturally linked to traded contracts. Each forward rate is modeled by a lognormal process under its forward measure, i.e. a Black (1976) model leading to a Black formula for interest rate caps (this equation was independently derived by Miltersen, Sandmann and Sondermann (MSS) (1997)[3] . This formula is the market standard to quote cap prices in terms of implied volatilities, hence the term "market model". In connection with the interest modeling it is impossible to get around Jamshidian, who is arguably one of the main pioneers. For example can mentioned the case of pricing using forward induction, see Jamshidian (1991). See also references in: http://wwwhome.math.utwente.nl/~jamshidianf/.

The LIBOR market model may be interpreted as a collection of forward LIBOR dynamics for different forward rates with spanning tenors and maturities, each forward rate being consistent with a Black interest rate caplet formula for its canonical maturity. One can write the different rates dynamics under a common pricing measure, for example the forward measure for a preferred single maturity, and in this case forward rates will not be lognormal under the unique measure in general, leading to the need for numerical methods such as monte carlo simulation or approximations like the frozen drift assumption.

The area of (Financial) Risk-Management

If you went to work this morning, you took a risk. If you rode your bicycle, walked, or drove a car, you took a risk. If you put your money in a bank, or in stocks, or under a mattress, you took other types of risk. If you bought a lottery ticket at the newsstand or gambled at a casino over the weekend, you were engaging in activities that involve an element of chance – something intimately connected with risk. The word risk has its roots in the old French word risqué, which means “danger, in which there is an element of chance” (Littré, 1863). The word hazard, another term integral to discussions of risk management, comes from a game of chance invented at a castle named Hasart, in Palestine, while it was under siege (Oxford English Dictionary, 1989).

---

[3] It is interesting to note, that the formula derived by MSS, was actually - though in a completely different setup – already derived by Rady and Sandmann (1994).

In his 1998 book Against the Odds, Peter Bernstein describes how thinking about risk evolved in part because of changes in mathematical numbering systems, an understanding of the statistical basis of probability, and the rise in popularity of gambling. Although games of chance and gambling were depicted in Egyptian tomb paintings from 3500 B.C.E., it wasn't until the Renaissance that a "scientific" or statistical basis for gambling was presented. This was because the Hindu-Arabic numbering system (the numerals 1, 2, 3, and so forth) appeared in Europe between 1000–1200 C.E. allowing calculations beyond simple addition and subtraction to be performed. It wasn't until the Renaissance, however, that the ten digits – 0 to 9 – that we take for granted fully replaced the more clumsy Roman numerals.

Financial Risk-Management as a discipline first really kicked off when J.P. Morgan (1992) (published in 1996 by Zangari) introduced the ground-breaking RiskMetrics™ and the concepts of DEAR, which is now named VaR – Value-at-Risk[4].

We define VaR as a category of probabilistic measures of market risk. Consider a portfolio with fixed holdings. Its current market value is known. Its markets value at some future time—say one day or one month in the future—is a random variable. As a random variable, we may ascribe it a probability distribution. A VaR metric is a function of:

1. that distribution and
2. the portfolio's current market value

With this definition, variance of return, standard deviation of P&L and .95-quantile of loss are all VaR metrics. We define a VaR measure as any procedure that, given a VaR metric, assigns values for that metric to portfolios.

What justifies our interpreting the output of those computations as, say, two-week 99% EUR VaR? The answer is the VaR model. The VaR model is the intellectual link between the computations of a VaR measure and the interpretation of the output of those computations, which is the VaR metric.

Early VaR measures developed along two parallel lines. One was portfolio theory, and the other was capital adequacy computations.

Markowitz (1999) have documented the history of VaR measures in the context of portfolio theory, whereas the development of VaR measures in the context of capital adequacy computations is documented in Holten (2002).

Since the introduction of VaR a lot of discussion have centered around the following key areas:

1. The distribution assumption
2. How to estimate volatility and indeed how to define it
3. How to measure the correlation or more generally, co-dependence (causality)

---

[4] Originally however, the concept of risk-management originated from the pioneering work of Markowitz (1952).

This debate is still an on-going debate, which probably never will end. After the financial crisis of 2008 new areas of how to look at risk have entered on the scene, where one very interesting direction is the use of Bayesian Networks, see Fenton and Neil (2012).

A major contribution which does not focus on any of the 3 points mentioned above but instead of the VaR measure itself is the highly technical article by Artzner, Delbaen, Eber and Heath (2001). They showed that VaR understood as a percentile is not a coherent risk measure. A coherent risk measure is a function ρ that satisfies properties of monotonicity, sub-additivity, homogeneity, and translational invariance. They specified the requirements for being part of the family of coherent risk measures and showed that the most simple candidate is ETL (expected tail-loss) – also termed CVaR (Conditional Value-of-Risk).

A practical example that illustrates this is the following:

- "*We have 2 identical bonds, A and B. Each default with probability 4%, and we get a loss of 100 if default occurs, and a loss of 0 if no default occurs. The 95% VaR for each bond is therefore 0, so VaR(A) = VaR(B) = VaR(A) + VaR(B) = 0. Let us now suppose that defaults are independent. Elementary calculations now establish that we get a loss of 0 with probability 0.962 = 0.9216, a loss of 200 with probability 0.042 = 0.0016, and a loss of 100 with a probability of 1 – 0.9216 – 0.0016 = 0.0768. That is VaR(A+B) = 100. Thus VaR(A+B) = 100 > 0 = VaR(A) + VaR(B), and the VaR violates the sub-additivity property – hence VaR is not a Coherent Risk Measure!*"

The next stage

In the past, mathematical models had a limited impact on finance practice. But since 1973 these models have become central in markets around the world. In the future, mathematical models are likely to have an indispensable role in the functioning of the global financial system including regulatory and accounting activities.

We need to seriously question the assumptions that make models of derivatives work: the assumptions that the market follows probability models and the assumptions underneath the mathematical equations. But what if markets are too complex for mathematical models? What if irrational and completely unprecedented events do occur, and when they do – as we know they do – what if they affect markets in ways that no mathematical model can predict? What if the regularity that all mathematical models assume ignores social and cultural variables that are not subject to mathematical analysis?

Any virtue can become a vice if taken to extreme, and just so with the application of mathematical models in finance practice. At times, the mathematics of the models becomes too interesting and we lose sight of the models' ultimate purpose. Futures and derivatives trading depend on the belief that the stock market behaves in a statistically predictable way; in other words, that probability distributions accurately describe the market. The mathematics is precise, but the models are not, being only approximations to the complex, real world. The practitioner should apply the models only tentatively, assessing their limitations carefully in each application. The belief that the market is statistically predictable (meaning that we have a well-specified stochastic model) drives the mathematical refinement, and this belief inspires derivative trading to escalate in volume every year.

Financial events since late 2008 show that many of the concerns of the previous paragraphs have occurred. In 2009, Congress and the Treasury Department considered new regulations on derivatives markets. Complex derivatives called credit default swaps appear to have been based on faulty assumptions that did not account for irrational and unprecedented events, as well as social and cultural variables that encouraged unsustainable borrowing and debt. Extremely large positions in derivatives that failed to account for unlikely events caused bankruptcy for financial firms such as Lehman Brothers and the collapse of insurance giants like AIG. The causes are complex, but some of the blame has been fixed on the complex mathematical models and the people who created them. This blame results from distrust of that which is not understood. Understanding the models is a prerequisite for correcting the problems and creating a future which allows proper risk management.

Wilmott and Derman (2009) as a reflection over the financial crisis in 2008 wrote what they termed the “Financial Modeler’s Manifesto”. It is a fitting end of the history of mathematical finance so for that reason I have included the full manifest here:

Manifesto

“In finance we study how to manage funds – from simple securities like dollars and yen, stocks and bonds to complex ones like futures and options, subprime CDOs and credit default swaps. We build financial models to estimate the fair value of securities, to estimate their risks and to show how those risks can be controlled. How can a model tell you the value of a security? And how did these models fail so badly in the case of the subprime CDO market? Physics, because of its astonishing success at predicting the future behavior of material objects from their present state, has inspired most financial modeling. Physicists study the world by repeating the same experiments over and over again to discover forces and their almost magical mathematical laws. Galileo dropped balls off the leaning tower, giant teams in Geneva collide protons on protons, over and over again. If a law is proposed and its predictions contradict experiments, it's back to the drawing board. The method works. The laws of atomic physics are accurate to more than ten decimal places.

It's a different story with finance and economics, which are concerned with the mental world of monetary value. Financial theory has tried hard to emulate the style and elegance of physics in order to discover its own laws. But markets are made of people, who are influenced by events, by their ephemeral feelings about events and by their expectations of other people's feelings. The truth is that there are no fundamental laws in finance. And even if there were, there is no way to run repeatable experiments to verify them.

You can hardly find a better example of confusedly elegant modeling than models of CDOs. The CDO research papers apply abstract probability theory to the price co-movements of thousands of mortgages. The relationships between so many mortgages can be vastly complex. The modelers, having built up their fantastical theory, need to make it useable; they resort to sweeping under the model's rug all unknown dynamics; with the dirt ignored, all that's left is a single number, called the default correlation. From the sublime to the elegantly ridiculous: all uncertainty is reduced to a single parameter that, when entered into the model by a trader, produces a CDO value. This over-reliance on probability and statistics is a *severe limitation. Statistics is shallow description, quite unlike the deeper cause and effect of physics, and can't easily capture the complex dynamics of default.*

Models are at bottom tools for approximate thinking; they serve to transform your intuition about the future into a price for a security today. It's easier to think intuitively about future housing prices, default rates and default correlations than it is about CDO prices. CDO models turn your guess about future housing prices, mortgage default rates and a simplistic default correlation into the model's output: a current CDO price.

Our experience in the financial arena has taught us to be very humble in applying mathematics to markets, and to be extremely wary of ambitious theories, which are in the end trying to model human behavior. We like simplicity, but we like to remember that it is our models that are simple, not the world.

Unfortunately, the teachers of finance haven't learned these lessons. You have only to glance at business school textbooks on finance to discover stilts of mathematical axioms supporting a house of numbered theorems, lemmas and results. Who would think that the textbook is at bottom dealing with people and money? It should be obvious to anyone with common sense that every financial axiom is wrong, and that finance can never in its wildest dreams be Euclid. Different endeavors, as Aristotle wrote, require different degrees of precision. Finance is not one of the natural sciences, and its invisible worm is its dark secret love of mathematical elegance and too much exactitude.

We do need models and mathematics – you cannot think about finance and economics without them – but one must never forget that models are not the world. Whenever we make a model of something involving human beings, we are trying to force the ugly stepsister's foot into Cinderella's pretty glass slipper. It doesn't fit without cutting off some essential parts. And in cutting off parts for the sake of beauty and precision, models inevitably mask the true risk rather than exposing it. The most important question about any financial model is how wrong it is likely to be, and how useful it is despite its assumptions. You must start with models and then overlay them with common sense and experience.

Many academics imagine that one beautiful day we will find the 'right' model. But there is no right model, because the world changes in response to the ones we use. Progress in financial modeling is fleeting and temporary. Markets change and newer models become necessary. Simple clear models with explicit assumptions about small numbers of variables are therefore the best way to leverage your intuition without deluding yourself.

All models sweep dirt under the rug. A good model makes the absence of the dirt visible. In this regard, we believe that the Black-Scholes model of options valuation, now often unjustly maligned, is a model for models; it is clear and robust. Clear, because it is based on true engineering; it tells you how to manufacture an option out of stocks and bonds and what that will cost you, under ideal dirt-free circumstances that it defines. Its method of valuation is analogous to figuring out the price of a can of fruit salad from the cost of fruit, sugar, labor and transportation. The world of markets doesn't exactly match the ideal circumstances Black-Scholes requires, but the model is robust because it allows an intelligent trader to qualitatively adjust for those mismatches. You know what you are assuming when you use the model, and you know exactly what has been swept out of view.

Building financial models is challenging and worthwhile: you need to combine the qualitative and the quantitative, imagination and observation, art and science, all in the service of finding approximate patterns in the behavior of markets and securities. The greatest danger is the age-old sin of idolatry. Financial markets are alive but a model, however beautiful, is an artifice. No matter how hard you try, you will not be able to breathe life into it. To confuse the model with the world is to embrace a future disaster driven by the belief that humans obey mathematical rules.

MODELERS OF ALL MARKETS, UNITE! You have nothing to lose but your illusions.
The Modelers' Hippocratic Oath

1. *I will remember that I didn't make the world, and it doesn't satisfy my equations.*
2. *Though I will use models boldly to estimate value, I will not be overly impressed by mathematics.*
3. *I will never sacrifice reality for elegance without explaining why I have done so.*
4. *Nor will I give the people who use my model false comfort about its accuracy. Instead, I will make explicit its assumptions and oversights.*
5. *I understand that my work may have enormous effects on society and the economy, many of them beyond my comprehension."*

How is the Financial Modeler's Manifesto related to this thesis?

In the area of yield-curve modelling, not a lot of completely new work has been done since the BGM-Model (1997), but instead extensions to this framework, like for example the Displace Diffusion extension of the Libor-Market Model, see Joshi and Rebonato (2003) and the CEV extension of the Libor-Market-Model, see Andersen and Andreasen (2000) or the SABR model of Hagan, Kumar, Lesniewski and Woodward (2002).

Academically, the affine framework (including related transform methods for option pricing) was a key development form the late 90'ies up to the "outbreak" of the financial crisis 2008. Since then, we have been forced "back to fundamentals" including "what is a risk-free rates", which has led to multi-curve frameworks – in various guises, and related issues on the pricing of derivatives in this new multi-curve setting.

Even though the Market model (or versions hereof) probably is the most popular model for interest rate derivative pricing, the model is not without its limitations:

- As it is such a flexible model, one can fit any combination of Caps/Floors or Swaptions with it – the question that arises is, which one to actually use in the estimation?
- As it is a multi-factor model, in general it should be able to capture the correlation observed in the market, however as shown by Sidenius (1997) that a 10-factor model is needed in order to capture the observed correlation structure. This is true for all multi-factor models – capturing the observed correlation is not possible without a "huge" amount of factors – which complicated the use

- From a theoretical point of view, even though the model is consistent with the use of the Black (1976) model for the pricing of Caps it is not consistent with the Black (1976) model for the pricing of swaptions – which also is market standard. The reason being if the forward-rates are assumed log-normal, then an “average” of forward-rates cannot also be log-normal distributed. This has give rise to yet another version of the Libor-Market-Model, namely the Swap-Market-Model

In the wake of the financial crisis the valuation of interest rate derivatives has been even more complicated due to this new multi-yield-curve scenario, we now are in, see for example Pallaviciniy and Tarenghi (2010).

In the spirit of the Financial Manifesto we have instead decided – in the first paper (The normal class of arbitrage-free spot-rate models”) – to take a step backward in order to take a more detailed look at a type of model which belongs to the class of one-factor-models. The reasons being:

- Even though the math might be somewhat complex it is easy to understand the models limitations
- It does have all the nice ingredients, tractability and a automatic fit of the initial yield-curve, and non-negative rates

One can say that, yes we know the model does not take into account all the issues there are in the world of interest-rate derivative pricing, but it is still consistent and theoretical founded – and as the Black-Scholes model – the limitations for the user of this type of model is more easy to grasp.

The second paper (“Empirical Yield-Curve Dynamics, Scenario Símulation and Risk-Measures”) addresses methods for assessing risk for complex portfolios. Where we by complex portfolios mean portfolios that contain instruments that are not “plain-vanilla”, in our case it is Danish Mortgage Bonds.

Some citations from the crisis with respect to Risk-Management (VaR) are (see Madsen (2009)):

- ”Risk modelling did not help as much as it should have”, says Aaron Brown (former risk manager Morgan Stanley)
- ”VaR is a limited tool”, says Marc Groz (Risk Consultant)
- ”This is like an air bag that works all the time, except when you have a car accident”, says Nassim Taleb
- David Einhorn (founder of Greenlight Capital) has recently written that VaR was ”relatively useless as a risk management tool and potentially catastrophic when its use creates a false sense of security among senior managers and watchdogs”

One could easily argue that it is not so much VaR itself that are the inherent problem but also the Risk-Culture in financial companies and the way VaR is being calculated – maybe there is too much reliance on the numbers that are being calculated, and to little accept/knowledge about its limitations!

The focus of paper 2 in this connection is as follows:

- Show that it is possible to obtain much more reliable Risk-Estimates with limited additional complexity

In the last paper "Can Danish households benefit from Stochastic Programming models? – An empirical study of mortgage refinancing in Denmark" we focus on the following important practical question:

- Is it possible to improve and extend the rules of thumbs that are used for refinancing?

In this analysis we focus on introducing advanced optimization techniques to investigate if we can do better than simple rules of thumb. So, yes, we do add complexity, but we always keep in mind that the results should be intuitive. This paper is a very good example on the fact that mathematical models is an important tool in supporting financial decisions, but on the other hand one should always be able to understand from an economic point of view why (in this case) the model generate that particular refinancing decision!

Chapter 2: Summary of the papers

Paper 1: The normal class of arbitrage-free spot-rate models

When yield curve models are used in practice the traditional approach is to make sure that the parameters describing the process for the term structure satisfy certain requirements. The most important requirement in this connection is that the initial observed yield curve is exactly described by the chosen model - that is the model takes the initial yield curve as given. This has the desired characteristic that all other interest rate derivatives are priced by using a no-arbitrage argument.

The first yield curve model to satisfy this requirement was proposed by Ho and Lee (1986). Other models building on this basic idea are for example Black, Derman and Toy (1990), Black and Karasinski (1991) and Hull and White (1993a) - where these models are so called spot-rate models. The most general approach along these lines is the framework developed by Heath, Jarrow and Morton (HJM) (1991), which build an m-factor continuous time term structure model in the instantanous forward rates.

It is of course preferable to price interest rate contingent claims analytically instead of having to use more or less time-consuming numerical methods. Unfortunately, for many models this is not possible. In general, a normal distribution assumption is needed in order to have a rich analytical structure for the term structure model.

This assumption has that well known and (properly) undesired characteristic that interest rates can become negative. Recently, for that reason, a great deal of focus has been on models which contain some degree of tractability and at the same time do not allow for negative interest rates.

In the framework of HJM there is the log-normal Market Model from Brace, Gatarek and Musiela (1995) which is used by Miltersen, Sandmann and Sondermann (1995) to derive analytical expression for options on discount bonds. Another model in the HJM framework that precludes negative interest rates is the model from Flesaker and Houghton (1996) which builds a new equivalent martingale measure called the terminal measure in order to prevent negative interest rates.

In the class of spot-rate models there is the extended Cox, Ingersoll and Ross model from Jamshidian (1995) and the Quadratic interest rate model from Jamshidian (1996).

The main focus on this paper will be on the Quadratic interest rate model, we will however take the analysis much further than done in Jamshidian.

The first result of this paper is to - using the concepts of fundamental solutions - show an alternative/new method to determine the T-forward adjusted risk-measure for interest rate models which does not rely on the use of Girsanov’s theorem. With respect to obtaining the fundamental solution of linear PDEs we will here introduce a new method relying on the Fourier transformation.

Following along these lines we will focus on the one-factor Quadratic interest rate model, and in that connection we will show that the analytical structure of this model is very rich - nearly as rich as the extended Vasicek model from Hull and White (1993a). We will employ a different approach than the one advocated by Jamshidian which has that nice characteristic that we are capable of carrying the analysis of the Quadratic interest rate model much further. We will in this paper also present a new derivation of the extended Vasicek model from Hull and White (1993a), which uses the concept of Fourier transformation.

The second result of this paper is the derivation of analytical expression for the prices of discount bonds, options on discount bonds - which, it appears, can be expressed in terms of the cumulative normal distribution. In the paper it will also be shown how the model can be fitted to the initial term structure in the spirit of Hull and White (1990).

For the pricing of American style securities (like for example Bermudan options) we will specify a special discrete time version for the model and compare it to the approach from Hull and White (1994).

The last part of the paper will focus on pricing techniques in order to handle path-dependent interest-rate derivatives. In connection with Monte Carlo simulation we will only focus on one-factor spot-rate processes with time-dependent coefficients. Monte Carlo simulation of yield-curve models that by construction match the initial term structures has (as far as we know) only been applied to models in the Heath, Jarrow and Morton framework (see for example Zhenyu (1994)). Recently Andersen (1995) has designed a powerful application of importance sampling for simulating interest rates with time-homogenous parameters.
We will here design a procedure which ensures that a given Monte Carlo simulation method matches the initial yield-curve by construction - which is the third main result of the paper.

Paper 2: Empirical Yield-Curve Dynamics, Scenario Símulation and Risk-Measures

This paper has two objectives. First we will construct a general model for the variation in the term structure of interest rates, or to put it another way, we will define a general model for the shift function. Secondly, we will specify a Risk model which uses the shift function derived in the first part of the paper as its main building block.

This general model of the variation in the term structure of interest rates is assumed to belong to the linear class, and, moreover, the factors that determine the shift function are independent; the model is therefore comparable to the Ross (1976) APT (Arbitrage-Pricing-Theory) model.

The traditional approach to describe the dynamics of the term structure of interest rates is either by defining the stochastic process that drives one or more state variables, such as Cox, Ingersoll and Ross (1985), Vasicek (1977) and Longstaff and Schwartz (1991), or by postulating one or more volatility structures for determining the dynamics of the initial term structure of interest rates, such as Heath, Jarrow and Morton (1991).

However, the approach used in this paper to describe the dynamics of the term structure of interest rates is an empirical approach in order to derive the number of factors needed to describe the variation in the term structure of interest rates. Our reference period will here be 2 January 2002 - 2 January 2012.

The approach follows along the lines of Litterman and Scheinkman (1988) and in that connection we will relate the approach to the Heath, Jarrow and Morton framework. We show in that connection that the PCA method can be thought of as a tool for specifying/determining the spot rate volatility structure using a non-parametric approach.

In the second and by far the largest part of the paper we will turn our attention to Risk-Models. The reason being that Risk Measures have three very important roles within a modern financial institution:

1. It allows risky positions to be directly compared and aggregrated
2. It is a measure of the economic or equity capital required to support a given level of risk activities
3. It helps management to make the returns from a diverse risky business directly comparable on a risk adjusted basis

Our approach to the calculation of VaR/ETL is a simulation based methodology which relies on the scenario simulation framework of Jamshidian and Zhu (1997).

The general idea behind the scenario simulation procedure is to limit the number of portfolio evaluations by using the factor loadings derived in the first part of the paper and then specify particular intervals for the Monte Carlo simulated random numbers and assign appropriate probabilities to these intervals (states).

We find that the scenario simulation procedure is computationally efficient, because we with a limited number of states are capable of deriving robust approximations of the probability distribution. Compared to Monte Carlo simulation another important feature with the scenario simulation procedure is that we have more control over the tails of the distribution - which for Risk models is important.

The promising results we obtain here leads us to address the problem of Risk-Calculations for non-linear securities. More precisely we turn our attention to VaR/ETL for Danish MBBs, with as far as we are aware of, only has been considered by Jacobsen (1996). We conclude here that the methodology is both efficient and feasible to use for large portfolios of non-linear securities - because even for complex instruments like Danish MBBs the computational burden is acceptable.

Our overall conclusion is the following:

- The Jamshidian and Zhu scenario simulation methodology is best suited for the calculation of the Risk-Measure ETL - less for VaR
- We find that the scenario simulation procedure is computational efficient, because we with a limited number of states is capable of deriving robust approximations of the probability distribution
- We also find that it is very useful for non-linear securities (Danish Mortgage-Backed-Bonds MBBs), and argue that the method is feasible for large portfolios of highly complex non-linear securities - for example Danish MBBs
- Backtesting the Risk-Model setup during 2008 showed some very promising results as we were able to capture the extreme price-movements that were observed in the market

Paper 3: Can Danish households benefit from Stochastic Programming models? – An empirical study of mortgage refinancing in Denmark

This paper focuses on improving advisory standards for mortgage choice and refinancing for the Danish market. In particular the focus is on improving and extending the rules of thumbs that are used for refinancing.

Refinancing rules of thumbs have traditionally only dealt with refinancing among fixed rate mortgages, but given that most borrowers today opt for F1 loans there is a need for refinancing rules which reduce the risk of holding F1 loans by looking into strategies which allow movement from one type of loan to the other – that is strategies that takes into account the risk associated with this.

The contribution of this paper is the following:

1. Showing that the existing rules of thumb for refinancing are of little or no use
2. Introducing a model-based refinancing framework ensuring significant refinancing gains without adding risk
3. Reducing the risk of holding a F1 loan by introducing strategies which allow mixing fixed and adjustable rate loans

The paper is organized as follows: Initially we explain the existing rules of thumb for refinancing in Denmark. After that we introduce the model framework used in this paper for individual mortgage choice and refinancing. Next, we compare the performance of the model-based advising framework compared to the existing rules of thumb.

Finally we conclude suggesting that future innovations in the Danish mortgage market are likely to have special focus on risk management and advisory services. The existing mortgage products offer considerable potential for diversification and individualization in terms of risk and cost characteristics of the loans. This potential is however not taken advantage of due to lack of effective and personal consultancy tools and services. The framework suggested in this paper suggests a foundation for such tools and services.

Part II

Papers

Part II

Chapter 4

The normal class of arbitrage-free spot-rate models

THE NORMAL CLASS OF ARBITRAGE-FREE SPOT-RATE MODELS
- IMPLICATION AND IMPLEMENTATION

Claus Madsen

Abstract: In this paper we first show how to determine the T-forward adjusted risk-measure using the concept of fundamental solution to linear PDE's. After that, relying on Fourier transformation we derive bond-and bond-option prices for the Extended Vasicek model from Hull and White (1990) and the Quadratic Interest Rate model. With respect to the Quadratic Interest Rate model we succeed in carrying the analysis much further than Jamshidian (1996). A special discrete time model - which in some cases is appropriate for the Quadratic Interest Rate model - is also derived.

The last part of the paper analyse Monte Carlo techniques in connection with spot-rate models with a time-dependent drift. We also introduce a method - using the concept of forward induction - that constrain the Monte Carlo simulated spot-rate process for the matching of the initial yield-curve. For the pricing of path-dependent contingent claim, we deduce that, even though Monte Carlo is the natural method to use, it might not be the most efficient one - at least not when the spot-rate is Markovian.

Keywords: Arbitrage-free pricing, spot-rate models, Fourier transformation, forward adjusted risk-measure, closed form solutions, lattice-models, SDE discretization schemes, Monte Carlo methods, Markovian structures, positive interest rates

The normal class of arbitrage-free spot-rate models
- Implication and Implementation

1: Introduction

When yield curve models are used in practice the traditional approach is to make sure that the parameters describing the process for the term structure satisfy certain requirements. The most important requirement in this connection is that the initial observed yield curve is exactly described by the chosen model - that is the model takes the initial yield curve as given. This has the desired characteristic that all other interest rate derivatives are priced by using a no-arbitrage argument.

The first yield curve model to satisfy this requirement was proposed by Ho and Lee (1986). Other models building on this basic idea are for example Black, Derman and Toy (1990), Black and Karasinski (1991) and Hull and White (1993a) - where these models are so called spot-rate models. The most general approach along these lines is the framework developed by Heath, Jarrow and Morton (1991), which build an m-factor continuous time term structure model in the instantanous forward rates.

It is of course preferable to price interest rate contingent claims analytically instead of having to use time-consuming numerical methods. Unfortunately, for many models this is not possible. In general, a normal distribution assumption is needed in order to have a rich analytical structure for the term structure model.

This assumption has that well known and (properly) undesired characteristic that interest rates can become negative. Recently, for that reason, a great deal of focus has been on models which contain some degree of tractability and at the same time do not allow for negative interest rates.

In the framework of HJM there is the log-normal Market Model from Brace, Gatarek and Musiela (1995) which is used by Miltersen, Sandmann and Sondermann (1995) to derive analytical expression for options on discount bonds. Another model in the HJM framework that precludes negative interest rates is the model from Flesaker and Houghton (1996) which builds a new equivalent martingale measure called the terminal measure in order to prevent negative interest rates.

In the class of spot-rate models there is the extended Cox, Ingersoll and Ross model from Jamshidian (1995) and the Quadratic interest rate model from Jamshidian (1996).

The first result of this paper is to - using the concepts of fundamental solutions - show an alternative/new method to determine the T-forward adjusted risk-measure for interest rate models which does not rely on the use of Girsanov's theorem. With respect to obtaining the fundamental solution of linear PDEs we will here introduce a new method relying on the Fourier transformation.

Following along these lines we will focus on the one-factor Quadratic interest rate model, and in that connection we will show that the analytical structure of this model is very rich - nearly as rich as the extended Vasicek model from Hull and White (1993a). We will employ a different approach than the one advocated by Jamshidian which has that nice characteristic that we are capable of carrying the analysis of the Quadratic interest rate model much further[1]. We will in this paper also present a new derivation of the extended Vasicek model from Hull and White (1993a), which uses the concept of Fourier transformation.

The second result of this paper is the derivation of analytical expression for the prices of discount bonds, options on discount bonds - which, it appears, can be expressed in terms of the cumulative normal distribution. In the paper it will also be shown how the model can be fitted to the initial term structure in the spirit of Hull and White (1990).

For the pricing of American style securities (like for example Bermudan options) we will specify a special discrete time version for the model and compare it to the approach from Hull and White (1994).

The last part of the paper will focus on pricing techniques in order to handle path-dependent interest-rate derivatives. In connection with Monte Carlo simulation we will only focus on one-factor spot-rate processes with time-dependent coefficients. Monte Carlo simulation of yield-curve models that by construction match the initial term structures has (as far as we know) only been applied to models in the Heath, Jarrow and Morton framework (see for example Zhenyu (1994)). Recently, Andersen (1995) has designed a powerful application of importance sampling for simulating interest rates with time-homogenous parameters.

We will here design a procedure which ensures that a given Monte Carlo simulation method matches the initial yield-curve by construction - which is the third main result of the paper.

The paper is organized as follows: In section 2, we will define a general class of one-factor interest rate models and derive the partial differential equation that all interest rate contingent claims have to satisfy. In this section, we will also show how the T-forward adjusted risk-measure and fundamental solutions to PDEs are connected. In section 3, a method to obtain fundamental solutions from the PDE using the concept of Fourier transformation is introduced.

Section 4 will employ this technique for the extended Vasicek model and in that connection show how the price of discount bonds and options on discount bonds can be obtained. In section 5 we will turn our focus on the Quadratic interest rate model and show how to derive the price of discount bonds and options on discount bonds.

---

[1] This has independtly been done by Pelsser (1995).

We will in the spirit of Hull and White (1990) and Jamshidian (1996) show how it is possible to fit the model to the initial term structure. The next two sections will focus on pricing techniques when closed form solutions are not avaiable. First we will focus on a discrete time model for the Quadratic interest rate model and second we will analyse general Monte Carlo methods for spot-rate processes with time-dependent parameters.

2: Generalized one-factor models

In this paper, I consider a continuous trading economy with zero-coupon bonds and a money market account with a trading interval [0,τ], for a fixed τ > 0. In addition, it is assumed that money does not exist, i.e. that the agents in the economy are forced at all times to invest all their funds in assets. As usual, the uncertainty in the economy is characterized by the probability space (Ω,F,P), where Ω is the entire state space, P is a probability measure and F is the event space. At the same time, it is assumed that a one-dimensional Wiener process exists: $W = [W(t);0 < t \leq T < \tau]$, with a drift equal to zero (0) and a variance equal to one (1).

Let us now consider, like Hull and White (1993a), the following rather general single-factor yield curve model:

$$\begin{gathered} dy = \mu(\varphi(t),y,t)dt + \sigma(y,t)dW \\ \textit{where} \\ r = F(t,y) \end{gathered} \qquad \textbf{(1)}$$

Where μ(•) represents the drift coefficient, which may be a function of time and the state variable, and where φ(t) is a time-dependent function, σ(•) represents the diffusion coefficient, which may be a function of time and the state variable, and dW is a Wiener process with the following properties: $(dW)^2 = dt$, $dtdW = 0$ and $(dt)^2 = 0$. Furthermore, the spot rate is determined from the underlying process through the function F(t,y).

An important class of models arise under the assumption that the diffusion coefficient is independent of the state variable, i.e. σ(y,t) = σ(t) and that the drift coefficient is defined as $\mu(\varphi(t),y,t) = \varphi(t) - \kappa(t)y$ - where κ(t) is the mean-reversion parameter.

For this particular choice, the process y has a normal distribution. Therefore, I will denote this class of models as normal models. By making appropriate choices for F(t,y) it is possible to show that several existing models belong to this category of normal models. For F(t,y) = y we obtain the extended Vasicek model from Hull and White (1990). By assuming $F(t,y) = e^y$ we get the Black and Karasinski (1991) model and for $F(t,y) = y^2$, this will lead to the one-factor Quadratic model from Jamshidian (1996). Lastly, (for F(t,y) = y) we get the Ho and Lee (1986) model for κ(t) = 0.

Using the arbitrage-free argument, it can be shown that the price of g(y,t,T) of an interest rate derivative security at time t which has a payoff at time T satisfies the partial differential equation:

$$g + g_y[\mu(\varphi(t),y,t) - \lambda\sigma(y,t)] + \frac{1}{2}\sigma(y,t)^2 g_{yy} - F(t,y)g = 0 \qquad \textbf{(2)}$$

Where λ denotes the market price of risk.

The price g(y,t,T) of an interest rate derivative that has a payoff H(y(T)) at maturity T can be calculated by solving the partial differential equation subject to the boundary condition g(y,T,T) = H(y(T)) at time T - that is at expiry/maturity the value of the derivative security is equal to the payoff at the time of expiry.

The Feynman-Kac solution to formula no. 2 (if it exists) can be written as[2]:

$$\begin{aligned} g(y,t,T) &= E^Q\left[\left(\exp - \int_t^T F(s,y(s))ds\right) H(y(T))|F_t\right] \\ &= E^Q\left[\left(\exp - \int_t^T r(s)ds\right) H(y(T))|F_t\right] \end{aligned} \qquad \textbf{(3)}$$

Where the expectation is taken, conditional on the information avaiable at time t, with respect to the risk-adjusted process:

$$\begin{gathered} dy = \mu(\varphi(t),y,t) - a(t)y]dt + \sigma(y,t)d\tilde{W} \\ where \\ d\tilde{W} = dW - \lambda\sigma(y,t) \end{gathered} \qquad \textbf{(4)}$$

Where $\tilde{W}_i$ is a Wiener process on (Ω,F,Q), for dQ = ρdP and ρ is the Radon-Nikodym derivative.

Using the Feynman-Kac formula the price of any interest rate derivative is given by the risk-adjusted expectation of the discounted payoff. However, as the discounting term and the payoff term are two correlated stochastic variables, this expectation is, in general, difficult to evaluate. As shown by Karoui, Myneni and Viswanathan (1993) it is more efficient to use the change of numerator techniques which in a sense moves the discounting outside the expectation and then the expectation is taken under a different probability measure and only with respect to the payoff at maturity. In connection which formula 3 this means that the price g(y,t,T) can be written as:

(5)

---

[2] In general the term " Feynman-Kac" is considered a misnomer as it originally refers to the probabilistic solution to a narrower class of parabolic equations than the Cauchy problem, see Duffie (1996) Appendix E. Typically formula 3 would be called the probabilistic solution to the PDE.

Where Q_T is the probability measure that is associated with using the T-period zero-coupon bond as numerator.

The result from formula 5 is obtained by using Girsanov's theorem. In the rest of this chapter, I will show an alternative derivation of this result, using the concept of fundamental solutions of PDEs.

Let us now suppose that a function d(y,t,T,z) exists that satisfies a linear PDE in y and t (that is a PDE like the one in formula 2), with boundary condition:

$$d(y,T,T,z) = \delta(y - z) \tag{6}$$

for all z and T. Formula 6 means that at time T the function d(y,t,T,z) will collapse into a Dirac delta-function[3] centered at point z. If we now consider a boundary condition described by the function H(y(T)) at time T then we have that:

$$g(y,t,T) = \int_{-\infty}^{\infty} H(z)d(y,t,T,z)dz \tag{7}$$

satisfies the partial differential equation, due to the linearity of the differentiation operator.

Furthermore, the above equation also satisfies the boundary condition at t = T. This is true since at t = T we have:

$$g(y,T,T) = \int_{-\infty}^{\infty} H(z)d(y,T,T,z)dz = \int_{-\infty}^{\infty} H(z)\delta(y - z)dz = H(y(T)) \tag{8}$$

where the last equality follows from the definition of the Dirac delta-function.

This means that for any given boundary condition, we can use the function d(y,t,T,z) to construct a solution to the PDE. For this reason, the function d(y,t,T,z) is called the fundamental solution to the partial differential equation.

It is actually possible to give an economic interpretation of formula 7. The Dirac delta-function can be thought of as the continuous time version of the payoff of an Arrow-Debreu security. As δ(y - z) ≠ 0 if only y = z and $\int_{-\infty}^{\infty} \delta(y - z)dz = 1$, we could say that the Dirac delta-function gives a payoff worth 1 in the state y = z and zero (0) elsewhere. This property means that d(y,t,T,z) can be viewed as the price at time t in state y of an Arrow-Debreu security that has a payoff of 1 at time T in state y = z.

---

[3] See Wilmott, Dewynne and Howison (1993) chapter 5.

If we now consider a discount bond, it is well known that this is a security that has a payoff of 1 in all states at maturity T. From this follows that the price at time t of a zero-coupon bond that matures at time T (P(y,t,T))[4] is given by $P(y,t,T) = \int_{-\infty}^{\infty} d(y,t,T,z)dz$.

If the economy is arbitrage-free, the prices of Arrow-Debreu securities cannot become negative and prices of discount bonds are finite. These observations lead to another interpretation for formula 7. More precisely it is possible to write the following relationship:

$$p(y,t,T,z) = \frac{d(y,t,T,z)}{P(y,t,T)} \tag{9}$$

The function p(y,t,T,z) is positive and integrates by construction to 1 with respect to z. Any function that integrates to 1 and is non-negative can be interpretated as a probability density function. By this, it follows that formula 7 can be rewritten as:

$$g(y,t,T) = P(y,t,T)\int_{-\infty}^{\infty} H(z)p(y,t,T,z)dz = P(y,t,T)E^{Q_T}[H(z(T))|F_t] \tag{10}$$

Where the expectation is taken, conditional on the information avaiable at time t, with respect to the density p(y,t,T,z), the density p(y,t,T,z) will in general differ from the transition density of the risk-adjusted process of y from formula 4.

If we now combine this result with the relation from formula 3, where z represents y(T), we get the following result:

$$g(y,t,T) = E^Q\left[\left(\exp - \int_t^T r(s)ds\right) H(z)|F_t\right] = P(y,t,T)E^{Q_T}[H(z)|F_t]F_t] \tag{11}$$

That is, we have now expressed the Q-expectation as the discounted value of the Q_T-expectatation of the time T payoff. Under the expectation E^Q the prices $\frac{d(y,t,T,z)}{P(y,t,T)}$ of interest rate derivatives with maturity T denominated in T-period discount bonds are martingales. This also holds true under E^{Q_T} , since for each $t < T_1 < T$ we have:

---

[4] I will interchangely use the notation P(t,T), P(y,t,T) and P(q,t,T) for the price at time t for a disount bond that matures at time T - whenever applicable.

$$E^{Q_T}\left[\frac{g(y(T_1),T_1,T)}{P(y(T_1),T_1,T)}|F_t\right] = E^{Q_T}\left[\frac{P(y,T_1,T)E^{Q_T}[H(y(T))|F_{T_1}]}{P(y,T_1,T)}|F_t\right] = E^{Q_T}[E^{Q_T}[H(y(T))|F_{T_1}]|F_t] = E^{Q_T}[H(y(T))|F_t] = \frac{g(y(t),t,T)}{P(y(t),t,T)} \tag{12}$$

Where the fourth equality follows from the law of iterated expectations, and the second and fifth are obtained through the use of formula 10.

3: Deriving Fundamental Solutions to linear PDEs

When we want to find the fundamental solution to a linear PDE, we are actually seeking functions d(y,t,T,z) that satisfy the partial differential equation and at expiry/maturity collapse into the Dirac delta-function for all z. Instead of trying to obtain the fundamental solution directly, I will solve the partial differential equation for the Fourier transformation of the fundamental solution and find the fundamental solution from here.

Let us now consider the function D(y,t,T,ς) defined as:

$$D(y,t,T,\varsigma) = \int_{-\infty}^{\infty} e^{-i\varsigma z} d(y,t,T,z)dz \tag{13}$$

Where this formula represents the Fourier transformation in the variable z and d(y,t,T,z), where i is the imaginary number, for which $i^2 = -1$. The function D(y,t,T,ς) satisfies the same PDE as d(y,t,T,z) but the boundary condition is here given by:

$$D(y,T,T,\varsigma) = \int_{-\infty}^{\infty} e^{-i\varsigma z}\delta(y - z)dz = e^{-i\varsigma y} \tag{14}$$

Where the last equality follows from the definition of the Dirac delta-function. It follows from here that the boundary condition for the function D(y,t,T,ς) is of a very simple form.

We can obtain the fundamental solutions d(y,t,T,z) by inverting the Fourier transformation D(y,t,T,ς). This inversion is usually very simple if we use the following property. Using the relation from formula 9, we can write the fundamental solution as a product of the price of the T-period discound bond and a probability density function, hence we can rewrite formula 14 as:

$$D(y,tT,T,\varsigma) = \int_{-\infty}^{\infty} e^{-i\varsigma z} P(y,t,T)p(y,t,T,z)dz = P(y,t,T)\tilde{p}(y,t,T,\varsigma) \tag{15}$$

Where $\tilde{p}(y,t,T,\varsigma)$ is the Fourier transformation of the density function p(y,t,T,z). Apart from a minus sign in the exponent (and, possibly, a factor $\sqrt{\frac{1}{2\pi}}$), characteristic functions coincide with Fourier transforms of a probability density function in the continuous case and with Fourier series in the discrete case. Furthermore, it is well known that it is possible to obtain the moments of the underlying distribution directly from the charateristic function[5].

If we assume that $\varsigma = 0$, then we have that the boundary condition for D(y,t,T,ς) becomes identical to the boundary condition for a discount bond, hence:

$$D(y,t,T,0) = P(y,t,T) \Rightarrow \tilde{p}(y,t,T,\varsigma) = \frac{D(y,t,T,\varsigma)}{D(y,t,T,0)} \qquad \textbf{(16)}$$

Where the last equality follows directly from formula 14.

From formula 16 it can be seen that from D(y,t,T,ς) we can obtain the bond price P(y,t,T) and the Fourier transformation of the density function p(y,t,T,z) - that is $\tilde{p}(y,t,T,\varsigma)$ -and then we can derive the density p(y,t,T,z) by inversion of the characteristic function[6]. Once we have defined the bond price P(y,t,T) and the transitions densities p(y,t,T,z), it is possible to price interest rate derivative securities by using formula 10.

Before turning my attention to the Quadratic interest rate model, I will demonstrate this technique for a more simple model, namely the extended Vasicek model from Hull and White.

4: Deriving the Fundamental Solution for the Extended Vasicek Model

As was mentioned in section 2, we get the extended Vaseicek model from Hull and White (1990) by making the following specification of the drift-parameter, the diffusion-parameter and the function F(t,y):

$$\begin{aligned} F(t,y) &= y \\ \mu(\varphi(t),y,t) &= \varphi(t) - \kappa(t)y \\ \sigma(y,t) &= \sigma(t) \end{aligned} \qquad \textbf{(17)}$$

which results in the following risk-adjusted spot rate process:

---

[5] See Gut (1995).

[6] If it is not possibly to invert the characteristic function then we can at least get useful information about the distribution, such as the moments or approximating distributions.

$$dy = [\varphi(t) - \kappa(t)y]dt + \sigma(t)d\tilde{W} \quad \textit{where} \quad d\tilde{W} = dW - \lambda\sigma(t) \tag{18}$$

Let me employ the following change of variable $r = y + \psi(t)$ where $\varphi(t) = \psi'(t) + \kappa(t)\psi(t)$ and r(0) = ψ(0) as y(0) = 0. Using Ito-lemma we can deduce that

$$\psi(t) = e^{-k(t)}\int_0^t \varphi(s)e^{k(s)}ds \ \textit{for} \ k(t) = \int_0^t \kappa(s)ds .$$

Under this change of variable the partial differential equation from formula 2 becomes of the following form:

$$g_t - g_y[\kappa(t)y + \lambda\sigma(t)] + \frac{1}{2}\sigma(t)^2 g_{yy} - g[y + \psi(t)] = 0 \tag{19}$$

In order to obtain the price of a discount bond in this model, we have to solve the PDE in formula 19 with respect to the boundary condition $D(y,T,T,\varsigma) = e^{i\varsigma y}$.

We guess that the solution is of the affine exponential type[7]:

$$D(y,t,T,\varsigma) = e^{B(t,T,\varsigma) + A(t,T,\varsigma)y} \tag{20}$$

By substituting this functional form into the PDE we get[8]:

$$B_t + A_t y - A\lambda\sigma(t)] - A\kappa(t)y + \frac{1}{2}\sigma^2(t)A^2 - y - \psi(t) = 0 \tag{21}$$

The partial differential equation can be solved by simultaneously solving the following two ordinary differential equations with respect to each other subject to the boundary conditions A(T,T,ς) = iς and B(T,T,ς) = 0:

$$\begin{gathered} B_t \mp Ax(t) + \frac{1}{2}\sigma^2(t)A^2 - \psi(t) = 0 \\ A_t - A\kappa(t) - 1 = 0 \\ \textit{for} \\ x(t) = \kappa(t) - \lambda\sigma(t) \end{gathered} \tag{22}$$

---

[7] "Guess" is probably the wrong term to use in this context, as we know from Hull and White (1990) that this actually is the case.

[8] I will in this paper interchangely use the notation A(t,T,ς) and A, B(t,T,ς) and B and C(t,T,ς) and C - whenever applicable.

The solution for A(t,T,ς) and B(t,T,ς) can be shown to be given by[9]:

$$
\begin{gathered}
A(t,T,\varsigma) = -A(t,T) + i\varsigma D(t,T) \\
B(t,T,\varsigma) = B(t,T) - \int_t^T \left[i\varsigma D(s,T)[x(s) + \sigma^2(s)A(s,T)] + \frac{1}{2}\varsigma^2 D(s,T)^2\sigma^2(s) \right] ds \\
\textit{where} \\
D(t,T) = e^{-\int_t^T a(s)ds} \\
A(t,T) = \int_t^T D(t,s)ds \\
B(t,T) = \int_t^T \left[\frac{1}{2}\sigma^2(s)A(s,T)^2 \pm x(s)A(s,T) - \psi(s) \right] ds
\end{gathered}
\tag{23}
$$

Substituting this into formula 22 yields:

$$
\begin{gathered}
D(y,t,T,\varsigma) = e^{B(t,T) - A(t,T)y + i\varsigma M(t,T) - \frac{1}{2}\varsigma^2 V(t,T)} \\
\textit{where} \\
M(t,T) = D(t,T)y - \int_t^T D(s,T)[x(s) \pm \sigma^2(s)A(s,T)]ds \\
V(t,T) = \int_t^T D(s,T)^2\sigma^2(s)ds
\end{gathered}
\tag{24}
$$

From the boundary condition for D(y,t,T,ς) we get the boundary condition for discount bonds for ς = 0 and then we immediately obtain the price of a discount bond from equation 24, namely:

$$
P(t,T) = e^{B(t,T) - A(t,T)y} \tag{25}
$$

and furthermore we see that the characteristic function of the density p(y,t,T,z) is given by:

$$
\tilde{p}(y,t,T,\varsigma) = e^{i\varsigma M(t,T) - \frac{1}{2}\varsigma^2 V(t,T)} \tag{26}
$$

It can now be observed that formula 26 actually is the characteristic function for the normal distribution with mean M(t,T) and variance V(t,T). We therefore have that the fundamental solution d(y,t,T,z) for the extended Vasicek model is given by:

---

[9] See Appendix A where the actual derivation is done.

$$d(y,t,T,z) = P(y,t,T)\frac{1}{\sqrt{2\pi V(t,T)}}e^{-\frac{1}{2}\frac{[z - M(t,T)]^2}{V(t,T)}} \tag{27}$$

Now that we have the fundamental solution, it is possible to price interest rate derivatives with the joint use of formula 10, formula 16 and the result from formula 27.

To illustrate this, let me show how to derive the price at time t of a call-option with expiry date T^F, for $t < T^F < T$, written on a discount bond which matures at time T. By assuming that $y(T^F) = z$ we can write the payoff as:

$$H(z) = [e^{B(T^F,T) - A(T^F,T)z} - K]^+ \tag{28}$$

Where K is the strike-price.

The price of the option can now be written as:

$$C(t,T^F) = P(t,T^F)\int_{-\infty}^{\infty}\left[[e^{B(T^F,T) - A(T^F,T)z} - K]^+\frac{1}{\sqrt{2\pi V(t,T^F)}}e^{-\frac{1}{2}\frac{[z - M(t,T^F)]^2}{V(t,T^F)}}\right]dz \tag{29}$$

From formula 28, it can be deduced that the payoff H(z) is positive for all z that satisfies

$$z < \frac{B(T^F,T) - \ln K}{A(T^F,T)}.$$

If the integration in formula 29 is done over the region for positive values of z, we can write the price of the call-option in terms of the cumulative normal distribution, hence:

$$\begin{aligned} C(t,T^F) = {} & P(t,T^F)e^{B(T^F,T) - A(T^F,T)M(t,T^F) + \frac{1}{2}A(T^F,T)^2V(t,T^F)}N\left(\frac{\frac{B(T^F,T) - \ln K}{A(T^F,T)} - M(t,T^F) + A(T^F,T)V(t,T^F)}{\sqrt{V(t,T^F)}}\right) \\ & - P(t,T^F)KN\left(\frac{\frac{B(T^T,T) - \ln K}{A(T^F,T)} - M(t,T^F)}{\sqrt{V(t,T^F)}}\right) \end{aligned} \tag{30}$$

As $P(T^F,T) = e^{B(T^F,T) - A(T^F,T)M(t,T^F) + \frac{1}{2}A(T^F,T)^2V(t,T^F)}$ we get[10]

---

[10] Later in the paper, this will become obvious, see section 6 and Appendex E.

$$\begin{gathered} C(t,T^F) = P(t,T)N(d_1) - P(t,T^F)KNt(d_1 - A(T^F,T)\sqrt{V(t,T^F)}) \\ where \\ d_1 = \frac{\ln\left(\dfrac{P(t,T)}{P(t,T^F)K)}\right)}{A(T^F,T)\sqrt{V(t,T^F)}} + \frac{1}{2}A(T^F,T)\sqrt{V(t,T^F)} \end{gathered} \tag{31}$$

where the expression in formula 31 also can be found in Hull and White (1993c). The price of a put-option is now easily derived by using either the put-call parity or the symmetric properties of the normal distribution.

Now I will turn my attention to the Quadratic interest rate model, and in the next section show how the price of a discount bond can be derived using the principle of fundamental solutions. Furthermore, I will show how to derive the price of a call-option on a discount bond in a easier expression than the one given in Jamshidian (1996) as it turns out that it can be expressed in terms of the cumulative normal distribution. I will then briefly discuss the pricing of caps, swaptions and options on coupon bonds in the Quadratic interest rate model compared to the extended Vasicek model.

After that I will discuss a discrete time implementation of the Quadratic interest rate model in comparison with the Hull and White (1994) trinomial approach. In section 8 I will focus on Monte Carlo simulation of spot-rate processes with time-dependent parameters.

5: The Quadratic Interest Rate Model

As mentioned in section 2, we get the Quadratic interest rate model by assuming that F(t,y) = y^2. Furthermore, I will only allow for one time-dependent parameter in the SDE for y, that is:

$$\begin{gathered} dy = [\varphi(t) - \kappa y]dt + \sigma d\tilde{W} \\ where \\ d\tilde{W} = dW - \lambda\sigma \end{gathered} \tag{32}$$

This property to allow only for one time-dependent parameter in the stochastic differential equation and use that for matching the initial term structure is also the approach advocated by Hull and White (1994)[11].

Using the arbitrage-free argument, it can be shown that the price of g(y,t,T) of an interest rate derivative security at time t which has a payoff at time T satisfies the partial differential equation:

---

[11] In the case of the Quadratic Interest rate model we actually need a deterministic volatility for the underlying state-variable - this issue is addressed in Appendix F.

$$g + g_y[\varphi(t) - \kappa y - \lambda\sigma] + \frac{1}{2}\sigma^2 g_{yy} - F(t,y)g = 0 \tag{33}$$

We will employ the following change of variable[12]:

$$\begin{gathered} q = y - \psi(t) \\ \psi(t) = e^{-\kappa t}\left(\sqrt{r} + \int_0^t e^{\kappa s}\varphi(s)ds\right) \end{gathered} \tag{34}$$

Under the change of variable, we have that the spot rate is governed by the following SDE:

$$\begin{gathered} dq = -\kappa q dt + \sigma d\tilde{W} \\ where \\ r = [q + \psi(t)]^2 \\ d\tilde{W} = dW - \lambda\sigma \end{gathered} \tag{35}$$

Which means that the price of g(y,t,T) of an interest rate derivative security at time t has to satisfy the following PDE:

$$g_t - g_q[\kappa q + \lambda\sigma] + \frac{1}{2}\sigma^2 g_{qq} - g[q + \psi(t)]^2 = 0 \tag{36}$$

From the last sections, we know that in order to get a solution for an interest rate derivative security, we "just" have to find the expression for D(q,t,T,ς) as the price of discount bonds and the density p(q,t,T,z) can be found by inverting the characteristic function obtained through a Fourier transformation.

Using the technique employed in the last section we postulate that the functional form for D(q,t,T,ς) is given by:

$$D(q,t,T,\varsigma) = e^{A(t,T,\varsigma) - B(t,T,\varsigma)q - C(t,T,\varsigma)q^2} \tag{37}$$

which satisfies formula 36 subject to the boundary condition $D(q,T,T,\varsigma) = e^{i\varsigma q}$.

Plucking formula 37 into formula 36 we have that the partial differential equation can be solved by simultaneously solving the following three ordinary differential equations with respect to each other subject to the boundary conditions A(T,T,ς) = C(T,T,ς) = 0 and B(T,T,ς)

---

[12] As will be apparent later in this paper, this change of variable will make a discrete time version of the model easier to implement.

= -iς:

$$\begin{aligned} A_t + \frac{1}{2}\sigma^2 B^2 - \sigma^2 C - \psi(t)^2 &= 0 \\ B_t - \kappa B - 2\sigma^2 BC + 2\psi(t) &= 0 \\ C_t - 2\kappa C - 2\sigma^2 C^2 + 1 &= 0 \end{aligned} \qquad \textbf{(38)}$$

The solution for A(t,T,ς), B(t,T,ς) and C(t,T,ς) can be shown to be given by[13]:

$$\begin{aligned} C(t,T,\varsigma) &= C(t,T) \\ B(t,T,\varsigma) &= B(t,T) - i\varsigma D(t,T) \\ A(t,T,\varsigma) &= A(t,T) - \frac{1}{2}\sigma^2 \int_t^T [2i\varsigma D(s,T)B(s,T) + \varsigma^2 D(s,T)^2] ds \end{aligned} \qquad \textbf{(39)}$$

Where:

$$\begin{aligned} A(t,T) &= \int_t^T \left[\frac{1}{2}\sigma^2 B(s,T)^2 - \sigma^2 C(s,T) - \psi(s)^2\right] ds \\ B(t,T) &= 2\int_t^T \frac{e^{\gamma s} h(s,T)}{e^{\gamma t} h(t,T)} \psi(s) ds \\ C(t,T) &= \frac{e^{2\gamma(T-t)} - 1}{h(t,T)} \\ D(t,T) &= \frac{2\gamma e^{\gamma(T-t)}}{h(t,T)} \\ &\textit{where} \\ \gamma &= \sqrt{\kappa^2 + 2\sigma^2} \\ h(t,T) &= (\kappa + \gamma) e^{2\gamma(T-t)} + (\gamma - \kappa) \end{aligned} \qquad \textbf{(40)}$$

Substituting this into formula 37 yields:

---

[13] See Appendix B for the proof.

$$D(q,t,T,\varsigma) = e^{A(t,T) - B(t,T)q - C(t,T)q^2 + i\varsigma M(t,T) - \frac{1}{2}\varsigma^2 V(t,T)}$$

$$\textit{where}$$

$$M(t,T) = D(t,T)q - \sigma^2\int_t^T D(s,T)B(s,T)ds \qquad \textbf{(41)}$$

$$V(t,T) = \sigma^2\int_t^T D(s,T)^2 ds = \sigma^2 C(t,T)$$

From the boundary condition of D(q,t,T,ς), we can derive the price for a zero-coupon bond, and we obtain:

$$P(t,T) = e^{A(t,T) - B(t,T)q - C(t,T)q^2} \qquad \textbf{(42)}$$

As in the extended Vasicek case, we see from equation 41 that the terms containing ς can be recognized as the characteristic function for the normal distribution. This leads us to conclude that under the T-forward adjusted probability measure the probability density function p(q,t,T,z) is equal to a normal probability function, which means that the fundamental solution d(q,t,T,ς) for the Quadratic interest rate model can be expressed as:

$$d(q,t,T,z) = P(q,t,T)\frac{1}{\sqrt{2\pi V(t,T)}}e^{-\frac{1}{2}\frac{[z - M(t,T)]^2}{V(t,T)}} \qquad \textbf{(43)}$$

That is the density for the Quadratic interest rate model is identical to the density for the extended Vasicek model, except for the definition of M(t,T) and V(t,T) - as they are both of the normal class.

As in section 4, we are now in a position were we can price interest rate derivatives by a joint use of the density from formula 43 and formulas 10 and 16.

To illustrate this, let me again focus on the pricing at time t of a call-option with expiry date T^F, for $t < T^F < T$, written on a discount bond which matures at time T. By assuming that $y(T^F) = z$ we can write the payoff as:

$$H(z) = [e^{A(T^F,T) - B(T^F,T)z - C(T^F,T)z^2} - K]^+ \qquad \textbf{(44)}$$

Where K is the strike-price.

The price of the option can now be written as:

$$C(t,T^F) = P(t,T^F)\int_{-\infty}^{\infty}\left[[e^{A(T^F,T) - B(T^F,T)z - C(T^F,T)z^2} - K]^+ \frac{1}{\sqrt{2\pi V(t,T^F)}} e^{-\frac{1}{2}\frac{[z - M(t,T^F)]^2}{V(t,T^F)}}\right]dz \qquad \textbf{(45)}$$

If we integrate formula 45 for all z for which the payoff H(z) is positive, we will be able to express the price of the call-option in terms of the cumulative normal distribution.

It follows directly from formula 44 that z is defined as the root in a second order equation, more precisely we find that z is positive for values of z lying in the following interval:

(46)

From this relation it it clearly seen that for Dis ≤ 0, the payoff H(z) will never be positive - that is the value of the option has to be equal to zero (0). If instead Dis > 0, we can integrate formula 45 over the region Lower < z < Upper. If this is done, we obtain the following analytical expression for the price of a call-option[14]:

$$\begin{gathered} C(t,T^F) = P(t,T)\left[N\left(\frac{UpperF - E}{\sqrt{V(t,T^F)F}}\right) - N\left(\frac{LowerF - E}{\sqrt{V(t,T^F)F}}\right)\right] \\ - P(t,T^F)K\left[N\left(\frac{Upper - M(t,T^F)}{\sqrt{V(t,T^F)}}\right) - N\left(\frac{Lower - M(t,T^F)}{\sqrt{V(t,T^F)}}\right)\right] \\ where \\ E = M(t,T^F) - B(T^F,T)V(t,T^F) \\ F = 1 + 2C(T^F,T)V(t,T^F) \end{gathered} \qquad \textbf{(47)}$$

The same procedure can be used to derive the formula for a put-option, see Appendix C.

Quadratic interest rate models have also been analysed in for example Leblon and Moraux (2009) and Leippold and Wu (2001). However, they only consider the quadratic interest rate model for the case of constant parameters - that is their representations is not defined so it by construction match the initial yield-curve as is the case here.

Earlier examples of quadratic models - for the case of constant parameters - is the double square-root model of Longstaff (1989), the extended (and corrected) version from Beaglehole and Tenney (1991) and Beaglehole and Tenney (1992). One can also mention the SAINTS model of Constantinides (1992), where the pricing kernel is being defined as exogenously specified as a time-seperable quadratic function of the Markov process.

The paper (that to our knowledge) takes the analyses of the quadratic interest rate model furtherst is Leippold and Wu (2001). They for example manage to obtain closed-form solution for European options on bonds, which however require that one has to perform a numerical inversion of a Fourier transform. As comparison Jamshidian's (1996) expression for the price of an option on a discount bond was expressed in terms of the non-central chi-square distribution. See also Andersen and Piterbarg (2010), section 12.3.

---

[14] See Appendix C where the derivation has been made.

As it is possible to get a (nearly) closed form solution for both the price of a discount bond and the price of an option on a discount bond in the Quadratic interest rate model, it follows directly that a whole range of instruments can also be priced analytically. As examples, I can mention coupon bonds, Calls/Floors and Swaps. This is of course also possible in the extended Vasicek model but in the extended Vasicek model we are also capable of deriving closed form solution for the value of swaptions and options on coupon bonds, see Jamshidian (1989). In order to value swaptions or options on coupon bonds in the Quadratic interest model it does not seem possible to derive closed form solutions, instead we could use numerical integration in order to solve the following equation:

$$\begin{gathered} C(t,T^F) = P(t,T^F)\int_{-\infty}^{\infty}[P^k(T^F) - K]^+ p(y,t,T,z)dz \\ where \\ P^k(t) = \sum_{j=1}^{n} F_j P(t,T_j) \end{gathered} \tag{48}$$

where formula 48 gives the price at time t of a call-option that expires at time T^F, written on a coupon bond that matures at time T_n, for $t < T^F < T_j$ and $j \varepsilon \{1,2,....,n\}$.

Another possibility is to discretize the process in the manner of Hull and White (1994) and then price interest rate contingent claims by backwardation. This procedure is quite flexible and allows us to price claims with American style features[15] (backward path-dependency), so for that reason I will show how to employ this technique for the Quadratic interest rate model - this is done in section 7.

6: Arbitrage-free pricing in the Quadratic interest rate model

The reason for incorporating a time-dependent parameter in the drifts-specification as was done for the extended Vasicek model in formula 18 and for the Quadratic interest rate model in formula 32, was to be able to use the initial term structures as input and then price all other instruments using a no-arbitrage argument. The purpose of this section is to show how $\psi(t)$ is to be chosen so that the model by construction matches the initial term structure.

I will here only concentrate on the Quadratic interest model as the extended Vasicek model with respect to this reformulation has been analysed in Kijima and Nagayama (1995).

---

[15] For example Bermudan options.

The Heath, Jarrow and Morton framework is my starting point here as this framework by construction matches the initial term structure, and then I will rewrite the Quadratic interest rate model so it is consistent with the procedure of HJM.

Let me as HJM start by defining the forward rates:

$$r^{F}(t,T) = -\frac{\partial \ln P(t,T)}{\partial T} = -\frac{1}{P(t,T)}\frac{\partial P(t,T)}{\partial T} \quad for\ all\ t < T \tag{49}$$

We have that P(t,T) can be expressed as:

$$P(t,T) = E^{Q}\left[\left(\exp - \int_{t}^{T} r(s)ds\right)|F_{t}\right] \tag{50}$$

Plucking this expression into formula 49 yields:

$$\begin{aligned} r^{FP}(t,T) &= \frac{\partial E^{Q}\left[\left(\exp - \int_{t}^{T} r(s)ds\right)|F_{t}\right]}{\partial T}\frac{1}{P(t,T)} = E^{Q}\left[\frac{\partial \exp - \int_{t}^{T} r(s)ds}{\partial T}|F_{t}\right]\frac{1}{P(t,T)} \\ &= E^{Q}\left[\left(\exp - \int_{t}^{T} r(s)ds\right) r(T)|F_{t}\right]\frac{1}{P(t,T)} = P(t,T)E^{Q_{T}}[r(T)|F_{t}] \end{aligned} \tag{51}$$

where the last part in formula 51 follows directly from equation 5. As a consequence, it is obvious that under the Q_T- probability measure the expected instantaneous forward rate at time T is given by the spot rate at time T.

We have that $r(T) = [q(T) + \psi(T)]^2$ and furthermore, we know from the last equatio under the T-adjusted probability measure the instantanous forward rate is equal to the Q_T- expected spot rate, hence we can write $r^F(t,T)$ as[16]:

$$r^{F}(t,T) = \left[\psi(T) + D(t,T)q - \sigma^{2}\int_{t}^{T} B(s,T)D(s,T)ds - \frac{1}{2}V(t,T)\right]^{2} + \frac{1}{2}V(t,T) \tag{52}$$

For t = 0 we get that q(0) = 0 because $\psi(0)^2 = r(0)$, which means:

---

[16] In Appendix E the approach that leads to this result is shown with respect to the extended Vasicek model.

$$r^F(0,T) = \left[\psi(T) - \sigma^2 \int_0^T B(s,T)D(s,T)ds - \frac{1}{2}V(0,T) \right]^2 + \frac{1}{2}V(0,T) \qquad \textbf{(53)}$$

By re-arranging terms in formula 53 we find that ψ(T) is of the following form:

$$\psi(T) = F(T) + \sigma^2 \int_0^T D(s,T)B(s,T)ds$$

$$F(T) = \sqrt{r^F(0,T) - \frac{1}{2}V(0,T)} + \frac{1}{2}V(0,T) \qquad \textbf{(54)}$$

Which is only valid when $r^F(0,T) \geq \frac{1}{2}V(0,T)$.

This property clearly means that the model cannot be fitted to all initial term structures if $\frac{1}{2}V(0,T) > r^F(0,T)$.

From the last section, we know that B(t,T) is a function of ψ(T), hence formula 54 is an integral equation in ψ(T). In Appendix D, it is in this respect shown that it is possible to express ψ(T) as:

$$\psi(T) = F(T) + 2e^{-\kappa T} \int_0^T e^{\kappa s)} V(0,s)F(s)ds \qquad \textbf{(55)}$$

Now, following the approach in Jamshidian (1996), it is more efficient to express A(t,T) and B(t,T) in formula 40 in terms of the initial term structure and B(0,T) and A(0,T). As B(t,T) and A(t,T) are specified for now, we are in a position where we have to use a numerical integrations technique for every value we need of A(t,T) and B(t,T).

In the rest of this section, I will therefore show how to evaluate P(T_1,T) in terms of P(0,T_1) and P(0,T).

Proposition no. 1

We find generally that A(T_1,T), B(T_1,T) and C(T_1,T) at time t for $0 < t < T_1 < T$ can be written as:

$$
\begin{aligned}
A(T_1,T) &= [A(t,T) - A(t,T_1)] - \frac{1}{2}\ln F - \left[\frac{\frac{1}{2}B(T_1,T)^2V(t,T_1) + B(T_1,T)G - C(T_1,T)G^2}{F}\right] \\
B(T_1,T) &= \frac{[B(t,T) - B(t,T_1)]F}{D(t,T_1)} + 2C(T_1,T)G \\
C(T_1,T) &= \frac{C(t,T) - C(t,T_1)}{D(t,T_1)^2 - 2V(t,T_1)[C(t,T) - C(t,T_1)]} \\
&\quad where \\
F &= 1 + 2C(T_1,T)V(t,T_1) \\
G &= \sigma^2\int_t^{T_1} D(s,T_1)B(s,T_1)ds
\end{aligned}
\tag{56}
$$

For t = 0 we find that A(0,T), B(0,T) and C(0,T):

$$
\begin{aligned}
A(0,T) &= \ln P(0,T) \\
B(0,T) &= 2\int_0^T D(0,s)\sqrt{r^F(0,T) - \frac{1}{2}V(0,s)ds} \\
C(0,T) &= C(0,T)
\end{aligned}
\tag{57}
$$

Proof:

We have the following property (see Appendix E):

$$
P(t,T) = P(t,T_1)E^{Q_{T_1}}[P(T_1,T)|F\text{-}_t]
\tag{58}
$$

Plucking in the expression for P(t,T) from formula 37 and the definition for $E^{Q_{T_1}}[P(T_1,T)|F_t]$ *from formula 5 in Appendix C given* $T^F = T_1$ *and re-arranging terms we get:*

$$
\begin{aligned}
& e^{A(t,T) - A(t,T_1) - [B(t,T) - B(t,T_1)]q - [C(t,T) - C(t,T_1)]q^2} \\
&= \sqrt{F}e^{A(T^F,T) + \left[\frac{\frac{1}{2}B(T_1,T)^2V(t,T_1) - B(T_1,T)M(t,T_1) - C(T_1,T)M(t,T_1)^2}{F}\right]}
\end{aligned}
\tag{59}
$$

If we use the definition of $M(t,T_1)$ *from equation 41 (given* $T^F = T_1$*) and collecting terms of equal power in q, we get:*

$$\begin{aligned}
A(t,T) - A(t,T_1) &= A(T_1,T) + \frac{1}{2}\ln F + \left[\frac{\frac{1}{2}B(T_1,T)^2V(t,T_1) + B(T_1,T)G - C(T_1,T)G^2}{F}\right] \\
B(t,T) - B(t,T_1) &= \frac{D(t,T_1)[B(T_1,T) - 2C(T_1,T)G]}{F} \\
C(t,T) - C(t,T_1) &= \frac{C(T_1,T)D(t,T_1)^2}{F}
\end{aligned} \qquad \textbf{(60)}$$

Now equation 56 follows directly from here.

For t = 0 we have for all T > 0 that A(0,T) = ln P(0,T) (as q = 0 for t = 0) and we also have that $G_0 = \psi(T) - F(T)$, using this we can rewrite formula 60 as:

$$\begin{aligned}
A(T_1,T) &= \ln\left(\frac{P(0,T)}{P(0,T_1)}\right) - \frac{1}{2}\ln F_0 \\
&- \left[\frac{\frac{1}{2}B(T_1,T)^2V(0,T_1) + B(T_1,T)[\psi(T_1) - F(T_1)] - C(T_1,T)[\psi(T_1) - F(T_1)]^2}{F_0}\right] \\
B(T_1,T) &= \frac{[B(0,T) - B(0,T_1)]F_0}{D(0,T_1)} + 2C(T_1,T)[\psi(T_1) - F(T_1)] \\
C(T_1,T) &= \frac{C(0,T) - C(0,T_1)}{D(0,T_1)^2 - 2V(0,T_1)[C(0,T) - C(0,T_1)]}
\end{aligned} \qquad \textbf{(61)}$$

where F_0 is F evaluated at t = 0 and G_0 is G evaluated at t = 0.

The only unknown factor in this relationsship is B(0,T) for all T. In order to solve this problem we will employ the same technique we used in Appendix E with respect to the extended Vasicek model.

From equation 53 we have an expression for $r^F(0,T)$, but we are also able to describe the instantanous forward rate by using formula 49, hence:

$$-A_T(0,T) + B_T(0,T)q + C_T(0,T)q^2 = \left[\psi(T) + D(0,T)q - G_0 - \frac{1}{2}V(0,T)\right]^2 + \frac{1}{2}V(0,T) \qquad \textbf{(62)}$$

Using formula 62 and collecting terms of equal power in q we find that $B_T(0,T)$ can be written as follows:

$$B_T(0,T) = 2\psi(T)D(0,T) - 2D(0,T)G_0 - D(0,T)V(0,T() \qquad \textbf{(63)}$$

Using the knowledge that $G_0 = \psi(T) - F(T)$, *we can simplify the expression for* $B_T(0,T)$ *as:*

$$B_T(0,T) = 2D(0,T)\sqrt{r^F(0,T) - \frac{1}{2}V(0,T)} \tag{64}$$

Formula 57 now follows, which completes the argument.

Qed.

It follows from this that it is only necessary to valuate B(0,T) for all T - that is we can store the calculations.

7: A Discrete Time Version of the Quadratic Interest Rate Model

If we want to price instruments with American style features - that is path-dependent securities, or we need to price European style instruments where no analytical expression is available we can build a discrete model.

One possibility is to discretize the PDE using the trinomial model from Hull and White (1994) which of couse is directly applicable for the Quadratic model. It should nevertheless with respect to the Hull and White trinomial model be mentioned that it might not be possible to use the analytical expression for the arbitrage-free bond prices in connection with the trinomial approach[17].

What I mean is that if for example you want to price a 3-year option on a 10-year bond you might have to build the lattice all the way up to 10-years instead of just to 3-years. The reason for that is that the Hull and White trinomial model is modelled in the Δt-interest rate whereas in the analytical bond price expression you use the continous time interest rate. This difference in perspective means that you have to convert the Δt-interest rate given in the lattice to the continous time interest rate that you need to use in the analytical expression for the bond price. This conversion is in many cases not possible, as the continous time interest rate is the solution to a second order equation which might not have a positive discriminant.

For that reason, we will define a new method for building discrete time models for one-factor interest rate models. The idea is to build a lattice in the continous time interest rate instead of the Δt-interest rate.

We will construct a grid which is defined in changes in q and t. The steps along the y-axis will be Δq and along the x-axis Δt. We will define a node (i,j) on the lattice as a point where $t_i = i\Delta t$ and $q_j = j\Delta q$. Every node on the grid will be an approximation to the true value of the security, and will for node (i,j) be denoted as $g_{i,j}(t,q)$. From Clewlow (1992) we have that in the explicit finite difference technique the partial derivatives of g(t,q) at node (i,j) are approximated as follows:

---

[17] Which in general is a very efficient method.

$$\begin{aligned} g_q(t,q) &\approx \frac{g_{i+1,j+1}(t,q) - g_{i+1,j-1}(t,q)}{2\Delta q} \\ g_t(t,q) &\approx \frac{g_{i+1,j}(t,q) - g_{i,j}(t,q)}{\Delta t} \\ g_{qq}(t,q) &\approx \frac{g_{i+1,j+1}(t,q) + g_{i+1,j-1}(t,q) - 2g_{i+1,j}(t,q)}{(\Delta q)^2} \end{aligned} \qquad \textbf{(65)}$$

In order to solve the PDE from equation 36 (under the risk-adjusted process) we have to impose the restrictions from formula 65 on every node (i,j) - that is the approximations of the partial derivatives have to satisfy the partial differential equation exactly. By combining the last equation with formula 65 we find the following approximated value for g(t,q) in node (i,j):

$$\begin{aligned} g(t,q) &= \frac{1}{1 + [q_j + \psi(t_i)]^2 \Delta t}[p_u g_{i+1,j+1}(t,q) + p_m g_{i+1,j}(t,q) + p_d g_{i+1,j-1}(t,q)] \\ &\textit{where} \\ p_u &= \frac{1}{2}\left[\frac{\sigma^2}{(\Delta q)^2} \mp \kappa j\right]\Delta t \\ p_m &= \left[\frac{1}{\Delta t} - \frac{\sigma^2}{(\Delta q)^2}\right]\Delta t \\ p_d &= \frac{1}{2}\left[\frac{\sigma^2}{(\Delta q)^2} \pm \kappa j\right]\Delta t \end{aligned} \qquad \textbf{(66)}$$

Along the lines of Hull and White (1994) we impose[18] the following relationship $\Delta q = \sigma\sqrt{3\Delta t}$, hence we can rewrite p_u,p_m and p_d as[19]:

$$\begin{aligned} p_u &= \frac{1}{6} + \frac{1}{2}\kappa^* j\Delta t \\ p_m &= \frac{2}{3} \\ p_d &= \frac{1}{6} - \frac{1}{2}\kappa^* j\Delta t \end{aligned} \qquad \textbf{(67)}$$

---

[18] See Hull and White (1994) for the motivation.

[19] It should here be stressed that $\kappa^*\Delta t$ is defined as $e^{-\kappa\Delta t} - 1$.

From these equations it follows that the numbers p_u,p_m and p_d are positive and sum to 1 if $-\frac{2}{3}\frac{1}{\kappa\Delta t} < j < \frac{2}{3}\frac{1}{\kappa\Delta t}$, which allows us to conclude that they can be recognized as bra

probabilities. In order to prevent the probabilities from becoming negative we cannot use a lattice that is arbitrarily large, more precisely, we have to change branching process at some level j_{shift}.

We will therefore employ the following rule for the change of branching process: at some level $j^+ < \frac{2}{3}\frac{1}{\kappa\Delta t}$ we will from node (i,j) jump to either node (i+1,j), (i+1,j-1) or (i+1,j-2) that is the grid will be bounded at j^+ - where we will denote this the falling branching process. We will also at some level $j^- > -\frac{2}{3}\frac{1}{\kappa\Delta t}$ jump from node (i,j) to either node (i+1,j+2), (i+1,j+1) or (i+1,j) that is the grid will be bounded at j^- - where we will denote this the rising branching process.

For j respectively equal to -1 or +1 we find that the probabilities in the falling and the rising branching process are defined as:

$$
\begin{aligned}
&\mathit{TheFallingBranchingProcess}\\
&p_u = \frac{7}{6} + \frac{3}{2}\kappa^* j\\
&p_m = -\frac{1}{3} - 2\kappa^* j\\
&p_d = \frac{1}{6} + \frac{1}{2}\kappa^* j\Delta t\\
&\mathit{TheRisingBranchingProcess}\\
&p_u = \frac{1}{6} - \frac{1}{2}\kappa^* j\Delta t\\
&p_m = -\frac{1}{3} + 2\kappa^* j\Delta t\\
&p_d = \frac{7}{6} - \frac{3}{2}\kappa^* j
\end{aligned}
\qquad \textbf{(68)}
$$

Where the probabilities in the falling branching process are positive and sum to 1 if $\frac{1}{6}\frac{1}{\kappa\Delta t} < j^+ < \frac{1}{3}\frac{1}{\kappa\Delta t}$, and the probabilities in the rising branching process are positive sum to 1 if[20] $-\frac{1}{3}\frac{1}{\kappa\Delta t} < j^{-} < -\frac{1}{6}\frac{1}{\kappa\Delta t}$ ..

This trinomial approach defined above is quite similar to the trinomial lattice model from Hull and White (1994). Both numerical approximations methods are equal to the first order of Δt. As mentioned earlier, the main difference is that in the trinomial model that is implemented here the spot interest rate is modelled directly using $r = [q + \psi(t)]^2$, whereas in the Hull an White trinomial model it is the Δt-period interest rate that is modelled.

This difference has the following two implications - one positive and one negative: Firstly, in the method advocated here we are able to use the analytical expression for ψ(t) (from formula 55) which means that we do not need to use forward induction in building the lattice - which clearly speeds up the calibration. Secondly, the method devised here has that undesired property that the initial term structure is not matched by construction - as is the case when using forward induction in the Hull and White trinomial approach[21]. We will for this reason prefer (if possible) to use the Hull and White approach.

For the pricing of generel path-dependent interest rate derivatives a lattice method might not be the most efficient technique[22].

For the general problem of handling path-dependency, two different kinds of method can be employed:

Monte Carlo sampling from the lattice
Monte Carlo simulation

In this paper I will not discuss how to employ efficient sampling techniques in connection with lattice based methodology, this will be left for future research.

In section 8 I will for that reason show how MC techniques can be applied to one-factor interest rate models with a time-dependent drift part[23].

---

[20] In practise we define $j_{shift} = j^{+} = |j^{-}| = \left|\frac{1}{6}\frac{1}{\kappa\Delta t}\right|$ - that is we change branching pro given opportunity.

[21] However it is worth mentioning that the initial yield curve is only matched at discrete maturity-points.

[22] Hull and White (1993b) have shown that certain kinds of path-dependency can be handled in a lattice-based model. Their method is however not easy to generalise.

[23] Extending the method to the case of time-dependent diffusion parameters is however straightforward,

8: Time-dependent SDEs and Monte Carlo simulation

The main idea of the Monte Carlo method can be explained by a simple example. Consider the valuation of the expected value of a random variable X distributed according to its probability density f(x). Suppose we can, by whatever means, sample a point from the distribution of X, then a direct way to calculate the expectation of X is to independently sample a large number of points, say N points, from the distribution of X and compute the arithmetic mean of these points, as:

$$X_N = \frac{x_1 + x_2 + x_3 + \dots\dots + x_N}{N} \qquad \textbf{(69)}$$

Where x_i, for i = 1,2,....,N, are the points sampled from f(x). If the second moment of X is finite, then by the law of large numbers, $X_N \rightarrow E(X)$ as $N \uparrow \infty$, with probability 1. This means that for a sufficient large number of N and any particular realisation of the set of sampled points from f(x), the arithmetic mean of these points is very close to the mean of X. In fact, the mean of X_N is the same as the mean of X, and the variance of X_N is:

$$\begin{aligned} Var(X_N) &= E\left[\left(\sum_{n=1}^{N} \frac{1}{N} x_n - E(X)\right)^2\right] \\ &= \frac{1}{N^2}\sum_{n=1}^{N} E[x_n - E(X)]^2 + \frac{1}{N^2}\sum_{1 \le i < j \le N}^{N} E[(x_i - E(X))(x_j - E(X))] \end{aligned} \qquad \textbf{(70)}$$

Since all x_i are identically and independently sampled from f(x), the correlation among these x_i is zero and the variance of x_i is equal to that of X. Therefore $Var(X_N) = Var(X)/N$. This relationship is essential to the understanding of Monte Carlo. On one hand it ensures that as the number of sample points increases, the error of the estimated expected value of X decreases. On the other hand, it shows the difficulty of the Monte Carlo method in achieving a high degree of accuracy: the standard deviation descreases only as $O\left(N^{-\frac{1}{2}}\right)$.

As the complexity of traded instruments has grown, Monte Carlo techniques have become increasingly important, a trend that is likely to continue in the future. This is particularly true in the fixed income markets, where multiple factors and path-dependency are embedded in a wide variety of new structured or mortgage-based contracts. Despite its increasing importance and wide applicability, the Monte Carlo method has received much less attention in the litterature than lattice models. One reason for this is, of course, the low order of convergence $O\left(N^{-\frac{1}{2}}\right)$ which tends to make practical usage painfully slow.

In modern finance, the prices of the basic securities (the underlying state variables) are often modelled by specifying one or more Stochastic Differential Equations (SDEs). The pricing of, for example, an option that depends on one or more of these basic securities can, under the assumption of no arbitrage, be expressed as the expected value of its discounted payouts. Where this expectation is taken with respect to an appropriate transformation of the original probability measure.

The Monte Carlo method lends itself naturally (as shown above) to the evaluation of security prices represented as expectations. The Monte Carlo approach consist of the following steps:

> Simulate sample paths of the underlying state variable/s (underlying asset price and interest rate) over the relevant time horizon. Where the simulation is performed under the risk-neutral measure.
> Evaluate the discounted cash-flows of the security on each sample path, as determined by the structure of the security in question
> Average the discounted cash-flows over the sample paths

In principle this method computes a multi-dimensional integral - the expected value of the discounted payouts over the space of sample paths.

It might be worth pointing out that in the case of time-dependent SDEs for interest rate models we have to condition the path-sampling process - a method to solve this problem will be developed later in this section.

Before continuing, let me for a moment dwell on the pricing of contingent claims using the risk-neutral density pricing approach[24]. For a derivative asset with payoff defined by H(y(T)) and price g(y,t,T) we have:

$$g(y,t,T) = P(t,T)\int_{-\infty}^{\infty} H(y(T))p(y(T))dy(T) \qquad \mathbf{(71)}$$

Where p(y(T)) is the risk-neutral density function. Formula 71 represents the price at time t of a contingent claim that has a payoff defined by H(y(T)) at time T which is a function of the process that governs the underlying state-variable - here represented by y.

Since the risk-neutral density pricing approach expresses all asset prices as the discounted value of their expected payoff formula 71 is identical to formula 5. Equation 71 (and therefore equation 5) is very powerful due to the fact that any derivative security may be priced by this approach[25].

---

[24] The risk-neutral density function p(y(T)) is also known as Green's function or the fundamental PDE solution. p(y(T)) also satisfies the Kolmogorov forward (or Fokker-Planck) equation with a boundary condition given by the Dirac delta-function, see Cox and Miller (1965 chapter 5) and sections 2 and 3 in this paer.

[25] This is also true in the case of multivariate contingent claims. The only difference between valuing multivariate and univariate claims is that the payoff function for the multivariate claims involves the prices of multiple underlying assets, and the expectation is taken over all the underlying asset prices. The solution to formula 71 is now found by computing a multi-dimensional integral.

The central point in the valuation of derivative securities is the expression for the risk-neutal density function p(y(T)) - which is implied from the SDE/s for the underlying state variable/s.

In this paper we have considered two different assumptions about the spot-rate process - which both turned out to imply a normal risk-neutral density function[26].

In connection with the pricing of stock-options the most common assumption is to assume a lognormal risk-neutral density for the stock price - that is adopt the assumption from Black and Scholes (1973).

The normal risk-neutral density assumption for interest rate models and the lognormal risk-neutral density assumption for stock models is not neccesarily the appropriate assumption to use, as departures from both the normality and lognormality are well documented in many financial time-series, see for example Bollerslev, Engle and Nelson (1993). These assumptions are primarily used for practical reasons because of their great tractability.

An alternative approach could therefore be to use a non-parametric approach in the determintion of the density function. One approach could be to use the kernel estimator to determine the empirical density function, for additional information see Scott (1992 chapter 6), Ait-Sahalia (1996a,1996b) and Campbell, Lo and Mackinlay (1997 chapter 12 section 12.3)[27].

An important issue in connection with Monte Carlo simulation is how to improve the efficiency without the need to make N (the number of paths) unreasonable large - and therefore computation time too slow for practical purposes.

Using Monte Carlo simulation we can express the value of g(y,t,T) as:

$$g(y,t,T) \approx \frac{1}{N}\sum_{i=1}^{N} g_i(K) \tag{72}$$

Where K is the dimension of the problem - that is the number of time-steps used in the discrete time approximation of the SDE. $g_i(K)$ is the discounted value of the payoff conditioned on the ith sample path. Furthermore we have that $g_i(K)$ is derived from the random variables, as:

---

[26] Assuming a normal risk-neutral density function for the interest rate model does not neccesarly imply that interest rates can become negative - it does imply that the state variable can become negative but the probability of negative interest rates depends on the relationship between the state variable and the spot-rate (se section 2).

[27] A number of techniques for estimating state prices have been proposed in the context of option pricing, see for example Jackwerth and Rubenstein (1996) and Rosenberg (1997). Both these approaches imply the risk-neutral density from quoted options prices. Other related works are Andersen and Brotherton-Ratcliffe (1998) which utilises a Crank-Nicholson finite-difference scheme and Dupire (1994) who uses the explicit finite-difference scheme.

$$g_i(K) = S(z_1(i),\ z_2(i),\ \ldots.,\ z_K(i)) \qquad \textbf{(73)}$$

Where $z_1(i)$, $z_2(i)$,...,$z_K(i)$ is an independent sequence of standard Gaussian variables and S: $\Re^k \rightarrow \Re$ is a well-behaved function which is a function of the used discretization scheme[28].

To enhance the efficiency of the Monte Carlo simulation, two routes can be taken (or combinations hereof):

Reduce the variance of the sampled variance
Overcome the $O\left(N^{-\frac{1}{2}}\right)$ barrier

Different variance reduction techniques have been developed to increase precision. Two of the classical variance reduction methods are the control variate approach and the antithetic variate method. Recently, moment matching, importance sampling, empirical martingale simulation techniques, the Brownian-Bridge method, conditional Monte Carlo methods and variance reduction by Girsanovs transformation have been introduced in finance litterature, see for example Boyle, Broadie and Glasserman (1995), Duan and Simonato (1998), Caflisch, Morokoff and Owen (1997) and Schoenmakers and Heemink (1997)[29].

To overcome the $O\left(N^{-\frac{1}{2}}\right)$ barrier a Quasi-Monte Carlo simulation can be employed. This method uses a determinstic sequence that is more uniform than random[30].

This can best be explained by rewriting equation 73 as:

$$g_i(K) = S(z_1(i),\ z_2(i),\ \ldots.,\ z_K(i)) = Q(u_1(i),\ u_2(i),\ \ldots..,\ u_K(i)) \qquad \textbf{(74)}$$

Where $u_1(i)$, $u_2(i)$,...,$u_K(i)$ is an independent sequence of standard uniform U(0,1)-variables and Q: $\Re^k \rightarrow \Re$ is a well-behaved function[31].

---

[28] By this we mean - which kind of discretization has been used for the SDE - this will be explained in more detail later in this section.

[29] In Appendix G, I will briefly explain the different variance reduction techniques used later in this paper.

[30] Using the word random - does not mean truly random as the "random" numbers are generated by a deterministic algorithm and are described as pseudorandom numbers. The pseudorandom number generator used in the examples in section 9 is an example of the linear congruential kind introduced in 1948 by the mathematician D. H. Lehmer. It depends on two numbers, 16.807 and 2.147.483.647. The first is 7^5, a primitive root of the second, which is Eulers prime, 2^{31}. We have repeated the results obtained in this paper using the Mersenne twister (see http://en.wikipedia.org/wiki/Mersenne_twister), but the results we get are nearly identical.

[31] The transformation from Gaussian to uniform variates can be accomplished by inverting the cumulative Gaussian distribution or through the inverse of the Box-Mueller transformation.

Using 74 the expectation of g(y,t,T) can be written as an integral over the k-dimensional hypercube:

$$g(y,t,T) = E_t[g(K)] = \int_{[0;1]^K} Q(x_1, x_2, \ldots, x_K)dx_1dx_2\ldots dx_K \approx \frac{1}{N}\sum_{i=1}^{N} Q(x_i) \qquad \textbf{(75)}$$

Formula 75 is identical to formula 72 except that we have not specified which sampling scheme has been used. If the sampling scheme is Monte Carlo (that is, based on a pseudorandom number generator) we know that the expected error in formula 75 is independent of the dimension k and proportional to $N^{-\frac{1}{2}}$. To improve this convergence criterium, several deterministic sampling algorithms have been suggested instead of Monte Carlo simulation.

Specific algorithms for generating quasi-random numbers has been suggested, for example Sobol (1967), Halton (1960) and Faure (1982), see Press, Teukolsky, Vetterling and Flannery 1992 (chapter 7) for a good description of these methods.

Due to Niederreiter (1992 page 20) one can show (the Koksma-Hlawka Inequality) that under technical conditions on the function Q, the Sobol and Halton sequences generate errors which decrease with N^{-1} - at least as $\frac{(\log N)^K}{N}$[32].

For small dimensions the discrepancy for quasi-random number generators appears to be $O(N^{-1})$, ignoring logarithmetic factors, for all N. For large dimensions, the discrepancy behaves initially like $O(N^{-\frac{1}{2}})$ - as for the random number sequence - converging only to $O(N^{-1})$ for very large values of N.

The value of N appears to grow exponentially with the dimensions - which implies that in high dimensions - unless N is extremely high - quasi-random number sequences are no more uniform than random number sequences. Thus the Koksma-Hlawka bound does not imply that quasi-random number sequences are more efficient than random number sequences for moderate values of N and large dimensions K.

It is usually the case that Quasi-Monte Carlo methods are superior to Monte Carlo simulation

---

[32] Which is smaller than the error for Monte Carlo simulation because the error bound $O\left(\frac{(\log N)^K}{N}\right)$ is smaller than $O\left(N^{-\frac{1}{2}}\right)$ as N → ∞.

(or at least no worse), see for example Paskov (1996), Acworth, Broadie and Glasserman (1997) and Caflisch, Morokoff and Owen (1997) - even in high-dimensionel situations. Furthermore, the results from Paskov (1996) indicate that the Sobol sequence frequently outperforms other kinds of quasi-random number sequences.

So far we have discussed varies aspects in connection with Monte Carlo simulation and mentioned ways to improve the efficiency. Let us now turn our attention to the SDE we wish to simulate:

$$\begin{gathered} dy = (\varphi(t) - \kappa y)dt + \sigma dW \\ \text{where} \\ r = F(t,y) \end{gathered} \tag{76}$$

which can be recognized as being identical to formula 1 - for the diffusion coefficient defined as $\sigma(y,t) = \sigma$ and the drift coefficient defined as $\mu(\varphi(t),y,t) = \varphi(t) - \kappa y$.

For the purpose of simulation let us instead consider the following diffusion process (SDE):

$$\begin{gathered} dx = -\kappa x dt + \sigma dW \quad \text{for } x_0 = 0 \\ \text{where} \\ y = x + \varphi(t) \end{gathered} \tag{77}$$

Generally speaking, a diffusion process is an arbitrary strong Markov process with continous sample paths. In our framework, a diffusion process is given as a strong solution[33] of an SDE driven by the underlying Brownian motion W.

With respect to discrete time approximations of continous time processes, it is important to distinguish between strong and weak convergence. The weak convergence criterion does not (as the strong convergence) require a pathwise approximation of the Ito-process, but only an approximation of the probability distribution (see Kloeden and Platen (1995 chapter 4, section 9.6 and 9.7).

In Kloeden and Platen (1995 parts V and VI) it is shown that the strong and weak convergence criteria leads to the development of different sampling schemes which are only efficient with respect to one of the two criteria. This means that a given sample scheme usually has different orders of convergence with respect to the two criteria.

We will here consider three different strong convergence sampling schemes, the Euler scheme, the Milstein scheme and the 1.5 strong Taylor scheme. Futhermore we will consider two weak convergence sampling schemes, the Euler scheme and the 2. order Milstein scheme.

---

[33] A strong solution to formula 77 requires that the SDE satisfies the Lipschitz condition, the linear growth bound and initial value condition. It is furthermore required that the coefficient in the SDE is continuous - with imply measurability, see Kloeden and Platen (1995, sections 4.5 and 4.6).

Before I go through the three different sampling schemes let me make the following definitions.

We will here consider a time-discretization τ of the time interval [t,T]:

$$t = \tau_1 < \tau_1 < \tau_2 < .... < \tau_k < .... < t_K = T \tag{78}$$

Where K denotes the number of discretization intervals $\Delta_k = \tau_{k+1} - \tau_k$. As we only consider the case of equidistant discretization points, we have that $\delta = \Delta_k = \frac{T - t}{K}$ for all k.

8.1 The Euler scheme

The Euler discretization scheme is given by:

$$x_{k+1} = x_k - \kappa x_k \delta + \sigma \Delta W \tag{79}$$

Where $\Delta W = W_{\tau_{k+1}} - W_{\tau_k}$ is the N(0,δ) distributed increments of the Wiener process W o $[\tau_k, \tau_{k+1}]$.

This discretization scheme represents the simplest strong Taylor approximation and attains the order or convergence 0.5, see Kloeden and Platen (1995 section 10.2).

Furthermore, we have that the Euler weak approximation is of a form identical to equation 79, however, here the order of convergence is 1, see Kloeden and Platen (1995 section 14.1).

8.2 The Milstein scheme - strong 1. order Taylor approximation

The Milstein discretization scheme is defined as:

$$x_{k+1} = x_k - \kappa x_k \delta + \sigma \Delta W + \frac{1}{2}\sigma^2((\Delta W)^2 - \delta) \tag{80}$$

which (as seen from the equation) is obtained by adding the term $\frac{1}{2}((\Delta W)^2 - \delta)$ to the Euler scheme.

In Kloeden and Platen (1995 section 10.3) it is shown that the Milstein scheme converges strongly with order 1.

8.3 The order 1.5 strong Taylor scheme

The order 1.5 strong Taylor scheme can in abstract form be written as[34]:

$$x_{k+1} = x_k + \mu\delta + \sigma\Delta W + \frac{1}{2}\sigma\sigma_x((\Delta W)^2 - \delta) + \mu_x\sigma\Delta Z + \frac{1}{2}\left(\mu\mu_x + \frac{1}{2}\mu_{xx}\sigma^2\right)\delta^2 + \left(\mu\sigma_x + \frac{1}{2}\mu_{xx}\sigma^2\right)(\Delta W\delta - \Delta Z) + \frac{1}{2}\sigma\left(\sigma\sigma_x + (\sigma_x)^2\right)\left(\frac{1}{3}(\Delta W)^2 - \delta\right)\Delta W \quad \textbf{(81)}$$

$$\textit{where}$$

$$\mu = -\kappa x_k$$

Where the footsigns (x) represent partial derivatives with respect to x. The additional random variable ΔZ is defined by the following double integral:

$$\Delta Z = \int_{\tau_k}^{\tau_k+1}\int_{\tau_k}^{s_2} dW_{s_1} ds_2 \quad \textbf{(82)}$$

where we have that the random variable ΔZ is distributed as $N\left(0, \frac{1}{3}\delta^3\right)$. Furthermore we have that the covariance $E[\Delta Z\Delta W] = \frac{1}{2}\delta^2$.

For our particular SDE - given by formula 77 - we can simplify formula 81and get the following expression for the order 1.5 strong Taylor scheme:

$$x_{k+1} = x_k - \kappa x_k\delta + \sigma\Delta W - \kappa\sigma\Delta Z + \frac{1}{2}\kappa^2 x_k\delta^2 \quad \textbf{(83)}$$

8.4 The order 2 weak Milstein scheme

The order 2 weak Milstein scheme can in abstract form be written as[35]:

$$x_{k+1} = x_k + \mu\delta + \sigma\Delta W + \frac{1}{2}\sigma\sigma_x((\Delta W)^2 - \delta) + \mu_x\sigma\Delta Z + \frac{1}{2}\left(\mu\mu_x + \frac{1}{2}\mu_{xx}\sigma^2\right)\delta^2 + \left(\mu\sigma_x + \frac{1}{2}\mu_{xx}\sigma^2\right)(\Delta W\delta - \Delta Z) \quad \textbf{(84)}$$

$$\textit{where}$$

$$\mu = -\kappa x_k$$

Where the footsigns (x) represent partial derivatives with respect to x. The variable ΔZ is defined in formula 82. Comparing formula 84 to formula 81 it is easily seen that the order 2 weak Taylor approximation is equal to the order 1.5 strong Taylor approximation if the last term in line 2 in equation 81 is removed[36].

---

[34] The order of convergence is shown in Kloeden and Platen (1995 section 10.4).

[35] The order of convergence is shown in Kloeden and Platen (1995 section 14.2)

For our particular SDE, we actually have that the order 1.5 strong Taylor approximation is equal to the order 2 weak Taylor approximation - that is equal to formula 83.

Before deciding to use either the higher order strong-or weak-discrete approximations[37] it is important to clarify the aim of a simulation. The following question should be answered:

Is a good pathwise approximation of the Ito process required or is the approximation of some functional of the Ito process the real objective?

In many cases, also the examples considered in section 9, the weak discrete-time approximation is the appropriate one as we in many cases are just interested in the probability distribution and its moments at given times T[38].

An interesting point to emphasize at this point is that equation 77 is actually a fairly general formulation[39] for one-factor interest rate models with a time-dependent drift term as we by making appropriate choices for F(t,y) (as mentioned in the beginning of section 2) get the Extended Vasicek model from Hull and White (1990) (F(t,y) = y), Black and Karisinski (1991) ($F(t,y) = e^y$), the Quadratic interest rate model ($F(t,y) = y^2$) and the Cox, Ingersoll and Ross (1985) model for $F(t,y) = \sqrt{y}$.

From the points discussed so far it is possible to conclude that the efficiency of simulating the diffusion process defined by the SDE in formula 77 is a function of the following:

The discretization scheme
The Monte Carlo simulation procedure

In section 9 we will for the Extended Vasicek model and the Quadratic interest model for the case of a call and put option on a zero-coupon bond analyse some combination of discretization schemes and Monte Carlo methods. Furthermore we will as comparison price the same instruments analytically and by using a lattice approach.

---

[36] Higher order strong-and weak discretization schemes can of course also be derived but for strong schemes above 1.5 and weak schemes above 2 the expressions get pretty complicated. As the expansions to higher order approximations of the SDE leads to non-trivial multiple stochastic integrals. An atractive alternative to the truncated Taylor series expansion - at least in the case of weak discretizations shemes - is to use Richardson or Romberg extrapolation, see Kloeden and Platen (1995 page 285).

[37] We are here referring to discrete schemes of higher order than Euler schemes.

[38] A situation where we would be interested in a pathwise approximation of the Ito process is in the case of filtering.

[39] Expanding to the case of more complex drifts specifications or for that sake diffusion coefficients is relatively straightforward. This issue will however not be pursued here - as the formulation chosen here covers the typical assumptions used in the literature.

Before that we will in the rest of this section show a method to condition the path-sampling process in order to ensure that the initial yield-curve is matched.

8.5 Constrained Monte Carlo simulation

Our basic SDE is (as shown above) defined as:

$$dx = -\kappa x dt + \sigma dW \quad for\ x_0 = 0 \tag{85}$$

Let us assume that we have split the time interval [t,T] into K discretization intervals of an equal size, that is $\delta = \frac{T-t}{K}$. Furthermore we assume that we have performed N simulations of the SDE.

That is we have an NxK matrix of simulated values for x - which we will denote the diffusion-matrix.

It might be worth mentioning that the procedure described below is independent of the chosen discretization scheme and the Monte Carlo method employed in the construction of the diffusion matrix - which of course makes the procedure extremely general and quite flexible.

Given is also the yield-curve (or equally the discount function), which means that we know the price P(t,kδ), for all k = [1,K] and where P(t,kδ) = 1 for t = kδ.

We now wish to construct a new diffusion matrix (RD-matrix) for the spot-rate, F(t,y) = r, where y = x + φ(t). That is we now introduce the correct, time-varying drift. To do this we will change the values of each cell in the diffusion-matrix. The change of the diffusion-matrix will however be equal for all k's as φ(kδ) is only time-dependent and not path-dependent.

The method - which is a forward-induction principle - can be described as[40]:

Step 1: Column 1 in the RD-matrix is known from the yield-curve and is defined as $r_1 = \frac{-\ln P(t,\delta)}{\delta}$. Where we therefore have that φ(1) = r_1.

Step 2: We will now introduce the concept of Arreu-Debreu prices.

An Arreu-Debreu is a security that has a payoff of 1 at time kδ in state n = [1,N]. In general we have that the vector of Arreu-Debreu prices at time kδ can be written as:

---

[40] Forward-induction in connection with Monte Carlo and time-dependent spot-interest rate models might have been around for some time but until now this technique has not (as far as I know) been described in the literature.

$$Q(t,k\delta) = e^{-\tilde{r}_k\delta}Q(t,(k-1)\delta)$$
$$\text{where}$$
$$\tilde{r}_k = r_{k-1} + x_{k-1}\delta$$

We have that Q(t,kδ) = 1 for t = kδ, and $\tilde{r}_k$ represent the forecasted vector of spot-rates for the period kδ - (k+1)δ.

Step 3: The true vector of spot-rates can now be derived from the following equation:

$$P(t,k\delta) = Q(t,k\delta)e^{-F(t,\tilde{r}_k + \varphi(k\delta))\delta} = \frac{1}{N}\sum_{n=1}^{N} q_n(t,k\delta)e^{-F(t,\tilde{r}_k(n) + \varphi(k\delta))\delta}$$

In some case we can solve this expression directly for φ(kδ) - this is for example the case in the Extended Vasicek. When this is not the case we need to estimate φ(kδ) using a minimization algorithm - this can however (in most cases) be done in 1-2 iterations.

In some cases it is even possible to derive an analytical expression for φ(t). In these cases (as pointed out in section 7) we did not need to use forward-induction in calibrating the lattice as we could determine φ(t) analytically. This is however only possible for the Quadratic interest rate model, see formula 55 and for the extended Vasicek model, see Appendix E formula 6.

This approach could also be utilized here. If that was the case the same implications mentioned in connection with the lattice method in the end section 7 would apply here.

Step 4: Repeating Step 2-3 for all k > 1 builds the diffusion RD-matrix. Where the following equation by construction is now true for all k = [1,K]:

$$P(t,k\delta) = \frac{e^{-[RD_1 + RD_2 + \dots RD_k]\delta}}{N}$$

Where RD_k is the k'th row in the diffusion RD-matrix with a length equal to N. This equation just states that the average price from the sampled process is equal to the corresponding market price for all k = [1,K].

From the description above it follows that the method designed here - for the case of path-dependent simulation of the interest rate process - is very similar in spirit to the Hull and White (1994) lattice-building procedure.

9: Some pricing examples

In this section we will calculate European options-prices for both the Extended Vasicek model and the Quadratic Interest rate model.

In this connection we assume that the yield-curve is defined as: $0.08 - 0.05e^{-0.18(T-t)}$ Given this term structure we will now price 3-year call-and put-options where the underlying security is an 8-year zero-coupon bond.

From the yield-curve we can deduce that the forward-price P(3,8) = 67.53. Therefore we have chosen to calculate the option-prices for the following selection of strike-prices: 52.53, 62,53, 67.53, 72.53 and 82.53. For both models we have assumed that $\kappa = 0.05$ and $\sigma = 0.01$.

In table 1 we have shown the analytical values of the options for both the Extended Vasicek model and the Quadratic Interest rate model:

Table 1: **Analytical Option Prices**

	Call option prices		**Put option prices**	
Exercise Prices	**The Extended Vasicek Model**	**The Quadratic Interest Rate Model**	**The Extended Vasicek Model**	**The Quadratic Interest Rate Model**
52,53	*12,877*	*12,877*	*0,000*	*0,000*
62,53	*4,576*	*4,311*	*0,283*	*0,019*
67,53	*1,647*	*0,847*	*1,647*	*0,847*
72,53	*0,354*	*0,015*	*4,647*	*4,307*
82,53	*0,003*	*0,000*	*12,881*	*12,877*

In the case of the Extended Vasicek model we have also calculated the option prices by performing a simulation under the T^F-adjusted probability measure - where T^F symbolises the expiry-date of the option (that is $T^F = 3$). More precisely we have transformed the Extended Vasicek model into the Heath, Jarrow and Morton framework - which mean that it is possible to write the forward-price $P(T^F,T)$ as:

$$P(T^F,T) = \frac{P(0,T)}{P(0,T^F)}\exp\left[-\frac{1}{2}\int_0^{T^F}[\sigma_P(s,T) - \sigma_P(s,T^F)]^2 ds - \int_0^{T^F}[\sigma_P(s,T) - \sigma_P(s,T^F)]dW^{Q_{T^F}}(s)\right] \tag{86}$$

Where $\sigma_p(t,T)$ is the bond-price volatility - which is related to the forward-rate volatility through the following relationship:

$$\sigma_P(t,T) = \int_t^T \sigma^F(t,v)dv = \frac{\sigma}{\kappa}[1 - e^{-\kappa(T-t)}] \tag{87}$$

What we have done is simulated the forward-price $P(T^F,T)$ using equation 86. We have employed the following Monte Carlo simulation methods: Crude Monte Carlo, Antithetic Monte Carlo, Stratified Sampling and Empirical Martingale Simulation (EMS)[41].

We have performed B batches of N simulations, for B = 100 and N = 1000. Provided that B is sufficiently large, the Central Limit Theorem implies that the batch error is approximately Gaussian. The sampled variance can under this assumption be calculated as:

$$Var_B = \frac{1}{B(B-1)}\left(B\sum_{b=1}^{B}\hat{C}_b^2 - \left(\sum_{b=1}^{B}\hat{C}_b\right)^2\right) \qquad \textbf{(88)}$$

Where $\hat{C}_b$ is the estimated/simulated option-price for the b'th batch-run.

Table 2: Simulated Call-option prices in the Extended Vasicek Model

Excercise Price	**Crude Monte Carlo**	**Std. Error[42]**	**Antithetic Monte Carlo**	**Std. Error**	**Stratified Samp-ling**	**Std. Error**	**Empirical Monte Carlo Simulation**	**Std. Error**
52,53	*12,864*	*0,127*	*12,877*	*0,006*	*12,878*	*0,002*	*12,877*	*0,001*
62,53	*4,572*	*0,125*	*4,573*	*0,025*	*4,576*	*0,002*	*4,578*	*0,023*
67,53	*1,647*	*0,083*	*1,648*	*0,041*	*1,647*	*0,001*	*1,649*	*0,033*
72,53	*0,350*	*0,033*	*0,350*	*0,023*	*0,354*	*0,002*	*0,351*	*0,028*
82,53	*0,003*	*0,003*	*0,003*	*0,002*	*0,003*	*0,001*	*0,003*	*0,003*

---

[41] In Appendix G these methods are briefly explained.

[42] Std. Error is calculated as the square root of the sampled variance specified in formula 88.

Table 3: Simulated Put-option prices in the Extended Vasicek Model

Excercise Price	**Crude Monte Carlo**	**Std. Error**	**Antithetic Monte Carlo**	**Std. Error**	**Stratified Samp-ling**	**Std. Error**	**Empirical Monte Carlo Simulation**	**Std. Error**
52,53	*0,000*	*0,001*	*0,000*	*0,000*	*0,000*	*0,000*	*0,000*	*0,000*
62,53	*0,284*	*0,030*	*0,284*	*0,019*	*0,284*	*0,001*	*0,284*	*0,025*
67,53	*1,655*	*0,071*	*1,650*	*0,037*	*1,647*	*0,001*	*1,648*	*0,039*
72,53	*4,660*	*0,106*	*4,645*	*0,018*	*4,647*	*0,001*	*4,647*	*0,026*
82,53	*12,869*	*0,142*	*12,880*	*0,006*	*12,880*	*0,001*	*12,880*	*0,003*

From these two tables it can be concluded (as expected) that Crude Monte Carlo simulation performs worst. Furthermore it follows that of the variance reduction methods employed here Stratified sampling is the most efficient method, and EMS performs marginally better than Antithetic. It is usually the case that the Antithetic variates technique is the variance reduction method which is least efficient.

The reason why the EMS variance reduction method does not perform better than it does can be explained as follows: The main idea behind the EMS method is namely to adjust the sampling in such a way that the average price of the underlying instrument is equal to the expected price (forward-price) of the underlying instrument[43]. As the sampling using formula 86 is performed under the T^F-adjusted probability measure - which in a sense is to move the discounting outside the expectation - we have therefore already lowered the uncertainty in the sampling procedure, as we have[44]:

---

[43] Se Appendix G.

[44] The terms in the first expectation in formula 89 can be recognized as the Radon-Nikodym derivative (the likelihood ratio process) which under technical conditions for $\sigma_P(t,T)$ accordingly to Girsanov Theorem relates the exisisting probability measure Q to a new equivalent probability measure (in this example Q_{TF}), see Musiela and Rutkowski (1997 Appendix B.2).

$$E^{Q}\left[\exp\left[-\frac{1}{2}\int_0^{T^F}\sigma_P(s,T^F)^2 ds - \int_0^{T^F}\sigma_P(s,T^F)dW^{Q}(s)\right]\right] = 1$$

or equivalently **(89)**

$$E^{Q_{TF}}\left[\exp\left[-\frac{1}{2}\int_0^{T^F}[\sigma_P(s,T) - \sigma_P(s,T^F)]^2 ds - \int_0^{T^F}[\sigma_P(s,T) - \sigma_P(s,T^F)]dW^{Q_{TF}}(s)\right]\right] = 1$$

From this it follows that the technique employed using formula 86 in the simulation procedure is equal to the measure transformation method developed by Andersen (1995) in connection with Monte Carlo simulation of one-factor non-linear time-homogenous interest rate models.

Before we calculate the prices using Monte Carlo simulation for both the call-and put-options for the two different weak discretization schemes for the SDE from section 8 we will first calculate the option-prices using a lattice-model.

We will if nothing else is mentioned (if possible[45]) be using the Hull and White (1994) lattice approach.

Table 4[46]: Call-and Put-option prices in the Extended Vasicek Model using the Hull and White (1994) Lattice approach

Call-Option prices					
Number of time-steps	**8**	**16**	**32**	**64**	**128**
52,53	*12,880*	*12,879*	*12,878*	*12,878*	*12,878*
62,53	*4,603*	*4,579*	*4,574*	*4,579*	*4,576*
67,53	*1,631*	*1,645*	*1,649*	*1,650*	*1,649*
72,53	*0,383*	*0,362*	*0,357*	*0,358*	*0,356*
82,53	*0,004*	*0,003*	*0,003*	*0,003*	*0,003*

---

[45] See the discussion at the beginning of section 7 for cases where that is not possible. It is however of importance to mention here that the probability of not being able to use the Hull and White procedure to match the initial yield-curve - is larger in lattice-based models than in Monte Carlo based models. The reason for is because the interest-space that a lattice-based method spans is larger than the one covered in a Monte Carlo procedure. In contrast to this the interest-space that a Monte Carlo method covers is finer than the one covered by the lattice-based methodology. The conclusion now follows as the probability - of not being able to use the Hull and White procedure - is bigger the lower the interest rates, see formula 54.

[46] It is here worth noting that we in the calculation of the option-prices using the lattice approach have made use of the fact that we have a closed form solution for the bond-price, for further details see Appendix H.

Put-Option prices					
52,53	*0,000*	*0,000*	*0,000*	*0,000*	*0,000*
62,53	*0,308*	*0,285*	*0,281*	*0,286*	*0,284*
67,53	*1,628*	*1,643*	*1,649*	*1,650*	*1,649*
72,53	*4,673*	*4,653*	*4,648*	*4,650*	*4,648*
82,53	*12,880*	*12,880*	*12,880*	*12,880*	*12,880*

As is obvious from table 4, the convergence is very fast using a lattice-based methodology.

We have in table 5 performed the same calculations for the Quadratic interest rate model. As was the case for the Extended Vasicek model we have also here used the fact that it is possible to get a (semi) closed form solution for the bond-price[47].

Table 5: Call-and Put-option prices in the Quadratic Interest Rate Model using the Hull and White (1994) Lattice approach

Call-Option prices					
Number of time-steps	**8**	**16**	**32**	**64**	**128**
52,53	*12,879*	*12,878*	*12,878*	*12,877*	*12,877*
62,53	*4,313*	*4,313*	*4,311*	*4,312*	*4,311*
67,53	*0,841*	*0,845*	*0,846*	*0,847*	*0,847*
72,53	*0,015*	*0,017*	*0,015*	*0,015*	*0,015*
82,53	*0,000*	*0,000*	*0,000*	*0,000*	*0,000*
Put-Option prices					
52,53	*0,000*	*0,000*	*0,000*	*0,000*	*0,000*
62,53	*0,019*	*0,020*	*0,019*	*0,020*	*0,019*
67,53	*0,840*	*0,844*	*0,846*	*0,847*	*0,847*
72,53	*4,306*	*4,308*	*4,308*	*4,307*	*4,307*
82,53	*12,876*	*12,877*	*12,877*	*12,877*	*12,877*

---

[47] For elaboration see Appendix H.

Again we see, as was the case for the Extended Vaiscek model, a fast convergence to the true result.

As our first step we have now in comparison calculated the call-and put-option prices for both the Extended Vasicek model and the Quadratic interest rate model using crude Monte Carlo simulation for the following number of steps: k=[8,16,32,64,128]. We have performed the calculation for the two different weak discretizations scheme for the SDE which was shown in section 8.

We have performed B batches of N simulations, for B = 100 and N = 1000, and have used the constrained MC-technique presented in section 8.4 in order to ensure that the initial yield-curve is matched for all k for each batch. Furthermore we have - as when using the lattice-based approach - used the knowledge that we can obtain a (semi) closed form solution for the bond-price.

The calculated put-option prices for the Extended Vasicek model are shown in table 6 and for the Quadratic interest rate model in table 7. The calculated call-option prices for the two models are shown in Appendix I.

Table 6: Put-option prices in the Extended Vasicek Model using Constrained Crude Monte Carlo Simulation

	K=8	Std. Error	K=16	Std. Error	K=32	Std. Error	K=64	Std. Error	K=128	Std. Error
Euler Discretization scheme - order 1 weak Taylor approximation										
52,53	*0.000*	*0.001*	*0.000*	*0.000*	*0.000*	*0.001*	*0.000*	*0.001*	*0.000*	*0.001*
62,53	*0.278*	*0.019*	*0.280*	*0.025*	*0.285*	*0.023*	*0.276*	*0.023*	*0.275*	*0.021*
67,53	*1.552*	*0.033*	*1.609*	*0.043*	*1.627*	*0.040*	*1.640*	*0.036*	*1.637*	*0.036*
72,53	*4.466*	*0.025*	*4.569*	*0.031*	*4.603*	*0.020*	*4.630*	*0.021*	*4.632*	*0.020*
82,53	*12.639*	*0.008*	*12.765*	*0.006*	*12.825*	*0.005*	*12.854*	*0.006*	*12.866*	*0.008*
Milstein Discretization scheme - order 2 weak Taylor approximation										
52,53	*0.000*	*0.001*	*0.000*	*0.001*	*0.000*	*0.001*	*0.000*	*0.000*	*0.000*	*0.000*
62,53	*0.265*	*0.017*	*0.271*	*0.026*	*0.283*	*0.022*	*0.284*	*0.016*	*0.278*	*0.021*
67,53	*1.572*	*0.041*	*1.603*	*0.042*	*1.624*	*0.039*	*1.633*	*0.041*	*1.647*	*0.033*
72,53	*4.465*	*0.031*	*4.553*	*0.027*	*4.601*	*0.030*	*4.628*	*0.028*	*4.635*	*0.025*
82,53	*12.642*	*0.008*	*12.761*	*0.005*	*12.825*	*0.008*	*12.852*	*0.007*	*12.865*	*0.006*

Table 7: Put-option prices in the Quadratic Interest Rate Model using Constrained Crude Monte Carlo Simulation

	K=8	Std. Error	K=16	Std. Error	K=32	Std. Error	K=64	Std. Error	K=128	Std. Error
Euler Discretization scheme - order 1 weak Taylor approximation										
52,53	*0,000*	*0,000*	*0,000*	*0,000*	*0,000*	*0,000*	*0,000*	*0,000*	*0,000*	*0,000*
62,53	*0,017*	*0,006*	*0,018*	*0,006*	*0,018*	*0,005*	*0,021*	*0,005*	*0,018*	*0,004*
67,53	*0,763*	*0,024*	*0,804*	*0,016*	*0,819*	*0,020*	*0,835*	*0,022*	*0,844*	*0,022*
72,53	*4,089*	*0,005*	*4,199*	*0,004*	*4,255*	*0,004*	*4,280*	*0,005*	*4,292*	*0,004*
82,53	*12,650*	*0,002*	*12,766*	*0,002*	*12,822*	*0,002*	*12,849*	*0,002*	*12,864*	*0,002*
Milstein Discretization scheme - order 2 weak Taylor approximation										
52,53	*0,000*	*0,000*	*0,000*	*0,000*	*0,000*	*0,000*	*0,000*	*0,000*	*0,000*	*0,000*
62,53	*0,017*	*0,004*	*0,018*	*0,005*	*0,018*	*0,006*	*0,019*	*0,005*	*0,020*	*0,004*
67,53	*0,767*	*0,016*	*0,795*	*0,020*	*0,825*	*0,019*	*0,840*	*0,019*	*0,842*	*0,018*
72,53	*4,088*	*0,004*	*4,201*	*0,004*	*4,254*	*0,003*	*4,279*	*0,005*	*4,294*	*0,004*
82,53	*12,650*	*0,002*	*12,767*	*0,002*	*12,822*	*0,002*	*12,850*	*0,002*	*12,864*	*0,002*

First of all it is easy to see that the rate of convergence is much slower than in the lattice-based method.

From the put-option prices shown in tables 6 and 7 compared to the call-option prices in Appendix I, we have the following general indication:

> The convergence for put-option prices is a process which starts out with a price lower than the true price and then rises towards the true price as a function of the increasing number of steps.
> The convergence for call-option prices is a process which starts out with a price higher than the true price and then falls towards the true price as a function of the increasing number of steps.

The possible explanation for this phenomenon is that there is a tendency for the simulation procedure[48] to calculate/estimate the highest bond-price for the lowest number of steps - that is the convergence starts out with a higher bond-price and then fall towards the true price as a function of the increasing number of steps. This explains the indications mentioned above.

---

[48] This is actually also the case for the simulations I have performed for time-homogenous SDE spot-rate models.

Looking at the results for the two different discretizations schemes it is difficult to conclude anything in general.

In the case of simulation of spot-rate models where the spot-rate process is governed by a time-homogenous SDE, we usually have a much clearer picture - namely that using higher order schemes in the sampling process in general matters[49]. From the results we obtain for time-dependent SDE's of the particular kind used here - no clear picture is available.

Which respect to time-dependent SDE's (of the kind analysed here) we cannot in general see a more efficient convergence rate when using higher order schemes. This must be because of the procedure that forces the average price for each k = [1,K] to match the price P(t,kδ) observed in the market. Remember, namely that this is done by adjusting each element in the basic diffusion-matrix column-wise - that is as a change of drift. This procedure to match the initial yield-curve must therefore dominate over the sampling scheme.

It could be interesting to look at a couple of variance reduction schemes to see if they make the convergence rate faster. With the above results in mind it is however not clear beforehand if this will increase the efficiency.

For this purpose we will consider two different kinds of variance reduction methods[50]:

Variance reduction by using the Brownian Bridge Process
Variance reduction by Girsanovs transformation

Both these methods are very efficient, in the case of simulating spot-rate processes under the assumption of a time-homogenous SDE[51].

Because of the unclear conclusion in connection with the use of different discretizations schemes for the SDE, we have only performed the calculations using the Euler discretizations method.

When using the Brownian Bridge method, we encountered some "strange" results. The results are shown in Appendix J. It is however worth pointing out that we have not been able to find any reason why the simulated options-prices in some cases seem to diverge instead of converge.

---

[49] These results can be obtained by contacting the author. The time-homogenous spot-interest rate model we have tested is the Ornstein-Uhlenbeck process of the kind specified in equation 85. Similar results has been obtained by Giles (2007) - for the valuation of a range of options in the Black-Scholes model.

[50] These two variance reduction techniques is explained in Appendix G.

[51] Results indication that these variance reduction methods is efficient for spot rate models that are time-homogenous can be obtained from the author. It is however in this connection worth mentioning that by far the most efficient method of those employed was the Girsanov's transformation.

Our results using the method on a time-homogenous spot-rate model behaved as expected. It was clearly superior to Crude Monte Carlo simulation using the Euler scheme and slightly better than Crude Monte Carlo simulation using the 2. order weak Milstein.

The only difference between our simulations with the time-homogenous spot-rate model and the result in Appendix J, is that we in Appendix J have constrained the Monte Carlo simulation in order to match the current yield-curve. Why that should make the Brownian Bridge method diverge in some cases is not obvious, and is left for further research.

In tables 8 and 9 we have shown the result for both the Extended Vasicek model and the Quadratic interest rate model - when using Girsanovs transformation in the simulation procedure.

Table 8: Call-and Put-option prices in the Extended Vasicek Model using Constrained Girsanovs Transformation Monte Carlo Simulation

	K=8	Std. Error	K=16	Std. Error	K=32	Std. Error	K=64	Std. Error	K=128	Std. Error
Call-Option prices										
52,53	*13,120*	*0,009*	*12,997*	*0,007*	*12,939*	*0,008*	*12,907*	*0,008*	*12,894*	*0,009*
62,53	*4,802*	*0,028*	*4,692*	*0,029*	*4,626*	*0,024*	*4,601*	*0,022*	*4,592*	*0,033*
67,53	*1,800*	*0,035*	*1,715*	*0,037*	*1,674*	*0,036*	*1,669*	*0,055*	*1,657*	*0,043*
72,53	*0,423*	*0,033*	*0,373*	*0,022*	*0,368*	*0,033*	*0,365*	*0,035*	*0,359*	*0,034*
82,53	*0,006*	*0,004*	*0,004*	*0,005*	*0,003*	*0,003*	*0,003*	*0,003*	*0,002*	*0,003*
Put-Option prices										
52,53	*0,000*	*0,000*	*0,000*	*0,000*	*0,000*	*0,001*	*0,000*	*0,000*	*0,000*	*0,001*
62,53	*0,271*	*0,025*	*0,272*	*0,021*	*0,284*	*0,024*	*0,277*	*0,023*	*0,280*	*0,024*
67,53	*1,553*	*0,032*	*1,601*	*0,033*	*1,621*	*0,028*	*1,649*	*0,045*	*1,632*	*0,033*
72,53	*4,464*	*0,025*	*4,567*	*0,031*	*4,602*	*0,026*	*4,623*	*0,029*	*4,639*	*0,023*
82,53	*12,639*	*0,008*	*12,761*	*0,005*	*12,822*	*0,007*	*12,850*	*0,008*	*12,865*	*0,008*

Table 9: Call-and Put-option prices in the Quadratic Interest Rate Model using Constrained Girsanovs Transformation Monte Carlo Simulation

	K=8	Std. Error	K=16	Std. Error	K=32	Std. Error	K=64	Std. Error	K=128	Std. Error
Call-Option prices										

52,53	*13,104*	*0,002*	*12,989*	*0,003*	*12,933*	*0,002*	*12,905*	*0,002*	*12,891*	*0,002*
62,53	*4,538*	*0,006*	*4,423*	*0,006*	*4,364*	*0,006*	*4,337*	*0,006*	*4,324*	*0,006*
67,53	*0,996*	*0,020*	*0,913*	*0,016*	*0,885*	*0,018*	*0,866*	*0,020*	*0,854*	*0,021*
72,53	*0,023*	*0,005*	*0,018*	*0,005*	*0,016*	*0,006*	*0,016*	*0,004*	*0,014*	*0,004*
82,53	*0,000*	*0,000*	*0,000*	*0,000*	*0,000*	*0,000*	*0,000*	*0,000*	*0,000*	*0,000*
Put-Option prices										
52,53	*0,000*	*0,000*	*0,000*	*0,000*	*0,000*	*0,000*	*0,000*	*0,000*	*0,000*	*0,000*
62,53	*0,017*	*0,004*	*0,019*	*0,005*	*0,019*	*0,005*	*0,020*	*0,005*	*0,018*	*0,005*
67,53	*0,764*	*0,016*	*0,804*	*0,019*	*0,827*	*0,020*	*0,836*	*0,018*	*0,845*	*0,017*
72,53	*4,089*	*0,004*	*4,201*	*0,004*	*4,254*	*0,004*	*4,281*	*0,004*	*4,293*	*0,004*
82,53	*12,650*	*0,002*	*12,767*	*0,002*	*12,822*	*0,002*	*12,851*	*0,002*	*12,864*	*0,002*

If the results in tables 8 and 9 are compared with respectively tables 6 and 7 and the results in Appendix I we can deduce the following:

No clear variance reduction is obtained using Girsanovs transformation, ie no substantial fall in Std. error

The convergence to the true option-price is not remarkably faster using Girsanonv’s transformation

Employing Monte Carlo simulation for spot-rate models with a time-dependent drift that through the constraining method ensures a perfect match of the yield-curve for all steps does not seem as straightforward as in the case of time-homogenous spot-rate models. With straightforward - I mean, variance reduction is not easily obtained - at least not through conventional techniques.

Analysing the pricing of options on zero-coupon bonds using Monte Carlo and the Hull and White lattice-based approach, we can conclude the following:

It seems as if the constraining method dominates over the variance reduction schemes (at least the ones employed here) and that it also dominates over the discretization schemes - that is a Crude Monte Carlo method using Eulers discretizations scheme does the work as well as anything else we came up with

The Hull and White lattice-based approach is superior to the Monte Carlo simulation

Of course it is possible that by employing other tricks in connection with Monte Carlo simulation we might get better results - maybe even results that can rival the lattice-based methods. Quasi-random number sequences is a next natural step to take, or other variance reduction techniques - such as stratified sampling - but that is left for further research.

The whole idea of introducing Monte Carlo methods for time-dependent spot-rate models was to be able to price path-dependent securities in a more natural setting, as lattice-based method is not directly suitable for this kind of problem. From the results in this paper on a fairly simple instrument, we however encountered some problems when we tried to reduce the variance in the Monte Carlo simulation procedure. Because variance reduction in connection with Monte Carlo method is extremely important, there is a need for finding techniques which are effective in connection with time-dependent spot-rate models. The methods we tried here - which usually are very efficient - did not perform as promised/expected.

The only direct difference between the Hull and White lattice-based method and the Constrained Monte Carlo method introduced here - is how each of the procedures constructs the time-space dimension for the basic interest rate process x. From this it seems clear that the controlled method that is inherent in the lattice-based methodology gives a superior representation of the process than what is possible with Monte Carlo. The only logical explanation for this phenomonen must be that the spot-rate process is Markovian.

From this we conclude:

> Markovian spot-rate models are best represented by a method that takes into account the Markovian nature in the process - that is lattice-based methods Monte Carlo are not as efficient as lattice-based procedures for time-dependent Markovian spot-rate models, as Monte Carlo methods do not take into account the Markovian nature - Monte Carlo is doing precisely the opposite - as Monte Carlo methods by nature are non-Markovian.

I do not have any knowledge of any research that compares Monte Carlo pricing methods to lattice-based pricing methods for time-homogenous Markovian spot-rate models - but relying on the conclusion above it would not be a surprise if the same conclusions were reached[52].

In general it is worth pointing out that the pricing of path-dependent contingent claims does not rely on the assumption that the spot-rate process is non-Markovian - that is, the need for a non-Markovian pricing technique does not neccesarily indicate that the spot-rate process needs to be non-Markovian.

This indicates that when pricing path-dependent instruments we have to decide on the following two things:

1: How do we most efficiently construct the time-space dimension of the interest rate process?

---

[52] This could be an interesting line of research.

2: Given that, how do we most efficiently utilize the constructed time-space dimension to price a path-dependent contingent claim?

For the pricing of path-dependent contingent claims there is a tradition for using Monte Carlo, because Monte Carlo methods by nature - like the problem at hand - is non-Markovian.

Doing that we do not distinguish between the two points mentioned above - but treat them as if they were one and the same. This is however not the case, as explained above. This does not neccesarily mean that Monte Carlo is not the most efficient method, it just indicates that we perhaps should re-think the problem - maybe there is a more efficient solution?.

10: Conclusion

Two of the main results in this paper were: First we showed how to determine the T-forward adjusted risk-measure using the concept of fundamental solution to linear PDEs. That is, it turned out to be possible to derive the T-forward adjusted risk-measure without the use of Girsanovs theorem, which is the traditional approach, see for example Karoui, Myneni and Viswanathan (1993).

The other main result was that we were able to carry the analysis of the Quadratic interest rate model further than Jamshidian (1996) by relying on the Fourier transformation in order to obtain the fundamental solution for the PDE.

In that connection we showed how it was possible to derive the price of a discount bond and the price of an option on a discount bond. This was also as an example done for the extended Vasicek model. Using the idea from Hull and White (1990), we showed how it was possible to fit the model to the initial term structure.

The next two parts of the paper focused on implemention issues.

In that connection, we designed a special discrete time model for the Quadratic interest rate model, as it in some cases was not possible to use the trinomial approach from Hull and White (1994).

After that we focused on pricing techniques for path-dependent interest rate contingent claims. We focused in that connection on Monte Carlo simulation of spot-rate models with a time-dependent drift. The third main result in the paper was to introduce a forward-induction technique that made it possible to constrain the Monte Carlo simulation for the matching of the initial yield-curve.

When using Monte Carlo simulation for time-dependent spot-rate models, we encountered problems obtaining variance reduction - by methods that in the case of time-homogenous spot-rate models are very effective. Further research is obviously needed in connection with Monte Carlo simulation of time-dependent spot-rate models.

Finally we compared lattice-based pricing method to Monte Carlo simulation procedures for the pricing of European put-and call-options on zero-coupon bonds for both the Extended Vasicek model and the Quadratic interest rate model.

In that connection we concluded:

> Markovian spot-rate models are best represented by a method that takes into account the Markovian nature in the process - that is lattice-based methods Monte Carlo are not as efficient as lattice-based procedures for time-dependent Markovian spot-rate models, as Monte Carlo methods do not take into account the Markovian nature - Monte Carlo is doing precisely the opposite - as Monte Carlo methods by nature are non-Markovian.

With respect to the pricing of path-dependent claims, no clear conclusion was drawn. It was however suggested that before deciding on a method (Monte Carlo or lattice-based method), it was of importance to separate the problem, as follows:

> How do we most efficiently construct the time-space dimension of the interest rate process?.
> Given that, how do we most efficiently utilize the constructed time-space dimension to price a path-dependent contingent claim?

From this we deduce that, even though Monte Carlo is the natural method to use when pricing path-dependent interest rate contingent claims, it might not be the most efficient one - at least not when the spot-rate is Markovian.

References:

Acworth, Broadie and Glasserman (1997) “A Comparison of some Monte Carlo and Quasi Monte Carlo Techniques for Option Pricing”, in Niederreiter et al. (Eds.), Monte Carlo and Quasi-Monte Carlo Methods 1996 - Springer-Verlag New York, Lectures Notes in Statistics, 1998, pp.1-18

Ait-Sahalia (1995a) “Testing Continuous-Time Models of the Spot Interest Rate”, Review of Financial Studies, 1996, 9, 385-426 (this paper received the Michael J. Brennan Award for the best paper published in the Review of Financial Studies in 1996)

Ait-Sahalia (1995b) “Nonparametric Pricing of Interest Rate Derivative Securities”, Econometrica, 1996, 64, 527-560

Andersen (1995) “Efficient Techniques for Simulation of Interest Rate Models involving Non-Linear Stochastic Differential Equations”, working paper Generel-Re-Financial Products

Andersen and Brotherton-Ratcliffe (1998) “The equity option volatility smile: an implicit finite-difference approach”, The Journal of Computational Finance, vol. 1, no. 2, page 5-37

Andersen and Piterbarg (2010) “Interest Rate Modeling”, Volume II, Atlantic Financial Press

Assefa (2007) “Calibration and pricing in a multi-factor Quadratic Gaussian Model”, Research Paper 197, UTS, ISSN 1441-8010

Barraquand (1994) “Numerical Valuation of High Dimensional Multivariate European Securities”, Management Science Vol. 41, No. 12, Dec., 1995

Beaglehole and Tenney (1991) “General Solution of some Interest Rate-Contingent Claim Pricing Equations”, Journal of Fixed Income 1, page 69-83

Beaglehole and Tenney (1992) “A Nonlinear Equilibrium Model of the Term Structure of Interest Rates: Corrections and Additions”, Journal of Financial Economics 32 (3), page 345-354

Black, Derman and Toy (1990) "A One-Factor Model of Interest Rates and its Application to Treasury Bond Options", Financial Analyst Journal, January-February 1990, page 33-39

Black and Karasinski (1991) "Bond and Option pricing when short rates are lognormal", Financial Analyst Journal, July-August 1991, page 52-69

Black and Scholes (1973) “The Pricing of Options and Corporate Liabilities”, Journal of Political Economy 81, page 637-654

Bloch and Assefa (2009) “Fast Calibration of interest rate claims in the quadratic gaussian model: 1 The Caplets”, working paper Mizuho Securities

Bloch and Assefa (2009) “Fast Calibration of interest rate claims in the quadratic gaussian model: 2 The Swaptions”, working paper Mizuho Securities

Bollerslev, Engle and Nelson (1993) “ARCH-Models”, working paper no. 154 - prepared for The Handbook of Econometrics, vol. 4

Boyle, Broadie and Glasserman (1995) “Monte Carlo Methods for Security Pricing”, Journal of Economic Dynamics and Control Volume 21, Issues 8-9, 29 June 1997, Pages 1267-1321

Brace, Gatarek and Musiela (1995) "The Market Model of Interest Rate Dynamics", Mathematical Finance, Vol. 7, No. 2 (April 1997), 127–147

Caflisch, Morokoff and Owen (1997) “Valuation of mortgage-backed securities using Brownian bridges to reduce effective dimensions”, The Journal of Computational Finance, vol. 1, no. 1, page 27-46

Campbell, Lo and Mackinlay (1997) “The Econometric of Financial Markets ”, Princeton University Press

Chen and Vincent Poor (2006) “Markovian Quadratic Term Structure Models for Risk-Free and Defaultable rates”, Working paper

Constantinides (1992) “A Theory of the Nominal Term Structure of Interest Rates”, Review of Financial Studies 5 (4), page 531-552

Cox and Miller (1965) “The Theory of Stochastic Processes”, Chapman and Hall. London

Duan and Simonato (1995) “Empirical Martingale Simulation for Asset Prices”, Management Science Vol. 44, No. 9, September 1998

Dupire (1994) “Pricing with a smile”, RISK 7, no. 1, page 18-20

Duffie (1996) "Dynamic Asses Pricing Theory", second edition, Princeton University Press

Flesaker and Houghston (1996) "Positive Interest: Foreign Exchange", working paper Merril Lynch, 10 of January 1996.

Gaspar (2004) “General Quadratic Term Structures of Bond, Futures and Forward Prices”, SSE/EFI Working paper series in Economics and Finance No 550, Stockholm School of Economics

Giles (2007) “Improved multilevel Monte Carlo convergence using the Milstein scheme”, page 343-358, in Monte Carlo and Quasi-Monte Carlo Methods 2006, Springer, 2007

Gut (1995) "An intermediate course in Probability", Springer-Verlag New York

Heath, Jarrow and Morton (1991) "Bond Pricing and the Term Structure of Interest Rates: A new Methodology for Contingent Claim Valuation", Econometrica, Vol. 60, No. 1, (January 1992), pp 77-105.

Heitmann and Trautmann (1995) "Gaussian Multi-factor Interest Rate Models: Theory, Estimation, and Implications for Option Pricing", working paper Johannes Gutenberg-University Mainz June 1995, http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.40.7649

Hull and White (1990) "Pricing Interest-Rate-Derivative Securities", The Review of Financial Studies 3, page 573-592

Hull and White (1993a) "One-Factor Interest Rate Model and the Valuation of Interest Rate Derivative Securities", Journal of Financial and Quantitative Analysis 28, 1993, page 235-254

Hull and White (1993b) "Efficient Procedures for Valuing European and American Path-Dependent Options", Journal of Derivatives vol. 1, no.1 1993, page 21-31

Hull and White (1993c) “Bond Option pricing based on a model for the evolution of bond prices”, Advanced in Futures and Options Research, Vol. 6, page 1-13

Hull and White (1994) "Numerical Procedures for Implementing Term Structure Models", The Journal of Derivatives Fall 1994, Vol. 2, No. 1: pp. 7-16

Jackwerth and Rubenstein (1996) “Recovering Probability Distributions from Option Prices”, The Journal of Finance, vol. 51, no. 5, page 1611-1631

Jamshidian (1989) "An exact Bond Option Formula", Journal of Finance 44, page 205-209

Jamshidian (1995) "A Simple Class of Square-Root Interest Rate Models", Applied Mathematical Finance, vol. 2, no. 1, March 1995, page 61-72

Jamshidian (1996) "Bond, future and option evaluation in the quadratic interest rate model", Applied Mathematical Finance, vol. 3, no. 2, June 1996, page 93-116.

Jiang and Yan (2009) “Linear-Quadratic term structure models - toward the understanding of jumps in interest rates”, Journal of Banking and Finance 33, page 473-485

Karatzas and Schreve (1988) “Brownian Motion and Stochastic Calculus“, Springer-Verlag NewYork Inc.

Karoui, Myneni and Viswanathan (1993) "Arbitrage pricing and hedging of interest rate claims with state variables", Working Paper, University of Paris IV

Kijima and Nagayama (1995) "Efficient Numerical Procedures for the Hull-White Extended Vasicek Model", The Journal of Financial Engineering, vol. 3, no. 3/4, 1995, page 275-292

Kloeden and Platen (1995) “Numerical Solution of Stochastic Differential Equations”, Springer-Verlag Berlin Heidelberg, second corrected edition

Laurance and Wang (2005) “Closed form solutions for Quadratic and Inverse Quadratic Term Structure Models”, working paper

Leblon and Moraaux (2009) “Quadratic Term Structure Models: Analysis and Performance”, Working paper Université de Rennes

Leippold and Wu (1998) “Quadratic Term Structure Models”, Working paper, university of St. Gallen and University of New York

Leippold and Wu (2001) “Design and Estimation of Quadratic Term Structure Models”, CRIF Working paper series, 6-14-2001

Leippold and Wu (2002) “Asset Pricing under the Quadratic Class”, Journal of Financial and Quantitative Analysis, Vol. 37, No. 2, June 2002.

Longstaff (1989) “A Nonlinear General Equilibrium Model of the Term Structure of Interest Rates”, Journal of Financial Economics 23, page 195-224

McDonnell (1998) “At Play with J: A Newer Random Link Generator”, Vector vol. 14, no. 4, April 1998

Matsumoto, M.; Nishimura, T. (1998). "Mersenne twister: a 623-dimensionally equidistributed uniform pseudo-random number generator". ACM Transactions on Modeling and Computer Simulation 8 (1): 3–30

Miltersen, Sandmann and Sondermann (1995) "Closed Form Solutions for Term Structure Derivatives with Log-Normal Interest Rates", The Journal of Finance, Vol. 52, No. 1, March 1997, pp. 409-430

Musiela and Rutkowski (1997) “Martingale Methods in Financial Modelling”, Springer-Verlag Berlin Heidelberg

Niederrieter (1992) “Random Number Generation and Quasi-Monte Carlo Method”, CBMS-NSF Regional Conference Series in Applied Mathematics 63

Paskov (1996) "New Methodologies for Valuing Derivatives", http://citeseer.ist.psu.edu/viewdoc/summary?doi=10.1.1.30.3094

Pelsser (1995) "An analytically tracable interest rate model that precludes negative interest rates", Springer Verlag, 2000

Pipkin (1991) "A Course on Integral Equations", Texts in Applied Mathematics 9, Springer-Verlag New York

Press, Teukolovski, Vetterling and Flannery (1992) “Numerical Recipes in C”, Cambridge University Press, second edition

Realdon (2006) “Quadratic Term Structure Models in Discrete Time”, working paper Department of Economics, University of York

Rosenberg (1997) "Pricing Multivariate Contingent Claims using Estimated Risk-Neutral Density Functions", http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1297773

Schoenmakers and Heemink (1997) “Fast valuation of financial derivatives”, The Journal of Computational Finance, vol. 1, no 1, page 47-62

Scott (1992) “Multivariate Density Estimation”, John Wiley and Sons, Inc.

Wilmott, Dewynne and Howison (1993) "Option Pricing - Mathematical models and computation", Oxford University Press

Zhenyu (1994) “Monte Carlo Simulation of the HJM Interest Rate Model”, working paper Cornell University

Appendix A

We seek a solution A(t,T,ς) and B(t,T,ς) that simultanously satisfies the following set of ordinary differential equation :

$$\begin{aligned} B_t - Ax(t) + \frac{1}{2}\sigma^2(t)A^2 - \psi(t) &= 0 \\ A_t - A\kappa(t) - 1 &= 0 \end{aligned} \qquad \textbf{(90)}$$

with respect to the boundary conditions A(T,T,ς) = iς and B(T,T,ς) = 0.

We will start with the second equation in formula 1 as we can here find a unique solution for A(t,T,ς). It is easily seen that A(t,T,ς) is of the following form:

$$\begin{gathered} A(t,T,\varsigma) = -\int_t^T e^{k(s)}ds + k \\ where \\ k(t) = \int_t^T \kappa(s)ds \end{gathered} \qquad \textbf{(91)}$$

where k is the constant which can be found by using the boundary condition, hence we get:

$$\begin{gathered} A(t,T,\varsigma) = -\int_t^T D(s,T)ds + i\varsigma D(t,T) \\ where \\ D(t,T) = e^{-\int_t^T \kappa(s)ds} \end{gathered} \qquad \textbf{(92)}$$

By plucking formula 3 into the first equation in formula 1 and rearranging terms we get:

$$\begin{aligned} B(t,T,\varsigma) &= i\varsigma\int_t^T D(s,T)x(s)ds - \int_t^T A(s,T)x(s)ds \\ &- \frac{1}{2}\int_t^T \sigma^2(s)[A(s,T)^2 - \varsigma^2 D(s,T)^2 - 2i\varsigma A(s,T)D(s,T)]ds + \int_t^T \psi(s)ds + k \end{aligned} \qquad \textbf{(93)}$$

Where k is the constant. Applying the boundary condition it can be seen that B(t,T,ς) is given by (as k = 0):

$$\begin{aligned} B(t,T,\varsigma) &= \int_t^T \left[\frac{1}{2}\sigma^2(s)A(s,T)^2 + A(s,T)x(s) - \psi(sa) \right] ds \\ &- \int_t^T \left[i\varsigma D(s,T)[x(s) + \sigma^2(s)A(s,T)] + \frac{1}{2}\varsigma^2\sigma^2(s)D(s,T)^2 \right] ds \end{aligned} \qquad \textbf{(94)}$$

Formula 23 in the main text now follows directly from equation 3 and 5.

Appendix B

We seek a solution A(t,T,ς), B(t,T,ς) and C(t,T,ς) that simultanously satisfies the following set of ordinary differential equation :

$$\begin{aligned} A_t + \frac{1}{2}\sigma^2 B^2 - \sigma^2 C - \psi(t)^2 &= 0 \\ B_t - \kappa B - 2\sigma^2 BC + 2\psi(t) &= 0 \\ C_t - 2\kappa C - 2\sigma^2 C^2 + 1 &= 0 \end{aligned} \tag{95}$$

with respect to the boundary conditions B(T,T,ς) = -iς and A(T,T,ς) = C(T,T,ς) = 0.

We will start with equation no. 3 in formula 1 as this is only a function of C(t,T,ς). C(t,T,ς) is known to be the solution to the following integral:

$$2(T - t) = \int_t^T \frac{1}{-\sigma^2 C(s,T,\varsigma)^2 - \kappa C(s,T,\varsigma) + \frac{1}{2}} dC \tag{96}$$

which by employing a few tricks yields:

$$C(t,T,\varsigma) = C(t,T) = \frac{e^{2\gamma(T - t)} - 1}{(\gamma + \kappa)e^{2\gamma(T - t)} + (\gamma - \kappa)} \tag{97}$$

We will now turn our attention to the second equation in formula 1 as the only unknown here is B(t,T,ς), more precisely we are seeking a solution to the following first order ODE:

$$B_t = \kappa B + 2\sigma^2 BC(t,T) - 2\psi(t) \tag{98}$$

where it follows that B is given by:

$$\begin{aligned} B(t,T,\varsigma) &= -e^{-G(t,T)} \int_t^T 2\psi(s) e^{G(s,T)} ds + k e^{-G(t,T)} \\ &\quad where \\ G(t,T) &= \int_t^T [2\sigma^2 C(s,T) + \kappa] ds \end{aligned} \tag{99}$$

The only problem here is the integral $\int_t^T C(s,T)ds$. In order to faciliate a solution let me first rewrite C(t,T) as:

$$\begin{aligned} C(t,T) &= \frac{\gamma - \kappa}{2\sigma^2}\left(\frac{(\kappa + \gamma)(e^{2\gamma(T - t)} - 1)}{h(t,T)}\right) \\ &= \frac{1}{2\sigma^2}\frac{h'(t,T)}{h(t,T)} - \frac{\kappa + \gamma}{2\sigma^2} \\ &\quad where \\ h(t,T) &= (\kappa + \gamma)e^{2\gamma(T - t)} + (\gamma - \kappa) \\ h'(t,T) &= 2\gamma(\kappa + \gamma)e^{2\gamma(T - t)} \end{aligned} \qquad \textbf{(100)}$$

We can now easily derive $\int_t^T C(s,T)ds$ using the substitutions-rule which yields:

$$\int_t^T C(s,T)ds = \frac{1}{2\sigma^2}\left[\ln\left(\frac{2\gamma}{h(t,T)}\right) \mp (\kappa + \gamma)(T - t)\right] \qquad \textbf{(101)}$$

By plucking formula 7 into formula 5 and doing a little algebra, we find that B(t,T,ς) can be written as:

$$B(t,T,\varsigma) = -2\int_t^T \frac{e^{\gamma s}h(s,T)}{e^{\gamma t}h(t,T)}\psi(s)ds + k\frac{2\gamma e^{\gamma(T - t)}}{h(t,T)} \qquad \textbf{(102)}$$

The constant k can be found by employing the boundary condition B(T,T,ς) = -iς, hence we find the following expression for B(t,T,ς):

$$\begin{aligned} B(t,T,\varsigma) &= B(t,T) - i\varsigma D(t,T) \\ &\quad where \\ B(t,T) &= 2\int_t^T \frac{e^{\gamma s}h(s,T)}{e^{\gamma t}h(t,T)}\psi(s)ds \\ D(t,T) &= \frac{2\gamma e^{\gamma(T - t)}}{h(t,T)} \end{aligned} \qquad \textbf{(103)}$$

Now we only have to determine A(t,T,ς) to have a solution to the system of ODE´s from formula 1, and it turns out that A(t,T,ς) with respect to the boundary condition can be written as:

$$A(t,T,\varsigma) = \int_t^T \left[\frac{1}{2}\sigma^2 B(s,T)^2 - \sigma^2 C(s,T) - \psi(s)\right] ds - \frac{1}{2}\sigma \int_t^T [2i\varsigma D(s,T)B(s,T) + \varsigma^2 D(s,T)^2] ds \quad \textbf{(104)}$$

Formula 39 and 40 in the main text now follows directly from equation 3,9 and 10.

Appendix C

From formula 45 in the main text, we have that the price at time t of a call-option with expiry at time T^F written on a discount bond that matures at time T, for t < T^F < T, in the Quadratic interest rate model is given by:

$$C(t,T^F) = P(t,T^F)\int_{-\infty}^{\infty}\left[[e^{A(T^F,T) - B(T^F,T)z - C(T^F,T)z^2} - K]^+ \frac{1}{\sqrt{2\pi V(t,T^F)}} e^{-\frac{1}{2}\frac{[z - M(t,T^F)]^2}{V(t,T^F)}}\right]dz \quad \textbf{(105)}$$

This can be rewritten as:

$$C(t,T^F) = P(t,T^F)\int_{-\infty}^{\infty} e^{A(T^F,T) - B(T^F,T)z - C(T^F,T)z^2} p(q,t,T^F,z)dz - P(t,T^F)K\int_{-\infty}^{\infty} p(q,t,T^F,z)dz$$

$$\text{where} \qquad \textbf{(106)}$$

$$p(q,t,T^F,z) = \frac{1}{\sqrt{2\pi V(t,T^F)}} e^{-\frac{1}{2}\frac{[z - M(t,T^F)]^2}{V(t,T^F)}}$$

If we integrate formula 2 for all z for which the payoff is positive, we will be able to express the price of the call-option in terms of the cumulative normal distribution. From the main text we know that z has to be found as the root in a second order equation, more precisely we find that z is positive for values of z lying in the following interval:

$$Lower = \frac{-B(T^F,T) - \sqrt{Dis}}{2C(T^F,T)} < z < \frac{-B(T^F,T) + \sqrt{Dis}}{2C(T^F,T)} = Upper$$

$$\text{where} \qquad \textbf{(107)}$$

$$Dis = B(T^F,T)^2 + 4C(T^F,T)[A(T^F,T) - \ln K]$$

From this relation it it clearly seen that for Dis ≤ 0, the payoff H(z) will never be positive - that is the value of the option has to be equal to zero (0). If instead Dis > 0 we can integrate formula 2 over the region Lower < z < Upper.

The last integral in formula 2 is easy found as this can be expressed as:

$$\int_{Lower}^{Upper} p(q,t,T^F,z)ds = N\left(\frac{Upper - M(t,T^F)}{\sqrt{V(t,T^F)}}\right) - N\left(\frac{Lower - M(t,T^F)}{\sqrt{V(t,T^F)}}\right) \quad \textbf{(108)}$$

From the first integral in formula 2 it follows that z has to multiplied by $1 + 2C(T^F,T)V(t,T^F()$. Which means that we need to multiply the variance and the b with $F = 1 + 2C(T^F,T)V(t,T^F)$, hence the first integral can be rewritten as[53]:

$$\int_{-\infty}^{\infty} e^{A(T^F,T) - B(T^F,T)z - C(T^F,T)z^2} p(q,t,T^F,z)dz$$

$$= \left[\sqrt{F}e^{A(T^F,T) + \frac{\frac{1}{2}B(T^F,T)^2V(t,T^F) - M(t,T^F)^2C(T^F,T) - M(t,T^F)B(T^F,T)}{F}} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi V(t,T^F)F}} e^{-\frac{1}{2}\frac{[zF - M(t,T^F) + B(T^F,T)V(t,T^F)]^2}{V(t,T^F)F}} dz\right] \tag{109}$$

$$= \left[\sqrt{F}e^{A(T^F,T) + \frac{\frac{1}{2}B(T^F,T)^2V(t,T^F) - B(T^F,T)M(t,T^F) - C(T^F,T)M(t,T^F)^2}{F}} \int_{\frac{LowerF - M(t,T^F) + B(T^F,T)V(t,T^F)}{\sqrt{FV(t,T^F)}}}^{\frac{UpperF - M(t,T^F) + B(T^F,T)V(t,T^F)}{\sqrt{FV(t,T^F)}}} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz\right]$$

As $P(T^F,T) = \sqrt{F}e^{A(T^F,T) + \frac{\frac{1}{2}B(T^F,T)^2V(t,T^F) - B(T^F,T)M(t,T^F) - C(T^F,T)M(t,T^F)^2}{F}}$

expression:

$$P(T^F,T)\left[N\left(\frac{UpperF - E}{\sqrt{FV(t,T^F)}}\right) - N\left(\frac{LowerF - E}{\sqrt{FV(t,T^F)}}\right)\right] \tag{110}$$

All in all this means that the call-options price is given by:

$$\begin{gathered}
C(t,T^F) = P(t,T)\left[N\left(\frac{UpperF - E}{\sqrt{V(t,T^F)F}}\right) - N\left(\frac{LowerF - E}{\sqrt{V(t,T^F)F}}\right)\right] \\
- P(t,T^F)K\left[N\left(\frac{Upper - M(t,T^F)}{\sqrt{V(t,T^F)}}\right) - N\left(\frac{Lower - M(t,T^F)}{\sqrt{V(t,T^F)}}\right)\right] \\
where \\
E = M(t,T^F) - B(T^F,T)V(t,T^F) \\
F = 1 + 2C(T^F,T)V(t,T^F)
\end{gathered} \tag{111}$$

Using a similar technique it is possible to derive the price of a put-option, hence we get:

---

[53] Se for example Gut (1995).

$$
\begin{aligned}
Put(t,T^F) &= P(t,T^F)K\left[1 - \left[N\left(\frac{Upper - M(t,T^F)}{\sqrt{V(t,T^F)}}\right) - N\left(\frac{Lower - M(t,T^F)}{\sqrt{V(t,T^F)}}\right)\right]\right] \\
&\quad - P(t,T)\left[1 - \left[N\left(\frac{UpperF - E}{\sqrt{V(t,T^F)F}}\right) - N\left(\frac{LowerF - E}{\sqrt{V(t,T^F)F}}\right)\right]\right] \\
&\qquad where \\
E &= M(t,T^F) - B(T^F,T)V(t,T^F) \\
F &= 1 + 2C(T^F,T)V(t,T^F)
\end{aligned}
\tag{112}
$$

Appendix D

We are interested in a solution ψ(T) that satisfies the following integral equation:

$$\begin{gathered}\psi(T) - K(s,v) = F(T)\\ where\\ K(s,v) = 2\sigma^2\int_0^T \frac{2\gamma e^{\gamma(T-s)}}{h(s,T)}\int_v^T \frac{e^{\gamma v}h(v,T)}{e^{\gamma s}h(s,T)}\psi(s)dvds\\ F(T) = \sqrt{r^F(0,T) - \frac{1}{2}V(0,T)} + \frac{1}{2}V(0,T)\end{gathered} \qquad \textbf{(113)}$$

If it is possible to write K(s,v) as $\int_0^T K(T,v)\psi(v)dv$ then equation no. 1 can be recognized as a linear second order Volterra integral with a seperable kernel[54]. In that connection, it is known that a linear second order Volterra integral with a continous and bounded kernel has a solution for every continous function F(T).

We are therefore seeking a solution of the following form:

$$\psi(T) - \int_0^T K(T,v)\psi(v)dv = F(T) \qquad \textbf{(114)}$$

Let us then take a more close look at the following integral equation:

$$\begin{gathered}2\sigma^2\int_0^T \frac{2\gamma e^{\gamma(T-s)}}{h(s,T)}\int_v^T \frac{e^{\gamma v}h(v,T)}{e^{\gamma s}h(s,T)}dvds\\ where\\ h(t,T) = [\kappa + \gamma]e^{2\gamma(T-t)} + [\gamma - \kappa]\end{gathered} \qquad \textbf{(115)}$$

This formula can (after lengthly calculation) be shown to be possible to rewrite as follows:

$$2\sigma^2\int_0^T \frac{2\gamma e^{\gamma(T-s)}}{h(s,T)}\int_v^T \frac{e^{\gamma v}h(v,T)}{e^{\gamma s}h(s,T)}dvds = \frac{2\sigma^2 e^{\gamma T}}{h(0,T)}\int_0^T [e^{\gamma v} - e^{-\gamma v}]dv \qquad \textbf{(116)}$$

We are then in a situation were we have to find a solution ψ(T) that satisfies the following linear second order Volterra integral:

---

[54] See Pipkin (1991) chapter 5.

$$\begin{gathered}\psi(T) - \int_0^T K(T,v)\psi(v)dv = F(T)\\ \textit{where}\\ K(T,v) = \frac{2\sigma^2 e^{\gamma T}}{h(0,T)}[e^{\gamma v} - e^{-\gamma v}]\end{gathered} \tag{117}$$

In the case of a seperable kernel (as here) the integral equation can be solved as follows. First we differentiate formula 5 with respect to T, which yields:

$$\begin{gathered}\psi_T(T) - K(T,T)\psi(T) - 2\sigma^2\frac{\gamma e^{\gamma T}h(0,T) - e^{\gamma T}h_T(0,T)}{h(0,T)^2}\int_0^T[e^{\gamma v} - e^{-\gamma v}]\psi(v)dv = F_T(T)\\ \textit{where}\\ h_T(0,T) = 2\gamma[\kappa + \gamma]e^{2\gamma T}\end{gathered} \tag{118}$$

If we then use formula 5, we can express $\int_0^T[e^{\gamma v} - e^{-\gamma v}]\psi(v)dv$ as:

$$\int_0^T[e^{\gamma v} - e^{-\gamma v}]\psi(v)dv = [\psi(T) - F(T)]\frac{h(0,T)}{2\sigma^2 e^{\gamma T}} \tag{119}$$

If we then pluck formula 7 into formula 6, we get:

$$\begin{gathered}\psi_T(T) - \frac{2\sigma^2 e^{\gamma T}}{h(0,T)}[e^{\gamma T} - e^{-\gamma T}]\psi(T) - [\psi(T) - F(T)]\frac{\gamma h(0,T) - h_T(0,T)}{h(0,T)} = F_T(T)\\ \Rightarrow \psi_T(T) - \psi(T)\left[\frac{2\sigma^2 e^{\gamma T}}{h(0,T)}[e^{\gamma T} - e^{-\gamma T}] + \gamma - \frac{2\gamma[\kappa + \gamma]e^{2\gamma T}}{h(0,T)}\right]\\ = F(T)\left[\gamma - \frac{2\gamma[\kappa + \gamma]e^{2\gamma T}}{h(0,T)}\right] + F_T(T)\\ \Rightarrow \psi_T(T) + \psi(T)\kappa = F_T(T) - F(T)\left[\gamma - \frac{2\gamma[\kappa + \gamma]e^{2\gamma T}}{h(0,T)}\right]\\ \Rightarrow \psi_T(T) + \kappa\psi(T) = F(T) + F_T(T)W(T)\\ \textit{where}\\ W(T) = \gamma - \frac{2\gamma[\gamma - \kappa]}{h(0,T)}\end{gathered} \tag{120}$$

That is ψ(T) can be found by solving an ordinary differential equation subject to the boundary condition that ψ(0) = F(0). The solution to this equation can with respect to that be written as:

$$\psi(T) = e^{-\kappa T}\left[F(0) + \int_0^T e^{\kappa s}W(s)F(s)ds + \int_0^T e^{\kappa s}F_s(s)ds\right] \quad \textbf{(121)}$$

The last integral in formula 9 can be rewritten as follows by using partial integration:

$$\int_0^T e^{\kappa s}F_s(s)ds = e^{\kappa T}F(T) - F(0) - \kappa\int_0^T e^{\kappa s}F(s)ds \quad \textbf{(122)}$$

Which means that:

$$\psi(T) = F(T) + e^{-\kappa T}\int_0^T e^{\kappa s}[W(s) - \kappa]F(s)ds \quad \textbf{(123)}$$

If we pluck the expression for W(T) from formula 8 into this formula and simplify we finally find that ψ(T) for all T is given by:

$$\psi(T) = F(T) + 2e^{-\kappa T}\int_0^T e^{\kappa s}V(0,s)F(s)ds \quad \textbf{(124)}$$

Which is identical to formula 55 in the main text.

Appendix E

From equation 51 in the main text we have that $r^F(t,T) = E^{Q_T}[r(T)|F_t]$ and furthermo

have from formula 24 that the state-variable (y) in the extended Vasicek model is governed by the following stochastic process under the T-forward-adjusted probability measure:

$$y(T) = D(t,T)y_t - \sigma^2\int_t^T D(s,T)A(s,T)ds \mp \sigma\int_t^T D(s,T)d\tilde{W}^{Q_T} \qquad \textbf{(125)}$$

As r(T) = y(T) + ψ(T) we have:

$$r(T) = D(t,T)[r_t - \psi(t)] + \psi(T) - \sigma^2\int_t^T D(s,T)A(s,T)ds \mp \sigma\int_t^T D(s,T)d\tilde{W}^{Q_T} \qquad \textbf{(126)}$$

Now we want to express the stochastic process for the instantaneous forward rate under the Q_T-probability measure in terms of the value of the state-variable (including ψ(T)) at time T, that is:

$$r^F(t,T) = r(T) \pm \sigma\int_t^T D(s,T)d\tilde{W}^{Q_T} \qquad \textbf{(127)}$$

Combining formula 2 and formula 3 yields:

$$r^F(t,T) = D(t,T)[r_t - \psi(t)] + \psi(T) - \sigma^2\int_t^T D(s,T)A(s,T)ds + \sigma\int_t^T D(s,T)d\tilde{W}^{Q_T} \qquad \textbf{(128)}$$

For t = 0 we then find that ψ(T) is defined by:

(129)

Which under the assumption that κ(t) and σ(t) are time-independent yields:

$$\psi(T) = r^F(0,T) + \frac{\sigma^2}{2\kappa^2}[1 - e^{-\kappa T}]^{T2} \qquad \textbf{(130)}$$

If we want to find an expression for B(t,T) and A(t,T) in the extended Vasicek model it is most efficient to use the following relationship:

$$r^{F}(t,T) = -B_{T}(t,T) + A_{T}(t,T)r_{t} = D(t,T)[r_{t} - \psi(t)] + \psi(T) - \sigma^{2}\int_{t}^{T} D(s,T)A(s,T)ds \qquad \textbf{(131)}$$

Which means that $B_T(t,T)$ and $A_T(t,T)$ can be found by collecting powers of r_t, which yields:

$$\begin{gathered} A_{T}(t,T) = D(t,T) \\ B_{T}(t,T) = D(t,T)r^{F}(0,t) + \sigma^{2}D(t,T)\int_{0}^{t} D(s,t)A(s,tT)ds - r^{F}(0,T) - \sigma^{2}\int_{0}^{t} D(s,T)A(s,T)ds \end{gathered} \qquad \textbf{(132)}$$

If this expression is integrated from 0 - T (under the assumption of constant κ(t) and σ(t)) and plucked into formula 25 in the main text the result will be identical the result from Hull and White (1993c).

Another way to derive the expression for B(t,T) and A(t,T) which actually is easier as it does not involve much integration because we are cabable of writing the definitions of B(t,T) and A(t,T) in terms of B(0,T), B(0,t) and B(0,T), B(0,t), the technique is as follows:

We have the following well known result:

$$P(t,T) = P(t,T_{1})E^{Q_{T_{1}}}[P(T_{1},T)|F_{t}] \qquad \textbf{(133)}$$

This can alternatively be written as:

$$e^{B(t,T) - B(t,T_{1}) - [A(t,T) - A(t,T_{1})]y} = E^{Q_{T_{1}}}[P(T_{1},T)|F_{t}] \qquad \textbf{(134)}$$

We have from formula 30 in the main text (for $T_1 = T^F$) that the price $P(T_1,T)$ under the probability measure Q_{T_1} can be written as:

$$e^{B(T_{1},T) + \frac{12}{2}A(T_{1},T)^{2}V(t,T_{1}) - A(T_{1},T)M(t,T_{1})} \qquad \textbf{(135)}$$

Combining this equation with formula 10 and collecting terms of equal power in y, we get the following formulas for $B(T_1,T)$ and $A(T_1,T)$:

$$\begin{gathered} A(T_{1},T) = \frac{A(t,T) - A(t,T_{1})}{D(t,T_{1})} \\ B(T_{1},T) = [B(t,T) - B(t,T_{1})] - \frac{1}{2}A(T_{1},T)^{2}V(t,T_{1}) - A(T_{1},T)[\psi(T_{1}) - r^{F}(0,T_{1})] \end{gathered} \qquad \textbf{(136)}$$

For t = 0 we have that B(0,T) = ln P(0,T) more precisely we get the following expressions for $B(T_1,T)$ and $A(T_1,T)$:

$$A(T_1,T) = \frac{A(0,T) - A(0,T_1)}{D(0t,T_1)}$$

$$B(T_1,T) = \ln\left(\frac{P(0,T)}{P(0,T_1)}\right) - \frac{1}{2}A(T_1,T)^2V(0t,T_1) - A(T_1,T)[\psi(T_1) - r^F(0,T_1)] \qquad \textbf{(137)}$$

In order to have a fully defined model, we now only have to find the expression for A(0,T) for all T.

We can get the expression for A(0,T) by using formula 8 and find that A(0,T) is defined as:

$$A(0,T) = \int_0^T D(0,s)ds \qquad \textbf{(138)}$$

It might not be obvious but formula 13 (together with formula 14) evaluated at $t = T_1$, and under the assumption that κ(t) and σ(t) are time-independent yields identical results for B(t,T) and A(t,T) as the expression from formula 8.

Appendix F

In terms of forward-rates we can in abstract form write the Quadratic Interest rate model as[55]:

$$r^F(t,T) = b_1(t,T)y_t^2 + b_2(t,T)y_t + a(t,T) \tag{139}$$

Where the process for y_t is defined by the following SDE[56]:

$$\begin{gathered} dy_t = [\varphi(t) - \kappa y_t]dt + \sigma d\tilde{W}_t \\ where \\ d\tilde{W}_t = dW_t - \lambda\sigma \end{gathered} \tag{140}$$

The term structure appears as a quadratic function of the factor y_t - which follows a Vasicek type process. Let us in this connection remark that the unconditional correlation of y_t^2 and y_t is not 1 - as a consequence of this the Quadratic Interest rate model behaves - from an empirical point of view - like a linear two-factor model.

Let us therefore consider the following two-factor linear model[57]:

$$r^F(t,T) = a_1(t,T)f_{1t} + a_2(t,T)f_{2t} + b(t,T) \tag{141}$$

From this it can be seen that the Quadratic Interest rate model can be thought of as a linear factor model, where $f_{1t} = f_{2t}^2$.

In general we can write the vector process for f_t as:

$$df_t = [\Phi(t) - Bf_t]dt + S(F_t)d\tilde{W}_t \tag{142}$$

In particular we have:

---

[55] The expression for $b_1(t,T)$, $b_2(t,T)$ and $a(t,T)$ can be derived from formula 37 in the main text.

[56] Se formula 32 in the main text.

[57] Where the functional form for $b_1(t,T)$, $b_2(t,T)$ and $a(t,T)$ can be derived from the closed form solution for zero-coupon bond prices which is implied by the SDE for the two-factors f_{1t} and f_{2t} and their linear connection with the spot-rate, i.e, $r_t = x^T f_t$ - where x is a Boolean vector.

$$df_{2t} = [\varphi_2(t) - \kappa_1 f_{1t} - \kappa_2 f_{2t}]dt + \sigma_2(f_t)^T d\tilde{W}_t \qquad \textbf{(143)}$$

And, since $f_{1t} = f_{2t}^2$ we have:

$$df_{1t} = [2f_{2t}[\varphi_2(t) - \kappa_1 f_{1t} - \kappa_2 f_{2t}] + \sigma_2(f_t)^2]dt + 2f_{2t}\sigma_2(f_t)^T d\tilde{W}_t \qquad \textbf{(144)}$$

Equation 5 and 6 imply that:

$$S(f_t)S(f_t)^T = \sigma_2(f_t)^2 \begin{pmatrix} 4f_{2t}^2 & 2f_{2t} \\ 2f_{2t} & 1 \end{pmatrix} \qquad \textbf{(145)}$$

From this we can deduce that the matrix $(S(f_t)S(f_t)^T)$ is linear in terms of f_t - if and only if $\sigma_2(f_t)^2$ does not depend on f_t. This is the explanation why the Quadratic Interest rate model needs a deterministic volatility for the state-variable. Furthermore, the drift for f_{1t} should be linear in f_t, and then $\kappa_1 = 0$.

This means that the Quadratic Interest rate model is equivalent to a two-factor linear framework where there is only one state-variable f_{2t} - following a Vasicek process. That is, the term structure depends linearly on f_{2t} and f_{2t}^2 .

To conclude these results, let me remark that the Quadratic Interest rate model appears a particular kind of two-factor linear model, satisfying:

$$r^F(t,T) = b(t,T)^T f_t + a(t,T) \qquad \textbf{(146)}$$

where there exists a quadratic function g such that:

$$\begin{gathered} g(f_t) = 0 \\ \textit{for} \\ g(f_t) = f_{2t}^2 - f_{1t} \end{gathered} \qquad \textbf{(147)}$$

That is the Quadratic Interest rate model can be thought of as a particular case of the class of linear factor models.

Appendix G

In this Appendix I will shortly present the different variance reduction methods that I have employed in some of the examples in section 9.

This Appendix will consist of two parts.

The first part will discuss the variance reduction methods I am using in simulating the forward-price in the Extended Vasicek model under the T^F-adjusted probability measure. The problem considered here can be recognized as a 1-dimensional simulation problem.

The second part will present the variance reduction methods which were used in connection with sampling the spot-rate process from a discretization of the SDE. This problem can for this reason considered a K-dimensional simulation problem, where K is the number of steps. As we only consider equidistant discretization points, we have that the time-step δ is equal to $\frac{T - t}{K}$, for T-t being the total sampling interval.

This does not neccesarily mean that the methods employed in part 1 cannot be used on the problem considered in part 2. Actually there is nothing to stop us from doing exactly that.

On the other hand the approaches I will discuss in part 2 are not useful for the particular type of 1-dimensional problem I am considering. This will be obvious from the presentation in part 2.

The variance reduction method I will discuss in part 1 is:

Antithetic variates
Stratified sampling
Empirical Martingale simulation

For the more complex simulation problem considered in part 2 I have chosen two new methods which looks promising:

The Brownian Bridge technique
Girsanovs transformation

Part 1: Variance Reduction Techniques - I[58]

The equation we are considering in this example is given by formula 86 in the main text, ie.:

$$P(T^F,T) = \frac{P(0,T)}{P(0,T^F)}\exp\left[-\frac{1}{2}\int_0^{T^F}[\sigma_P(s,T) - \sigma_P(s,T^F)]^2 ds - \int_0^{T^F}[\sigma_P(s,T) - \sigma_P(s,T^F)]dW^{Q_{TF}}(s)\right] \quad \textbf{(148)}$$

Where $dW^{Q_{TF}} \approx N(0,1)$. Let us in that connection denote z_n, for n = [1,N], the vector of random-numbers simulated from the distribution N(0,1).

Antithetic variates:

One of the simplest and most widely used variance reduction methods in finance is the method of antithetic variates.

The method of antithetic variates is based on the observation that if z_n, has a standard normal distribution, then so does $-z_n$.

Let us now assume that we have simulated the bond-price $P(T^F,T)$ for a given N - let us denote this simulated price as $P^+(T^F,T)$. If we now also simulate the price where we have replaced each z_n with $-z_n$ - and denote this price $P^-(T^F,T)$ - then we have that an unbiased estimator of the forward price is given by:

$$P(T^F,T) \approx \tilde{P}(T^F,T) = \frac{1}{N}\sum_{n=1}^{N}\frac{P^+(T^F,T) + P^-(T^F,T)}{2} \quad \textbf{(149)}$$

We have that the random inputs obtained from the collection of antithetic pairs are more regularly distributed than a collection of 2N independent samples. In particular, we have that the sample mean over the antithetic pairs will always equal the population mean of 0, whereas the mean over finitely many independent samples is almost surely different from 0.

---

[58] The discussion in this section relies mainly on Duan and Simonato (1998) and Boyle, Broadie amd Glasserman (1995).

More precisely we have the following relationship, because $P^+(T^F,T)$ and $P^-(T^F,T)$ has the same variance:

$$Var\left[\frac{P_n^+(T^F,T) + P_n^-(T^F,T)}{2}\right] = \frac{1}{2}[Var[P_n^+(T^F,T)] + Cov[P_n^+(T^F,T), P_n^-(T^F,T)]] \qquad \textbf{(150)}$$

From formula 3 it follows that $Var[\tilde{P}(T^F,T)] \le Var[P^+(T^F,T)]$ if $Cov[P_n^+(T^F,T), P_n^-(T^F,T)] \le Var[P_n^+(T^F,T)]$.

Because we need twice as many replications to calculate/estimate $\tilde{P}(T^F,T)$ - that is twice the amount of work. We require the following inequality to hold in order for antithetic variates to have increased the efficiency, $2Var[\tilde{P}(T^F,T)] \le Var[P^+(T^F,T)]$ - which b 3 simplifies to the following requirement; $Cov[P_n^+(T^F,T), P_n^-(T^F,T)] \le 0$ - simulated prices should be negatively correlated[59].

Stratified sampling:

The idea behind stratified sampling is as follows: stratified sampling seeks to make the inputs to the simulation more regular than random inputs. In particularly it forces certain empirical probabilities to match theoretical probabilities[60].

Let us consider N simulated normal random numbers. The empirical distribution of an independent sample (z_1, z_2, z_3,...., z_N) will look only roughly like the normal density function, the tails of the distribution - which often is the most important part - will inevitably be under represented.

Using stratified sampling we can force exactly one observation to lie between the (n - 1)th and nth percentile, for n = [1,N]. This will - by construction - produce a better match to the normal distribution.

---

[59] A proof that this requirement is always met is given in Boyle, Broadie and Glasserman (1995).

[60] This is in a sense similar to the idea behind moment matchning methods - which force the empirical moments to match the theoretical moments, see Barraquand (1994).

The way we have chosen to implement this method is as follows: First we generate N independent uniform random variates ($u_1, u_2, u_3, \ldots, u_N$) - uniform on [0,1]. The vector of independent normal random variates can now be derived from the following formula:

$$z_n = N^{-1}\left(\frac{n + u_n - 1}{N}\right) \qquad \textit{for } n \in [1,N] \tag{151}$$

Where N^{-1} is the inverse of the cumulative normal distribution.

The reason that this works is because $\frac{n + u_n - 1}{N}$ falls between the (n - 1)th and nth percentile of the uniform distribution and percentiles are preserved by the inverse transformation.

Empirical Martingale simulation:

The basic idea behind the Empirical Martingale Simulation (EMS) method from Duan and Simonato (1998) can be explained this way:

Lets us for this purpose consider the exponential part of equation 1:

$$\tilde{A} = \frac{1}{N}\exp\left[-\frac{1}{2}\int_0^{T^F}[\sigma_P(s,T) - \sigma_P(s,T^F)]^2 ds - z_n\int_0^{T^F}[\sigma_P(s,T) - \sigma_P(s,T^F)]\right] \tag{152}$$

Under the Q_{TF}-probability measure we have that the expected value of $\tilde{A}$ is 1. Of course for a finite number of simulated normal random variates the average value calculated in formula 5 is almost surely different from 1.

The idea in the EMS-method is, because of this observation, to adjust each element in the vector ($a_1, a_2, a_3, \ldots, a_N$), for $\tilde{A} = \frac{1}{N}\sum_{n=1}^{N} a_n$ - that is each a_n can be recognized as a particular realisation of the exponential part in equation 5. The adjustment is done to ensure that we for all N have that: $\tilde{A} = 1$.

From this it follows that EMS bears some resemblance to the moment matching method of Barraquand (1994). However, the EMS method is entirely different from the moment matching method because of the following reason: For the EMS the correct first moment in simulation is ensured by the use of a multiplicative adjustor instead of an additive one as in moment matching[61]. This difference is very important as asset prices typically (as here) are modelled as exponential martingales. From this it follows that the multiplicative adjustment ensures no domain violation, whereas the additive adjustment cannot.

Part 2: Variance Reduction Techniques - 2[62]

The equation we are considering here is given by formula 84 in the main text, ie:

$$dx_t = -\kappa x_t dt + \sigma dW_t \qquad for\ x_0 = 0 \tag{153}$$

where $dW_t \approx \Delta W_k = W_{(k+1)\delta} - W_{k\delta}$ is distributed as N(0,δ). Therefore we h

$\Delta W_k = z_k\sqrt{\delta}$.

As mentioned above we assume that K is the number of steps, and the time discretization points are of an equal length δ. Furthermore we denote each discretization point as kδ, for k = [1,K]. Lastly we assume that we for each k perform N simulations.

The Brownian Bridge technique:

As far as I know the only other work that uses the Brownian Bridge method to reduce the effective dimensionality is Caflisch, Morokoff and Owen (1997). Let me for that reason explain the idea behind using the Brownian Bridge method in some detail[63]:

For convenience I will assume that the number of time-steps is a power of 2. Let Z(n) = (z_1(n), z_2(n), z_3(n),...., z_K(n)), for n = [1,N], be a vector of independent normal random variates - that is Z(n) can be seen as a particular path for the SDE over the interval [δ,Kδ].

This vector we will use to construct a Brownian Bridge simulated Wiener process, which we will denote B_k, for k = [0,K]. In particular we have that $B_0 = 0$. We will in the construction assume that the quality of the simulated normal random variates drops with the dimension number[64].

---

[61] When moment matching is applied to asset prices.

[62] The discussion in this section relies mainly on Caflisch, Morokoff and Owen (1997) and Schoenmakers and Heemink (1997).

[63] The descriptions of the Brownian Bridge method rely on Caflisch, Morokoff and Owen (1997) and Karatzas and Shreve (1988, section 2.3).

[64] This is generally the case, at least in the case of quasi-random number sequences.

As the variance of a Wiener process is linear in the time-length - the first point to generate will be B_K, ie:

$$B_K = z_K(n)\sqrt{K\delta} \tag{154}$$

We have now got both the starting point and the end point for the Wiener process. Viewing these as fixed points, we can think of the Wiener process as being "tied-down" - this can be recognized as a Browning Bridge process.

The next point to simulate will be the one that has the highest variance, which for the Brownian Bridge process is the midpoint of the interval. Conditioned on B_0 and B_K, the midpoint is normally distributed as $N\left(\frac{1}{2}B_K, \frac{1}{4}K\delta\right)$, which mean that we can simulate the midpoint by setting:

$$B_{\frac{K}{2}} = \frac{B_K}{2} + z_{\frac{K}{2}}(n)\sqrt{\frac{K\delta}{4}} \tag{155}$$

We now have two endpoints and a midpoint. This can in a sense be thought of as two consecutive Brownian Bridge processes - conditioned on each of the intervals.

Again we will simulate the process where the variance is highest, that is at the midpoints of each of the two intervals.

In general we have that each B_k can be derived from the following formula (if K is of the power of 2)[65]:

$$B_k = \frac{B_{(k-1)} + B_{(k+1)}}{2} + z_k(n)\sqrt{\frac{K\delta}{2^v}} \quad for\ v\ \varepsilon \left(1,\ 1 + \frac{\ln K}{\ln 2}\right) \tag{156}$$

After having simulated our K+1-dimensional vektor B_k for each k = [0,K], we can construct our simulated Wiener process from this. Actually we have that we can form K-simulated Wiener increments as follows:

$$\Delta W_k = B_k - B_{k-1} \quad for\ k\ \varepsilon\ [1,K] \tag{157}$$

---

[65] See also equation 5.2 in Caflisch, Morokoff and Owen (1997).

Accordingly to Corollary 2.3.4 in Karatzas and Shreve (1988) this should produce paths that are asymptotical Wiener paths.

This new constructed vector of Wiener increments can now directly be used instead of $\Delta W_k = z_k\sqrt{\delta}$ in our simulation procedure.

Variance reduction using Girsanovs transformation:

I will now turn my attention to the Girsanovs transformation technique as a means of reducing the variance in the simulation procedure.

This method is based on a Girsanov transformation of the original SDE for x.

Equation 6 can alternatively be written as:

$$x_t = x_0 - \int_0^t \kappa x_s ds + \int_0^t \sigma dW_s \quad for\ x_0 = 0 \tag{158}$$

The basic idea in the weak discretizations scheme is to approximate the functional:

$$u(s,x) = E[h(x_T)|x_s = x] \tag{159}$$

where h is a specific function and time s = 0. If we have that the function h and the drift-and diffusion-coefficients in formula 11 are sufficiently smooth, then the function u satisfies the Kolmogorov backward equation, ie:

$$L^0 u(s,x) = 0 \tag{160}$$

Subject to the end condition $u(T,x) = h(x)$ for all x ε $\Re$, and where the operator L^0 is defined as:

$$L^0 = \frac{\partial}{\partial s} + \mu\frac{\partial}{\partial x} + \frac{1}{2}\sigma^2\frac{\partial^2}{\partial x^2} \qquad where \qquad \mu = -kx \tag{161}$$

Using Girsanovs theorem we can transform the underlying probability measure P, so that the process $\tilde{W}$ is defined by:

$$\tilde{W} = W - \int_0^t v(s,x_s)ds \tag{162}$$

is a Brownian motion with respect to the transformed probability measure $\tilde{P}$ with Radon-Nikodym derivative:

$$\frac{d\tilde{P}}{dP} = \frac{\Theta_t}{\Theta_0} \tag{163}$$

Under the change of probability measure the Ito process $\tilde{x}$ satisfies the stochastic differential equation:

$$\tilde{x}_t = \tilde{x}_0 - \int_0^t [\kappa\tilde{x}_s + v(s,\tilde{x}_s)\sigma]ds + \int_0^t \sigma dW_s \quad for\ \tilde{x}_0 = 0 \tag{164}$$

Furthermore we have that the process Θ satisfies the equation:

$$\Theta_t = \Theta_0 + \int_0^t \Theta_s v(s,\tilde{x}_s)dW_s \quad Where\ \Theta_0 \neq 0 \tag{165}$$

It is obvious that the process $\tilde{x}$ in equation 17 is an Ito process with respect to $\tilde{P}$ with the same drift-and diffusion coefficients as the Ito process x in formula 11. From this and equation 16 it follows that:

$$u(s,x) = E[h(x_T)] = \frac{E[h(\tilde{x}_T)]\Theta_T}{\Theta_0} \tag{166}$$

This indicates that by estimating the expectation of the random variable om the right hand side of equation 19 we are evaulating the functional given in 12.

Furthermore this result is independent of the function v. The question is therefore how to choose v to get the most efficient variance reduction.

Given the Radon-Nikodym derivative, we can by using Ito's lemma deduce that the product $u(s,\tilde{x})\Theta_t$ under the probability measure P, follows:

$$d(u(s,\tilde{x})\Theta_t) = \Theta_t\big(u(s,\tilde{x})v(s,\tilde{x}) + u_x(s,\tilde{x})\sigma\big)dW_s \tag{167}$$

What we hope for is that the variance of the right side of equation 19 is much smaller than the variance of u(s,x). It follows immediately from formula 20 that this variance is reduced to zero if $v(s,\tilde{x})$ satisfies:

$$v(s,\tilde{x}) = -\frac{u_x(s,\tilde{x})\sigma}{u(s,\tilde{x})} \tag{168}$$

In which case we have[66]:

$$u(0,x) = \frac{E[h(\tilde{x}_T)]\Theta_T}{\Theta_0} \tag{169}$$

That is the variable is non-random and therefore the variance is reduced to zero.

For the construction of the parameter function in 21 we need to know the solution u of the Kolmorogov backward equation.

In our particular case it happens to be possible to obtain the solution - the procedure is as follows:

Let us assume that x represents the spot-rate and we want to derive the price of a zero-coupon bond - $P_b(t,T)$, for all T = [1,τ] where τ is the maximum maturity date. The end condition for the Kolmogorov backward equation[67] will in that case be $P_b(T,T) = u(T,x) = h(x) = 1$

From Duffie (1992 Appendix E) we have that the Feynman-Kac solution to the Kolmogorov backward equation under the above assumption - if it exists - can be written as:

---

[66] The proof is here obmitted, but can be derived from the SDE's from equation 17 and 18, together with 21 and 13, see Kloeden and Platen (1995 section 16.2).

[67] It should here be stressed that the Kolmogorov backward equation in the case of the pricing of zero-coupon bonds given a spot-rate process is not identical to the one specified in formula 14. For our purpose - namely figuring out the functional from for v - this is not of importance, as will be apparent in the derivation. It would be important if we tried to solve the PDE directly - but here we are utilising a probabilistic approach.

$$P_b(t,T) = u(t,x) = E\left[\exp - \int_t^T x_s ds\right] \quad \textbf{(170)}$$

As x is normally distributed, the solution to 23 can be written as:

$$P_b(t,T) = u(t,x) = e^{-E[\int_t^T x_s ds] + \frac{1}{2}V[\int_t^T x_s ds]} \quad \textbf{(171)}$$

We know that the expected value and variance for $Y(t,T) = \int_t^T x_s ds$, can be written as:

$$E[Y(t,T)] = \int_t^T e^{-\kappa(v - t)}\left[x + \int_t^v e^{\kappa(s - t)}\mu ds\right]dv$$

$$V[Y(t,T)] = \int_t^T e^{-2\kappa(v - t)}\left(\int_t^v e^{\kappa(s - t)}\sigma ds\right)^2 dv \quad \textbf{(172)}$$

where we on purpose have specified the drift in an abstract form.

From this we can easily derive the functional form for v as $\frac{u_x(s,\tilde{x})}{u(s,\tilde{x})} = \frac{u_x(s,x)}{u(s,x)}$ - it follows directly from formula 21, 24 and 25 - and the result is:

$$v(s,\tilde{x}) = \sigma\int_t^T e^{-\kappa(s - t)}ds \quad \textbf{(173)}$$

The two processes we are interested in, can now be expressed as:

$$
\begin{gathered}
d\tilde{x}_t = [-\kappa\tilde{x}_s - \sigma^2 \int_t^T e^{-\kappa(s-t)}ds]dt + \sigma dW_t \qquad for\ \tilde{x}_0 = 0 \\
and \\
d\Theta_t = \Theta_t \left(\int_t^T e^{-\kappa(s-t)}ds \right) dW_t \qquad for\ \Theta_0 = 1
\end{gathered}
\tag{174}
$$

From 22 we have that $P_b(t,T) = E\left[e^{-\int_t^T x_s ds}\right] = E\left[\Theta_T e^{-\int_t^T \tilde{x}_s ds}\right]$, for all T = [t,τ].

It should be stressed that $P_b(t,T)$ is not the price we are interested in, instead I will refer to $P_b(t,T)$ as our basic price - just as x is our basic interest rate process. The “true” price is derived by appropriate adjustment of the basic process x for each T in order for $P_b(t,T)$ to be equal to the market price P(t,T).

It now follows that using Girsanovs transformation in the simulation procedure requires the simultanous simulation of the two processes in formula 27.

In the main text - section 8 - we showed what the different discretizations schemed looked like for the SDE for x. Deriving the discretizations schemes for the process $\tilde{x}$ is straightforward and is for that reason left for the reader.

Deriving the actual form for the correction process Θ_t for the Euler scheme is straightforward, for that reason I will just show the order 1.5 strong Taylor discretization scheme as the others can be derived from here, ie:

$$
\begin{gathered}
\Theta_{k+1} = \Theta_k + \Theta_k\sigma_P(t,T)\Delta W + \frac{1}{2}\Theta_k\sigma_P^2(t,T)((\Delta W)^2 - \delta) \\
+ \frac{1}{2}\Theta_k\sigma_P^3(t,T)[\Theta_k - 1]\left(\frac{1}{3}(\Delta W)^2 - \delta\right)\Delta W \\
where \\
\sigma_P(t,T) = \int_t^T e^{-\kappa(s-t)}ds
\end{gathered}
\tag{175}
$$

It follows now that the order 1 strong Milstein will be equal to the order 2 weak Milstein, and furthermore is given by the first line in equation 28[68].

---

[68] Symbolizing v by $\sigma_P(t,T)$ is not a coincidence, as the procedure here is based on a reverse application of Girsanovs transformation for a shift of probability measure. It is namely well known from the theory of contingent claim pricing that arbitrage-free pricing under the T-adjusted probability measure - for the Ornstein-Uhlenbeck process - results in a likelihood-ratio process for $\sigma_P(t,T)$. $\sigma_P(t,T)$ can because of that be recognized as the bond-price volatility in the Vasicek model and in the Extended Vasicek model - because the bond-price volatility is independent of the long-term mean value (θ or $\theta(t)$).

Appendix H

We have that both the Extended Vasicek model and the Quadratic interest rate model is analytically very tractable.

The (semi) analytical expression for the Quadratic interest rate model is given in proposition no. 1 in the main text.

With respect to the Extended Vasicek model, can we from equation 13 in Appendix E derive the analytical solution for the price of a zero-coupon bond, ie.

$$\begin{aligned} P(t,T) &= e^{B(t,T) - A(t,T)r} \\ &\text{for} \\ A(t,T) &= \frac{1 - e^{-\kappa(T-t)}}{\kappa} \\ &\text{and} \\ B(t,T) &= \ln\frac{P(0,T)}{P(0,t)} - A(t,T)\frac{\partial \ln P(0,t)}{\partial t} - \frac{\sigma^2}{4\kappa^3}[e^{-\kappa T} - e^{-\kappa t}]^2[e^{2\kappa t} - 1] \end{aligned} \qquad \textbf{(176)}$$

Which is identical to the analytical expression for the Hull and White model, see Hull (1993 page 404).

Using the analytical properties for these two models when pricing derivative securities is clearly efficient - both in terms of calculation speed and convergence rate. The reason for this can be explained as follows:

Let us now assume we for example wish to calculate the price of a 3-year Bermudan[69] call-option on a 10-year zero-coupon bond.

If no analytical expression is available for the spot-rate process[70] we choose as our basic pricing model - then this means we have to simulate[71] the spot-rate process all the way out to

---

[69] The holder of a Bermudan option, also known as a limited exercise, mid-Atlantic or semi-American option, has the right to exercise the option on one or more possible dates prior to its expiry.

[70] Which for example is the case for the Black and Karisinski model.

[71] Using the phrase “simulate”, we are referring to either a Monte Carlo simulation procedure or a lattice-based method.

10 years. If on the other hand an analytical expression is available for the bond-price then we only have to simulate the spot-rate process out to the time of expiry of the option - in this case 3-years.

In the analytical expression for the bond price in formula 1, the variable r is the instantaneous spot-rate. This is of course also true for the Quadratic interest rate model - namely that the spot-rate in the analytical expression is the instantaneous spot-rate.

However; in the Hull and White initial yield-curve matching-procedure the interest rates are Δt-period rates - that is equal to δ in the examples in section 9.

This means that we cannot just simulate the spot-rate process out to - for example - 3 years and then directly use the vector of simulated spot-rates and pluck them into the analytical expressions.

First we have to transform the Δt-period spot-rates into instantaneous spot-rates.

Let now r_Δ be the Δt-period spot-rate at time T^F and r be the instantaneous spot-rate at time T^F. Using equation 1 we have the following relationship in the case of the Extended Vasicek model:

$$e^{-r_\Delta \Delta t} = e^{B(T^F, T^F + \Delta t) - A(T^F, T^F + \Delta t) r} \tag{177}$$

So that:

$$r = \frac{B(T^F, T^F + \Delta t) + r_\Delta \Delta t}{A(T^F, T^F + \Delta t)} \tag{178}$$

In the case of the Quadratic interest rate model, we have:

$$\begin{gathered} e^{-r_\Delta \Delta t} = e^{A(T^F, T^F + \Delta t) - B(T^F, T^F + \Delta t) q - C(T^F, T^F + \Delta t) q^2} \\ where \\ q = (\sqrt{r} - \psi(T^F)) \end{gathered} \tag{179}$$

From this we can deduce that r can expressed as follows as a function of the Δt-period spot-rate:

$$\begin{gathered} r = \left(\max[V_1,\ V_2] + \psi(T^F)\right)^2 \\ \textit{where} \\ V_1 = \frac{-B(T^F,T^F + \Delta t) - \sqrt{d}}{2C(T^F,T^F + \Delta t)} \\ \textit{and} \\ V_2 = \frac{-B(T^F,T^F + \Delta t) + \sqrt{d}}{2C(T^F,T^F + \Delta t)} \\ \textit{for} \\ d = B(T^F,T^F + \Delta t)^2 + 4C(T^F,T^F + \Delta t)[r_\Delta \Delta t - A(T^F,T^F + \Delta t)] \end{gathered} \qquad \textbf{(180)}$$

In the calculation performed in section 9 for both the Extended Vasicek model and the Quadratic interest rate model, we have used equation 3 and 5 respectively in order to transform the Δt-period spot-rates into instantaneous spot-rates.

These derived instantaneous spot-rates are then used in connection with the analytical expression for each of the models to calculate the forward-price from the options exercise date to the bonds maturity date.

This technique was used both in connection with the lattice-based pricing approach and the constrained Monte Carlo simulation approach.

Appendix I

Table 1: Call-option prices in the Extended Vasicek Model using Constrained Crude Monte Carlo

	K=8	Std. Error	K=16	Std. Error	K=32	Std. Error	K=64	Std. Error	K=128	Std. Error
Euler Discretization scheme - order 1 weak Taylor approximation										
52,53	*13,120*	*0,008*	*12,995*	*0,006*	*12,939*	*0,007*	*12,907*	*0,007*	*12,892*	*0,007*
62,53	*4,810*	*0,025*	*4,685*	*0,025*	*4,620*	*0,023*	*4,600*	*0,024*	*4,592*	*0,030*
67,53	*1,799*	*0,051*	*1,735*	*0,042*	*1,682*	*0,038*	*1,667*	*0,036*	*1,663*	*0,046*
72,53	*0,412*	*0,032*	*0,386*	*0,026*	*0,369*	*0,028*	*0,354*	*0,030*	*0,357*	*0,033*
82,53	*0,005*	*0,004*	*0,003*	*0,003*	*0,004*	*0,003*	*0,004*	*0,003*	*0,004*	*0,004*
Milstein Discretization scheme - order 2 weak Taylor approximation										
52,53	*13,124*	*0,009*	*12,996*	*0,007*	*12,937*	*0,007*	*12,909*	*0,006*	*12,889*	*0,009*
62,53	*4,798*	*0,029*	*4,692*	*0,029*	*4,635*	*0,030*	*4,601*	*0,027*	*4,597*	*0,031*
67,53	*1,794*	*0,046*	*1,721*	*0,041*	*1,679*	*0,037*	*1,658*	*0,040*	*1,662*	*0,044*
72,53	*0,421*	*0,030*	*0,381*	*0,029*	*0,360*	*0,028*	*0,345*	*0,033*	*0,356*	*0,026*
82,53	*0,005*	*0,004*	*0,004*	*0,003*	*0,004*	*0,005*	*0,003*	*0,004*	*0,003*	*0,003*

Table 2: Call-option prices in the Quadratic Interest Rate Model using Constrained Crude Monte Carlo

	K=8	Std. Error	K=16	Std. Error	K=32	Std. Error	K=64	Std. Error	K=128	Std. Error
Euler Discretization scheme - order 1 weak Taylor approximation										
52,53	*13,104*	*0,003*	*12,988*	*0,002*	*12,932*	*0,002*	*12,905*	*0,002*	*12,890*	*0,002*
62,53	*4,537*	*0,007*	*4,422*	*0,006*	*4,365*	*0,006*	*4,338*	*0,008*	*4,324*	*0,006*
67,53	*0,986*	*0,023*	*0,921*	*0,025*	*0,875*	*0,020*	*0,865*	*0,024*	*0,860*	*0,019*
72,53	*0,024*	*0,005*	*0,018*	*0,005*	*0,016*	*0,005*	*0,016*	*0,004*	*0,016*	*0,004*
82,53	*0,000*	*0,000*	*0,000*	*0,000*	*0,000*	*0,000*	*0,000*	*0,000*	*0,000*	*0,000*
Milstein Discretization scheme - order 2 weak Taylor approximation										
52,53	*13,104*	*0,002*	*12,988*	*0,002*	*12,932*	*0,002*	*12,904*	*0,002*	*12,891*	*0,002*

62,53	*4,535*	*0,005*	*4,421*	*0,006*	*4,366*	*0,007*	*4,339*	*0,005*	*4,325*	*0,005*
67,53	*0,991*	*0,019*	*0,913*	*0,022*	*0,882*	*0,018*	*0,868*	*0,016*	*0,855*	*0,014*
72,53	*0,024*	*0,005*	*0,018*	*0,005*	*0,015*	*0,005*	*0,015*	*0,004*	*0,016*	*0,005*
82,53	*0,000*	*0,000*	*0,000*	*0,000*	*0,000*	*0,000*	*0,000*	*0,000*	*0,000*	*0,000*

Appendix J

Table 1: Call-and Put-option prices in the Extended Vasicek Model using Constrained Brownian Bridge Monte Carlo Simulation

	K=8	Std. Error	K=16	Std. Error	K=32	Std. Error	K=64	Std. Error	K=128	Std. Error
Call-Option prices										
52,53	13,135	0,008	13,006	0,010	12,950	0,007	12,918	0,009	12,904	0,006
62,53	4,840	0,027	4,723	0,034	4,677	0,028	4,640	0,031	4,629	0,031
67,53	1,902	0,042	1,805	0,047	1,756	0,049	1,745	0,052	1,715	0,047
72,53	0,455	0,034	0,427	0,038	0,408	0,035	0,398	0,035	0,392	0,024
82,53	0,006	0,004	0,007	0,006	0,005	0,003	0,005	0,004	0,004	0,004
Put-Option prices										
52,53	0.000	0.000	0.000	0.001	0.000	0.000	0.000	0.001	0.000	0.000
62,53	0.296	0.025	0.309	0.025	0.311	0.027	0.310	0.026	0.318	0.024
67,53	1.627	0.038	1.668	0.036	1,680	0.036	1.709	0.045	1.687	0.038
72,53	4.503	0.035	4.590	0.025	4.630	0.025	4.654	0.028	4.665	0.019
82,53	12.629	0.008	12.757	0.007	12.810	0.006	12.842	0.005	12.857	0.006

Table 2: Call-and Put-option prices in the Quadratic Interest Rate Model using Constrained Brownian Bridge Monte Carlo Simulation

	K=8	Std. Error	K=16	Std. Error	K=32	Std. Error	K=64	Std. Error	K=128	Std. Error
Call-Option prices										
52,53	13,109	0,002	12,992	0,003	12,936	0,002	12,908	0,003	12,895	0,002
62,53	4,543	0,006	4,429	0,007	4,374	0,008	4,344	0,006	4,332	0,007
67,53	1,038	0,020	0,956	0,022	0,917	0,025	0,904	0,026	0,889	0,024
72,53	0,028	0,005	0,023	0,006	0,021	0,006	0,020	0,005	0,019	0,005
82,53	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
Put-Option prices										

52,53	*0,000*	*0,000*	*0,000*	*0,000*	*0,000*	*0,000*	*0,000*	*0,000*	*0,000*	*0,000*
62,53	*0,021*	*0,005*	*0,021*	*0,004*	*0,022*	*0,006*	*0,022*	*0,007*	*0,022*	*0,005*
67,53	*0,797*	*0,020*	*0,839*	*0,020*	*0,855*	*0,020*	*0,875*	*0,024*	*0,876*	*0,021*
72,53	*4,090*	*0,005*	*4,203*	*0,005*	*4,256*	*0,005*	*4,281*	*0,003*	*4,294*	*0,004*
82,53	*12,646*	*0,002*	*12,763*	*0,002*	*12,818*	*0,002*	*12,846*	*0,002*	*12,861*	*0,002*

Part II

Chapter 5

Empirical Yield-Curve Dynamics, Scenario Símulation and Risk-Measures

EMPIRICAL YIELD-CURVE DYNAMICS, SCENARIO SIMULATION AND RISK-MEASURES

Claus Madsen

Abstract: This paper has two objectives. First we will construct a general model for the variation in the term structure of interest rates, or to put it another way, we will define a general model for the shift function. Secondly, we will specify a Risk model which uses the shift function derived in the first part of the paper as its main building block.

Using Principal Component Analysis (PCA) we show that it takes a 4 factor model to explain the variation in the term structure of interest rates over the period from the beginning of 2002 to early-2012. These 4 factors can be called a Slope factor, a Short-Curvature Factor, a Short Factor and a Curvature Factor.

Using the methodology of Heath, Jarrow and Morton (1990) we now specify a 4-factor model for the dynamic in the term structure of interest-rates.

This 4-factor model is afterwards being extended to have a stochastic volatility part, which we assume is to be modelled with a GARCH(1,1) process. The resulting 4-factor yield-curve model belongs to the class of USV (unspanned stochastic volatility) models, as the volatility part is un-correlated with the PCA model for the variation in the yield-curve.

Our Risk-Model relies on the scenario simulation procedure of Jamshidian and Zhu (1997). The general idea behind the scenario simulation procedure is to limit the number of portfolio evaluations by using the factor loadings derived in the first part of paper and then specify particular intervals for the Monte Carlo simulated random numbers and assign appropriate probabilities to these intervals (states).

Our overall conclusion is the following:

- The Jamshidian and Zhu scenario simulation methodology is best suited for the calculation of the Risk-Measure ETL - less for VaR
- We find that the scenario simulation procedure is computational efficient, because we with a limited number of states is capable of deriving robust approximations of the probability distribution
- We also find that it is very useful for non-linear securities (Danish Mortgage-Backed-Bonds MBBs), and argue that the method is feasible for large portfolios of highly complex non-linear securities - for example Danish MBBs
- Backtesting the Risk-Model setup during 2008 showed some very promising results as we were able to capture the extreme price-movements that were observed in the market

Keywords: Multi-factor models, PCA, empirical yield-curve dynamics, APT, VaR, ETL, Risk-Model, Stochastic Volatility, Monte Carlo simulation, scenario simulation, non-linear securities - Danish MBBs

Empirical Yield-Curve Dynamics, Scenario Simulation and Risk-Measures[1]

1. Introduction

This paper has two objectives. First we will construct a general model for the variation in the term structure of interest rates, or to put it another way, we will define a general model for the shift function. Secondly, I will specify a Risk model which uses the shift function derived in the first part of the paper as its main building block.

This general model of the variation in the term structure of interest rates is assumed to belong to the linear class, and, moreover, the factors that determine the shift function are independent; the model is therefore comparable to the Ross (1976) APT model.

The traditional approach to describing the dynamics of the term structure of interest rates is either by defining the stochastic process that drives one or more state variables, such as Cox, Ingersoll and Ross (1985), Vasicek (1977) and Longstaff and Schwartz (1991), or by postulating one or more volatility structures for determining the dynamics of the initial term structure of interest rates, such as Heath, Jarrow and Morton (1991).

However, the approach used in this paper to describe the dynamics of the term structure of interest rates is an empirical approach in order to derive the number of factors needed to describe the variation in the term structure of interest rates. Our reference period will here be 2 January 2002 - 2 January 2012.

The approach follows along the lines of Litterman and Scheinkman (1988) and in that connection we will relate the approach to the Heath, Jarrow and Morton framework. We show in that connection that the PCA method can be thought of as a tool for specifying/determining the spot rate volatility structure using a non-parametric approach.

In the second and by far the largest part of the paper we will turn our attention to Risk-Models. The reason being that Risk Measures have three very important roles within a modern financial institution:

1. It allows risky positions to be directly compared and aggregrated
2. It is a measure of the economic or equity capital required to support a given level of risk activities
3. It helps management to make the returns from a diverse risky business

---

[1] I thank Kostas Giannopoulos for comments to the GARCH estimation and VaR in general.

directly comparable on a risk adjusted basis

Our approach to the calculation of VaR/ETL is a simulation based methodology which relies on the scenario simulation framework of Jamshidian and Zhu (1997).

The general idea behind the scenario simulation procedure is to limit the number of portfolio evaluations by using the factor loadings derived in the first part of the paper and then specify particular intervals for the Monte Carlo simulated random numbers and assign appropriate probabilities to these intervals (states).

We find that the scenario simulation procedure is computational efficient, because we with a limited number of states are capable of deriving robust approximations of the probability distribution. Compared to Monte Carlo simulation another important feature with the scenario simulation procedure is that we have more control over the tails of the distribution - which for Risk models is important.

The paper is organized as follows: In section 2 we will specify the relationship between the shift function and price sensitivities in a fairly general way. After that in section 3 we will specify a multi-factor term structure model in the Heath, Jarrow and Morton framework and show how the volatility structure is related to the shift function.

Section 4 and 5 will be focusing on the estimation of the non-parametric volatility structure (the shift function) using PCA. After that, in section 6, we will compare price sensitivities in the traditional one-factor duration model with price sensitivities derived from the empirical 4-factor yield-curve model.

The rest of the paper (section 7) will concentrate on scenario simulation and VaR/ETL. We will here start with a short introduction to VaR/ETL and at the same time discuss some of the different approaches that has been proposed in the litterature.

Next we will turn our attention to a practical example using a simple portfolio of government bonds. For that portfolio we will compare the scenario simulation model with the full Monte Carlo simulation procedure.

The promising results we obtain here leads us to address the problem of Risk-Calculations for non-linear securities. More precisely we turn our attention to VaR/ETL for Danish MBBs, with as far as we are aware of, only has been considered by Jacobsen (1996). We conclude here that the methodology is both efficient and feasible to use for large portfolios of non-linear securities - because even for complex instruments like Danish MBBs the computational burden is acceptable.

2. The traditional approach - Duration models

The traditional approach to calculating the sensitivities of interest-rate-contingent claims as a function of changes in the initial term structure is to apply a fairly basic assumption - namely that term structure movements only appear as additive shifts to the initial term structure.

The price of a coupon bond can generally be expressed as follows:

$$P^{k}(t,T) = \sum_{j=1}^{n} F_j e^{-R(t,T_j)(T_j - t)} \tag{1}$$

Where R(t,T) is the initial term structure and F_j is the j'th cash flow of the bond.

An expression of the marginal change in this bond, assuming that the term structure movements are defined by a shift function S(t,T), can be formulated as follows[2]:

$$\begin{aligned} dP^{k}(t,T) &= \frac{dP^{k}(t,T)}{d\tau}\Delta\tau + \frac{dP^{k}(t,T)}{d\lambda}\lambda + \frac{1}{2}\frac{d^{2}P^{k}(t,T)}{d\lambda^{2}}\lambda^{2} \\ &= -\Theta_{o}^{k}\Delta\tau - k^{k}(S(t,T))\lambda + \frac{1}{2}Q^{k}(S(t,T))\lambda^{2} \qquad \text{for } \lambda = 1 \end{aligned} \tag{2}$$

Where $k^k(S(t,T))$ is the price risk, $Q^k(S(t,T))$ is the curvature, Θ_o^k is the time sensitivity of the bond, i.e. the bond theta, $\lambda S(t,T)$ is the size of the impact on the initial term structure, τ is the time to maturity, $\Delta\tau$ is the chosen time-change unit for which the time sensitivity is desired to be computed, and S(t,T) is a specific term structure shift function - as of now left unspecified.

2.1 The relationship between the initial term structure and the shift function

Initially the price of a bond is given by formula 1. Following the impact on the initial term structure caused by the shift function S(t,T) the price of this bond can be formulated as:

$$\tilde{P}^{k}(t,T) = \sum_{j=1}^{n} F_j e^{-[R(t,T_j) + S(t,T_j)](T_j - t)} \tag{3}$$

Now define a function f(λ) as follows:

$$f(\lambda) = \sum_{j=1}^{n} F_j e^{-[R(t,T_j) + \lambda S(t,T_j)](T_j - t)} - \sum_{j=1}^{n} F_j e^{-R(t,T_j)(T_j - t)} \tag{4}$$

Where f(λ) is an expression of the change in the bond price when the initial term structure changes from R(t,T) to λS(t,T). In addition, as for λ = 1, $f(\lambda) = \tilde{P}^{k}(t,T) - P^{k}(t,T)$, implies that we want to determine the second-order approximation to f(λ) in point λ = 1.

The f(λ) function can be written as a second-order Taylor expansion around λ = 0, as follows:

---

[2] If we disregard yield-curve changes of higher order than 2 and time changes of higher order than 1.

$$f(\lambda) = \frac{df(0)}{d\lambda}\lambda + \frac{1}{2}\frac{d^2f(0)}{d\lambda^2}\lambda^2 \qquad \textbf{(5)}$$

As f(0) = 0 (see formula 3).

Based on formula 4, it can be deduced that a computation of $\frac{df(0)}{d\lambda}\lambda$ for $\lambda = 1$, yields the desired first-order approximation, and the corresponding argumentation can be used for the second-order approximation.

Bearing this in mind, the price risk ($k^k(S(t,T))$) and the curvature ($Q^k(S(t,T))$) can be formulated as follows:

$$\begin{gathered} k^k(S(t,T)) = -\sum_{j=1}^{n} F_j(T_j - t)S(t,T_j)e^{-R(t,T_j)(T_j - t)} \\ og \\ Q^k(S(t,T)) = \sum_{j=1}^{n} F_j(T_j - t)^2 S(t,T_j)^2 e^{-R(t,T_j)(T_j - t)} \end{gathered} \qquad \textbf{(6)}$$

Where the traditional approach is to let S(t,T) be equal to 0.01, i.e. an additive shift in the term structure of 1%.

Now it can be seen that S(t,T) = 0.01 causes the sensitivities stated in formulas 2 to degenerate into the traditional key figures in a duration/convexity approach, however extended by including the sensitivity to a shortening of maturity[3].

Thus, in this connection it can be deduced that once the shift function is known (is determined/specified), it is possible to calculate relevant and consistent key figures.

3. A multi-factor model for the bond return

Let us first recall some properties about the HJM framework as this will be our starting point when specifying a bond return model.

The dynamic in the zero-coupon bond-prices P(t,T), for $t < T \leq \tau$, is assumed to be governed by an Ito process under the risk-neutral martingale measure Q:

---

[3] In the rest of paper I will however only focus on the sensitivity that is a function of the shift-function - thus I will disregard time-sensitivities.

$$\frac{dP(t,T)}{P(t,T)} = rdt - \sum_{i=1}^{m} \sigma_P(t,T;i) d\tilde{W}_i(t) \quad for \quad \tilde{W}_i(t) = W_i(t) - \Gamma_i(t) \tag{7}$$

Where we have that P(0,T) is known for all T and P(T,T) = 1 for all T. Furthermore r is the risk-free interest rate, and $\sigma_P(t,T;i)$ represents the bond-price volatility, which can be associated with the i'th Wiener process, where $\tilde{W}_i$ is a Wiener process on (Ω,F,Q), for dQ = ρdP and ρ is the Radon-Nikodym derivative. We also have that $\Gamma_i(t)$ represents the market-price of risk that can be associated with the i'th Wiener process.

In order to derive the following results it is not necessary to assume that $\sigma_p(t,T;i)$ for i = {1,2,...m} is deterministic. It is sufficient to assume that $\sigma_p(t,T;i)$ is bounded, and its derivatives (which are assumed to exist) are bounded.

Formula 7 can be rewritten as:

$$d\ln P(t,T) = \left[r - \frac{1}{2}\sum_{i=1}^{m} \sigma_P^2(t,T;i) \right] dt - \sum_{i=1}^{m} \sigma_P(t,T;i) d\tilde{W}_i(t) \tag{8}$$

The solution to this process can be expressed as:

$$\ln P(t,T) = \ln P(0,T) + \int_0^t \left[r(s) - \frac{1}{2}\sum_{i=1}^{m} \sigma_P^2(s,T;i) \right] ds - \sum_{i=1}^{m} \int_0^t \sigma_P(s,T;i) d\tilde{W}_i(s) \tag{9}$$

and:

$$0 = \ln P(0,t) + \int_0^t \left[r(s) - \frac{1}{2}\sum_{i=1}^{m} \sigma_P^2(s,t;i) \right] ds - \sum_{i=1}^{m} \int_0^t \sigma_P(s,t;i) d\tilde{W}_i(s) \tag{10}$$

Where equation 10 follows from the horizon condition that P(T,T) = 1.

The drift in the process for the bond-price - r in formula 7 - can now be eliminated if we consider the difference between the process defined in formula 9 and the process that follows from the horizon condition (formula 10), ie:

$$\ln P(t,T) = \ln\frac{P(0,T)}{P(0,t)} - \sum_{i=1}^{m}\int_0^t \frac{1}{2}[\sigma_P^2(s,T;i) - \sigma_P^2(s,t;i)]ds - \sum_{i=1}^{m}\int_0^t [\sigma_P(s,T;i) - \sigma_P(s,t;i)]d\tilde{W}_i(s) \qquad \textbf{(11)}$$

The process for the forward-rates, can also be derived - namely by using formula 11, ie:

$$r^F(t,T) = r^F(0,T) + \sum_{i=1}^{m}\int_0^t \sigma^F(s,T;i)\sigma_P(s,T;i)ds + \sum_{i=1}^{m}\int_0^t \sigma^F(s,T;i)d\tilde{W}_i(s) \qquad \textbf{(12)}$$

Where $\sigma^F(t,T;i)$ is defined as $\frac{\partial\sigma_P(t,T;i)}{\partial T}$, and can be recognized as being a measure for the forward rate volatility.

We furthermore assume that the volatility function satisfies the usual identification hypothesis, that $\begin{pmatrix}\sigma^F(t,T_1)\\ \sigma^F(t,T_2)\\ .\\ .\\ \sigma^F(t,T_m)\end{pmatrix}$ is non singular for any t and any unique set of maturities $[T_1,T_2,...T_m]$.

The spot-rate process is easily found from the process for the forward rate, ie:

$$r(t) = r^F(0,t) + \sum_{i=1}^{m}\int_0^t \sigma^F(s,t;i)\sigma_P(s,t;i)ds + \sum_{i=1}^{m}\int_0^t \sigma^F(s,t;i)d\tilde{W}_i(s) \qquad \textbf{(13)}$$

That is, the spot-rate process is identical to the forward-rate process, except that in formula 13 we have simultanous variation in the time and maturity arguments.

It may be seen from formulas 12 and 13 that the process for the interest rates is fully defined by the initial yield-curve and the volatility structure, which is precisely the main result of the Heath, Jarrow and Morton (1991) model framework.

From this we can deduce that there are the following relationships between the bond price volatility structure and the shift-function:

$$S(t,T) = \sum_{i=1}^{m} \frac{\sigma_P(t,T;i)}{T - t} = \sum_{i=1}^{m} \sigma_R(t,T;i) \tag{14}$$

It can now be seen that the shift function is identical to the volatility structure of the spot rate structure. In this connection, it can therefore be deduced that the traditional 1-factor duration model given by additive impacts on the term structure of interest rates can be formulated as follows:

$$S(t,T) = \sigma_R \tag{15}$$

Where this formulation of the spot rate volatility structure is identical to the continuous-time version of the Ho and Lee model.

This result can be seen to clash with Ingersoll, Skelton and Weil (1978), who postulate that under the no arbitrage assumption the term structure of interest rates cannot change additively without the term structure of interest rates being flat. Thus, it can be concluded that this assertion is not valid, where this is also shown by Bierwag (1987b), however in a completely different framework

In determining the shift function S(t,T) there are, in principle, two methods, as was also shown by Heath, Jarrow and Morton (1990), namely to either make a pre-defined specification of the functional shape of the volatility structures (the implicit method)[4], or to estimate them historically - that is determine the volatility structure empirically[5].

The first method is the principle used in option theory and will not be the approach used in this paper; instead I intend to use historical data to determine the volatility structures - that is a non-parametric approach.

4. Principal Component Analysis (PCA)

It is now assumed that the shift function is to be determined by considering the historically observed movements in the term structure of interest rates; in addition, it is assumed that the estimation length (i.e. the number of times the term structure of interest rates is to be observed in the estimation) is equal to L, where l = {1,2,....,L} is the l'th observed variation of the term structure of interest rates. Furthermore, only a finite number of points on the term structure of interest rates are used. The number of points/interest rates is P, where p = {1,2,...,P} is the p'th interest rate and where the terms to maturity of these P interest rates cover the entire maturity spectrum from t-T.

---

[4] See for example Amin and Morton (1993).

[5] This is not to be understood in the sense that it is not possible to estimate the parameters that describe the functional shape of the volatility structures by using historical data, as this is of course possible; see Heath, Jarrow and Morton (1990) "Contingent Claim Valuation with a Random Evolution of Interest Rates".

It is now assumed that the shift function S(t,T) can be written as a linear factor model, as follows[6]:

$$S(t,T) = \sum_{i=1}^{m} v_p(t,T;i)dF_l(t,T;i) + \varepsilon_p(t,T) \tag{16}$$

Where $v_p(t,T;i)$ is a function consisting of m-independent risk factors, where a risk factor is defined for each p (i.e. each point/interest rate used in the estimation); in matrix form $v_p(t,T;i)$ can therefore be understood as being defined by P-rows and m-columns. In addition $dF_l(t,T)$ represents the change in the risk factors $v_p(t,T)$ across the entire estimation length L, i.e. in matrix form $dF_l(t,T)$ is therefore defined as a matrix with m-rows and L-columns. Furthermore, $\varepsilon_p(t,T)$ is an error element that is assumed to be normally distributed, so that $\varepsilon_p(t,T) \sim N(0,V)$ where it is assumed that the error elements are independent of the interest rate variation, so that V is a diagonal variance matrix, which can be formulated as follows in matrix form:

$$\begin{gathered} \varepsilon_p(t,T) \sim N(0,V) \\ for \\ V = \begin{bmatrix} V_1 & 0 & 0 & 0 & 0 \\ 0 & V_2 & 0 & 0 & 0 \\ 0 & 0 & V_3 & 0 & 0 \\ . & . & . & . & . \\ 0 & 0 & 0 & 0 & V_P \end{bmatrix} \end{gathered} \tag{17}$$

That is, $\varepsilon_p(t,T)$ is a PxP matrix. Where the more V deviates from the 0 matrix, the less correctly will the model describe the original data material.

The model constructed here can also be formulated as a function of the bond yield, as follows:

$$\begin{gathered} \ln\left[\frac{P(t+\alpha,\tau+\alpha)}{P(t,\tau)}\right] = \sum_{i=1}^{m} v_p(t,\tau;i)dF_l(t,\tau;i) + \omega_p(t,\tau) \\ for\ all\ \tau\ \varepsilon\ [T_j,T],\ for\ j\ \varepsilon\ [1,2....,P-1] \\ and\ \alpha\ \varepsilon\ \Delta l,\ for\ l\ \varepsilon\ [1,2,....,L] \end{gathered} \tag{18}$$

Where it can be seen that this model formulation is identical to the Ross (1976) APT model (The Arbitrage Pricing Theory), with the quantity $v_p(t,\tau;i)$ being better known as the i'th factor loading.

The difference between how we have formulated the APT model and the traditional APT model is that we have no factor-independent Rate-Of-Return (apart from the residual element).

---

[6] This is of course not the only formulation of the shift function imaginable; for a brief review of models related to this formulation, please refer to Appendix A.

We focus namely on modelling/estimating the total Rate-Of-Return of the bonds, and not the excess rate-of-return, where the excess rate-of-return is defined as the Rate-Of-Return that exceeds the risk-free interest rate.

In addition, it should be mentioned at this point that we have not made any kind of explicit definition of the individual factors F; the explanation is that this analysis will not focus on estimating spot rate volatility structures that have a specific pattern, but on the other hand on identifying the number of linear independent parameters that explain the variation in the term structure of interest rates historically.

In connection with the estimation of the model, we have used formula 16 as my starting point. The model is estimated by constructing a matrix of historical variations in the term structure of interest rates, after which the loading matrix $v_p(t,\tau;i)$ and the factor values $dF_l(t,\tau;i)$ are estimated using the Principal Component Analysis (PCA).

The underlying idea of PCA is to analyze the correlation structure (the correlation matrix), that is, the starting point is to find the correlation matrix on the basis of the matrix of historical interest changes, after which it is standardized in such a way that the diagonal consists of 1's, which means that the dispersion matrix is equal to its correlation matrix. The principal factor solution to the estimation problem is then:

$$K\Lambda^{\frac{1}{2}} = [\sqrt{\lambda_1}k_1, \sqrt{\lambda_2}k_2, \ldots\ldots, \sqrt{\lambda_m}k_m] \qquad \textbf{(19)}$$

where $[\lambda_1 \geq \lambda_2 \geq \ldots . \geq \lambda_m]$ *and* $[k_1, k_2, \ldots\ldots, k_m]$ respectively repre
correlation matrix and the associated standardized orthogonal eigenvectors. In addition, K is a matrix containing the eigenvectors by columns and Λ is a diagonal matrix with the eigenvalues in the diagonal. As the correlation matrix is an estimate of the squared loading matrix, the squared loading matrix is defined by formula 19.

In this formulation it is assumed that the number of factors is known in advance, namely m. If this is not the case, the factors that are attributable to the highest eigenvalues are selected until there is a satisfactory description of the data material, where Kaiser's theorem suggests that all the factors that have an eigenvalue higher than 1 are to be chosen.

As a factor loading is a vector of correlation coefficients, this means that the best interpretable factor loadings are achieved when they are either close to 0 (zero) or 1[8]. This can be achieved by rotating the factors, the main rule being that a new estimate for the loading matrix can be obtained - without changing the explanatory degree at row level, or for that matter at column

---

[7] In this connection we have used Jacobi's algorithm to solve the eigenvalue problem.

[8] It can therefore also be concluded that the individual loadings relate to the explanatory degree, in fact, the explanatory degree of the individual factors is given by the squared loadings.

level - by multiplying the principal factor solution by an orthogonal matrix. In this connection we have chosen to use Kaiser's Varimax method for rotating the factors.

The generation of the factor values is the last point missing. Where this can be done in the following way $F = D^{T}(s)L[L^{T}L]^{-1}$, for L being the loading matrix and D(s) being the o interest rate variation matrix (with the number of rows equal to P and the number of columns equal to L), however in a standardized form, i.e. each column has a mean value equal to 0 (zero) and a standard deviation of 1[9]. In conclusion, it should be stressed that these considerations regarding PCA are taken from Harman (1967).

Lastly, with respect to the model defined in formulas 16 and 17, the individual factor loadings are assumed to be constant, whereas the factor values (factor scores) are time-dependent.

This indicates that - assuming that the residual element is negligible - the factor values can be understood as being a time-dependent weighting parameter which, in principle, can be fixed at 1% when risk parameters are to be calculated. Where this means that the shift function can be formulated as follows:

$$S(t,T) \approx \sum_{i=1}^{m} \frac{v_p(t,T;i)}{100} \tag{20}$$

Which results in term structure shifts being measured in terms of standard deviations, which is also a relevant calculation unit when bearing formula 14 in mind. The determination of risk parameters in this multi-factor model will be discussed in more detail in section 6.

5. Estimation of the Volatility Structure

The analysis period has been selected to cover every Wednesday over the period 2 January 2002 - 2 January 2012, based on the yield-curve estimated using the methodology of B-Splines, see for example Webber and James (2000).

We are using the following vector of maturity dates as our key-maturity dates: [0.083,0.25,0.5,1,2,3,4,5,7.5,10,12.5,15,20,25,30] - that is P = 15.

In the table below are shown the ordered eigenvalues for each of the 15 potential factors.

Eigenvalue	Degree of Explanation
9,66	64,37
2,04	78,00
1,47	87,81
1,04	94,74

---

[9] In the construction of this expression of the factor values, the residual matrix has been disregarded, as will appear. This means that it is implicitly assumed that so many factors are being selected that the entire variation in the data material has been described.

0,27	96,56
0,23	98,09
0,10	98,76
0,07	99,23
0,05	99,54
0,03	99,72
0,02	99,87
0,01	99,93
0,01	99,99
0,00	99,99
0,00	100,00

According to Kaiser's theorem one should select all the factors that have an eigenvalue higher than 1, which in our case means the first 4 eigenvalues. Selecting 4-factors will explain approximate 95% of the total variation in the yield-curve over the analysed time-period - this is assumed to be sufficient.

Figure 1 below show the estimated factor loadings (for the 4 most important factors) before the varimax rotation and figure 2 show the factor loadings after we have performed the varimax rotation:

Figure 1 **Estimated Factor Loadings - Original (2. January 2002 - 2. January 2012)**

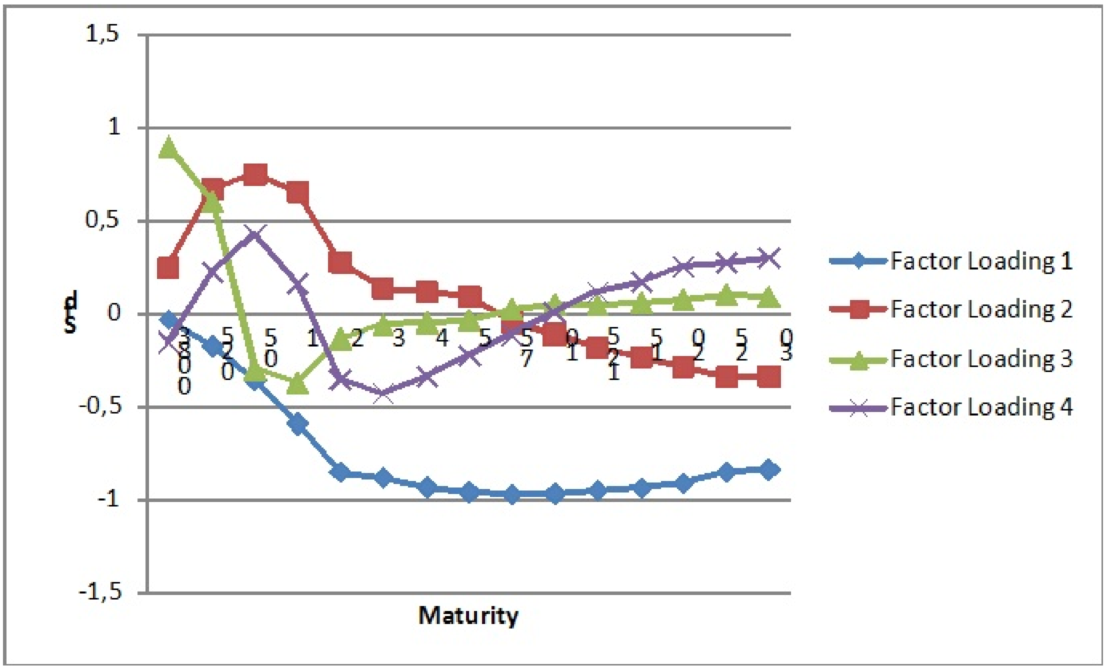
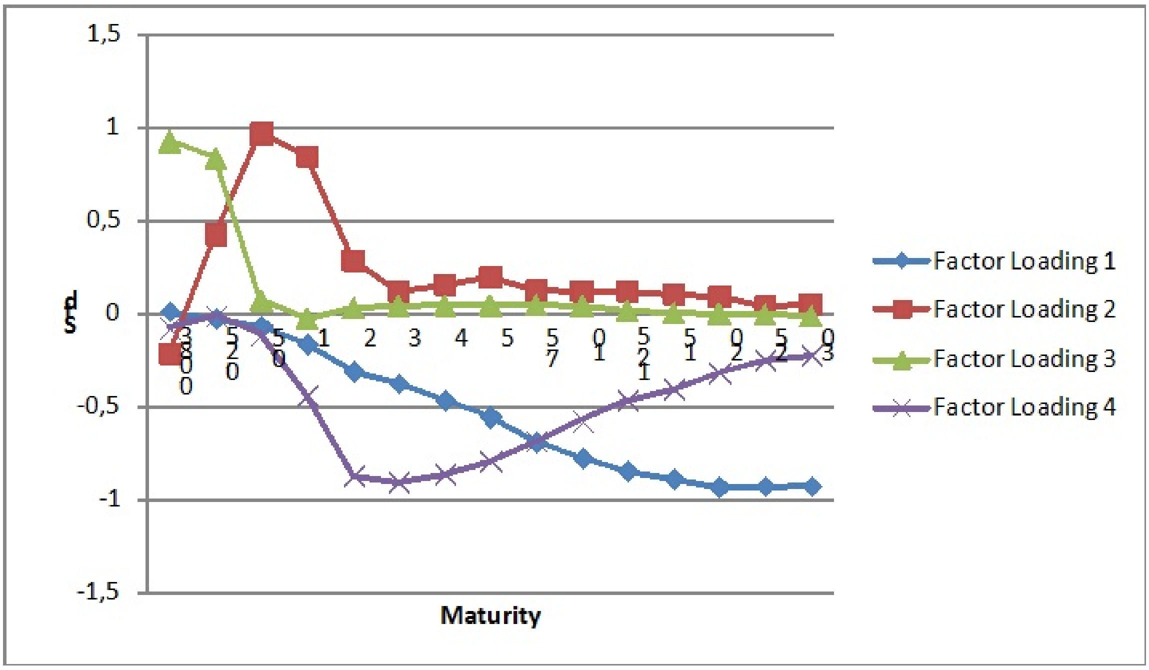


Figure 2 **Estimated Factor Loading - after Varimax Rotation (2. January 2002 - 2. January 2012)**

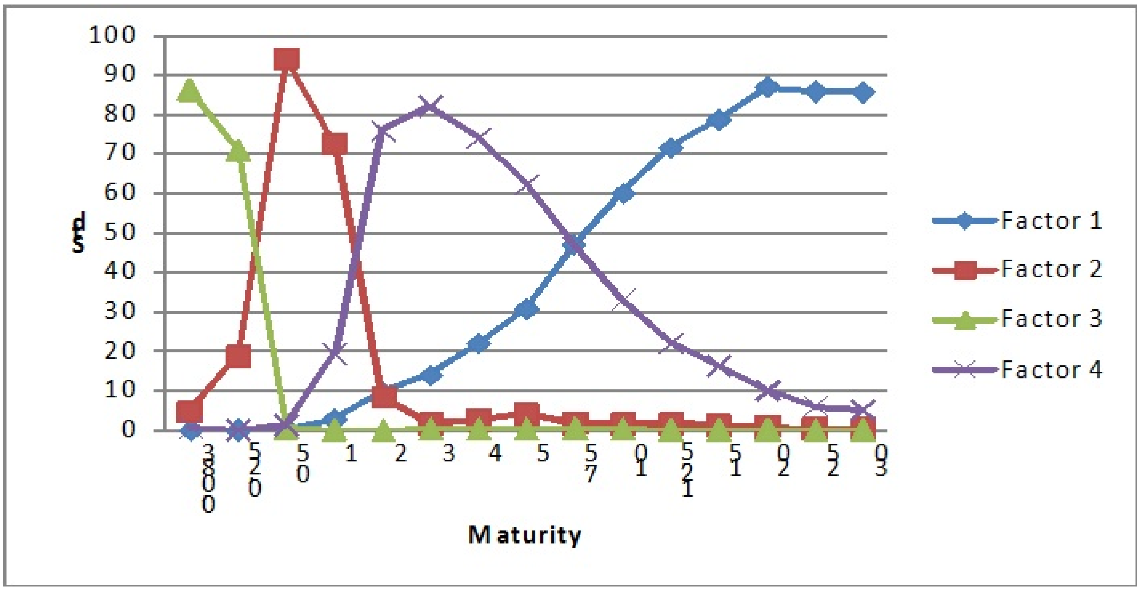


In figure 3 below we have shown the degree of explanation for each of the factors[10]:

Figure 3 **Degree of Explanation (2. January 2002 - 2. January 2012)**

---

[10] From now on when we use the phrase factor loading we refer to the factor loadings that are obtained after the varimax rotation.



These 4 factors can be interpreted as: a slope factor, a short-curvature factor, a short factor and a curvature factor. It is here interesting to mention, that no matter how we slice and dice our data, this meaning that if we for example select the period from 2. January 2002 to 2. January 2008 we still get the same 4 factors.

As can be seen in the graphs above, the result we obtain are somewhat different from what normally has been reported in the litterature, for example Dahl (1989) using Danish data, Litterman and Scheinkmann (1988) and Garbade (1986) using American data, and Caverhill and Strickland (1992) using English data, which conclude that empirically 3 factors exist that describe the dynamics of the term structure of interest rates.

Lord and Pelsser (2006) concluded the following in their study of the fact that traditionally 3-factors are observed, namely: A Level Factor, a Slot Factor and a Curvature Factor:

"...we analysed the so-called level, slope and curvature pattern one frequently observes when conducting a principal components analysis of term structure data. A partial description of the pattern is the number of sign changes of the first three factors, respectively zero, one and two. This characterisation enables us to formulate sufficient conditions for the occurrence of this pattern by means of the theory of total positivity. The conditions can be interpreted as conditions on the level, slope and curvature of the correlation surface. In essence, the conditions roughly state that if correlations are positive, the correlation curves are flatter and less curved for larger tenors, and steeper and more curved for shorter tenors, the observed pattern will occur.

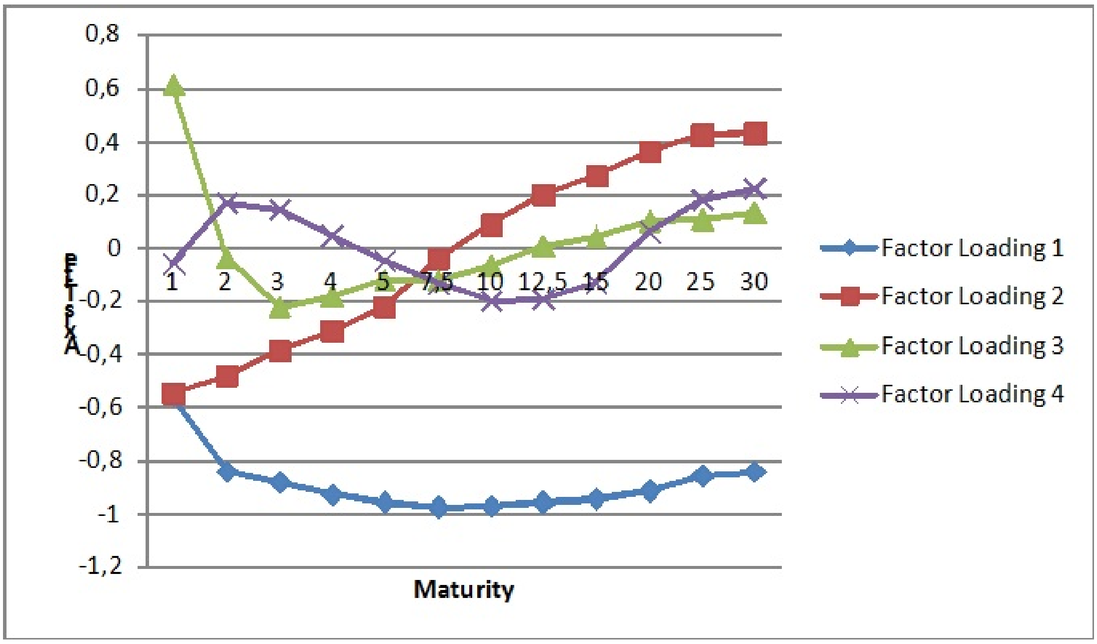
As a by-product of these theorems, we prove that if the correlation matrix is a Green's or Schoenmakers-Coffey matrix, level, slope and curvature is guaranteed. An unproven conjecture at the end of this paper demonstrates that at least slope seems to be caused by two stylised empirical within term structures: the correlation between two contracts or rates *decreases as a function of the difference in tenor between both contracts, and the correlation between two equidistant contracts or rates increases as the tenor of both contracts increases.*

...we can conclude that the level, slope and curvature pattern is part fact, and part artefact. It is caused both by the order and positive correlations present in term structures (fact), as well as by the orthogonality of the factors and the smooth input we use to estimate our correlations (artefact).”

So the reason for the difference we observe here, is probably given by the fact that we are employing a B-Spline estimation algorithm (which are very flexible and able to capture both the long end of the yield-curve and the short end simultanously), while the other studies use (if it is mentioned at all) less flexible functional forms for the yield-curve, like for example the Nelson and Siegel (1988) model or extensions hereof like the Svensson (1994) model.

To investigate that observation we have re-done the estimations of the PCA-model for the same period 2. January 2002-2. January 2012, but now for the following set of interest-rate maturities: [1,2,3,4,5,7.5,10,12.5,15,20,25,30]. If our observation is correct we would expect to get 3 factors, namely: A Level Factor, a Slot Factor and a Curvature Factor. The result are shown below:

Estimated Factor Loadings (No Rotation) and a subset of maturities (2. January 2002 - 2. January 2012)



The first 5 eigenvalues are:

Eigenvalue	Degree of Explanation
9,52434	79,36953
1,44311	91,39544

0,54039	95,89866
0,25575	98,02987
0,09443	98,81678

From the above we can actually see that now the results obtained are identical with what is normal reported in the litterature - namely 3-factors that explan more than 95% of the variation in the yield-curve which can be interpreted as: A Level Factor, a Slope Factor and a Curvature Factor.

6. Measuring of Risk in a multi-factor shift function model

The shift function in this 4-factor term structure model can now be formulated as follows:

$$S(t,T) = \frac{v_1 dw_1 + v_2 dw_2 + v_3 dw_3 + v_4 dw_4}{100}; \quad for\ dw_1 = dw_2 = dw_3 = dw_4 = 1 \quad \textbf{(21)}$$

where v_m, for m = {1,2,3,4} is the vector of factor loadings and dw_m is the vector of factor scores.

In this connection, the vector sensitivities can be formulated as follows for the coupon bond $P^k(t,T)$:

$$\frac{\partial P^k(t,T)}{\partial dw_m} \frac{1}{P^k(t,T)} = -\frac{\sum_{j=1}^{n} F_j(T_j - t) v_m(j) e^{-R(t,T_j)(T_j - t)}}{P^k(t,T)}; \quad for\ m\ \varepsilon\ [1,2,3,4] \quad \textbf{(22)}$$

where this can be regarded as the factor duration.

The factor convexity has the following form:

$$\frac{\partial^2 P^k(t,T)}{\partial dw_m^2} \frac{1}{P^k(t,T)} = \frac{\sum_{j=1}^{n} F_j(T_j - t)^2 v_m(j)^2 e^{-R(t,T_j)(T_j - t)}}{P^k(t,T)}; \quad for\ m\ \varepsilon\ [1,2,3,4] \quad \textbf{(23)}$$

These two equations deserve a few comments.

To fully understand these relations we first note that if we disregard v_m, for each m, then the equations degenerate into the portion of returns of the bond which result from a unit change (a standard deviation, S(t,T) = 1%) in the whole yield-curve. The factor model on the other hand tells us that a unit change to the factor does not change the yield-curve by one percent - but by v_m percent.

From this we can deduce that in order to find the total impact on bond returns from a factor change, we need to scale by a weight that is exactly equal to the appropriate factor loading.

6.1 A practical example - risk factors in the 4-factor term structure model compared with the risk factors in the traditional 1-factor duration model

In order to illustrate the difference between the traditional approach in calculating price sensitivities (i.e. S(t,T) = 0.01), and the price sensitivities generated by this 4-factor model the table below shows the risk calculated using modified duration and the factor duration for a wide range of Danish Government bonds.

Table 1: Duration and Factor Sensitivities (2. January 2012)

Isin	Instrument-Name	Mod. Duration	Factor 1 Dur	Factor 2 Dur	Factor 3 Dur	Factor 4 Dur
DK0009922593	4 Ink st 2012	-0,869	-0,119	0,768	-0,001	-0,307
DK0009920894	5 Ink st 2013	-1,824	-0,524	0,684	0,043	-1,466
DK0009922833	2 Ink st 2014	-2,812	-1,020	0,432	0,121	-2,523
DK0009921439	4 Ink st 2015	-3,656	-1,632	0,606	0,171	-3,151
DK0009922759	2.5 Ink st 2016	-4,636	-2,477	0,909	0,220	-3,705
DK0009902728	4 Ink st 2017	-2,800	-1,242	0,667	0,132	-2,217
DK0009921942	4 Ink st 2017	-5,352	-3,126	0,964	0,256	-4,052
DK0009922403	4 Ink st 2019	-6,930	-4,685	0,965	0,334	-4,708
DK0009922676	3 Ink st 2021	-8,673	-6,484	1,149	0,383	-5,198
DK0009918138	7 Ink st 2024	-9,524	-7,577	1,263	0,236	-4,941
DK0009922320	4.5 Ink st 2039	-18,076	-16,151	1,268	-0,002	-5,517

The most interesting to note in this table is the fact that the degree of sensitivity a given bond has to a particular (risk) factor has got nothing to do with the importance of this (risk) factor - this issue will be discussed later in section 7.4.

One can from the table see that this 4-factor model measure the risk at the relevant segments/parts of the yield-curve where the risks of each securities are more logically attributable than what is the case for the traditional 1-factor duration model. For example the short bullet bond 4 Ink st 2012 is more sensitive to factor 2 than the other 3-factors as the maturity date of that bond occur at a time where factor 2 has more influence than the other 3-factors, see Figure 2.

This ends the first part of the paper - namely determining the number of factors that drive the evolution in the yield-curve using an empirical approach.

In the next section we will explain how we - for the current analysis - have defined to setup our Risk-Model framework.

7. Risk-Models - a Short Survey

VaR is probably the most important development in risk management. This methodology has been specially designed to measure and aggregate diverse risky positions across an entire institution using a common conceptual framework. Even though these measures come under difference disguises e.g. Banker Trusts Capital at Risk (CaR), J.P. Morgans Value at Risk (VaR) and Daily Earnings at Risk (DeaR), they are all based on the same foundation. Even though different institutions has come up with their own names the one that seems to be most commonly used is VaR - which is the name we will be using here.

Definition 1:

Risk Capital is defined as the maximum possible loss for a given position (or portfolio) within a known confidence interval over a specific time horizon.

VaR plays three important roles within a modern financial institution:

- It allows risky positions to be directly compared and aggregrated
- It is a measure of the economic or equity capital required to support a given level of risk activities
- It helps management to make the returns from a diverse risky business directly comparable on a risk adjusted basis

Even though there are some open issues regarding its calculation, VaR is nevertheless a very useful tool for helping management to steer and control diverse risk operation.

The problem with VaR is that there is a variety of different ways to implement the definition of Risk Capital (see Definition 1) each having distinct advantages and weaknesses. In order to get a better grasp of the trade-offs implicit in each method - it is important to understand which kind of components Risk Capital is built around.

Risk Capital comprises of the following two (2) distinct parts:

- The sensitivity of a position's (portfolio's) value to changes in market rates
- The probability distribution of changes in the market rates over a predefined reporting horizon

Given these assumption - VaR is (usually) defined as the maximum loss within the 99% confidence interval over at 10-day time period (Basel II) - for Market Risk calculations[11].

In the following table we have listed the 3 most common methods used when calculating VaR - together with their advantages and disadvantages:

---

[11] Recently there has been talk about changing from the use of VaR to the use of ETL, for more about ETL see below.

Method	**RiskMetric^{TM}**	**Delta-Gamma Methods[12]**	**Simulation Based Methods**
Description	Assume that asset returns are normally distributed, implying linear pay-off profiles and normally distributed portfolio returns[13]: $VaR = \alpha\sqrt{\Delta t}\sqrt{VCV^T}$	Assume that asset returns are normally distributed, but pay-off profiles are approximated by local second order terms	Approximate probability distribution for asset returns based on simulated rate movements, either historically or model based
Advantage	Simplicity.	Simplicity. Captures second order effects.	Captures local and non-local price movements. Takes into account fat tails, skewness and kurtosis. With Monte Carlo simulation, we have flexibility to select a probability distribution. With historical simulation we do not need to infer a probability distribution.
Disadvantage	Assuming normality of returns ignores fat tails, skewness and kurtosis. Ignores higher order moments in sensitivities. Captures only risks of local movements for linear securities. Does not capture risks of non-local movements.	Assuming normality of returns ignores fat tails, skewness and kurtosis. Does not capture risks of non-local movements.	Computationally expensive - for Monte Carlo simulation. With Monte Carlo simulation then we need to select a probability distribution. With historical simulation then we cannot select a probability distribution.

---

[12] This method is due to Wilson (1994).

[13] $\alpha = 2.54$ if we wish to calculate VaR at the 99% confidence level, Δt = unwind period, V = vector of volatilities and C is the correlation matrix.

It is not directly mentioned in this table but in general it is assumed that the expected return is zero (0), ie: $E_{t-1}\left[\frac{P_t - P_{t-1}}{P_{t-1}}\right] = 0$. This assumption is related to the fact that it is the standa measure Risk over a "short" time interval, for longer periods (unwinding periods) one should not employ the zero expected return assumption!

The issue that has been mostly widely discussed in the literature with respect to VaR is how to derive appropriate volatility estimates? - see for example Alexander (1996). We will not here discuss that issue but will address it in section 7.3.

Another very important issue in connection with Risk-Calculations is how to derive a reliable correlation matrix? The problem with estimating and modelling the correlation matrix can be summarized as follows:

- Correlations coefficients are highly unstable and their signs are ambiguous
- If for example we have 15 interest rates we need to keep track of (model) $\prod_{1}^{14} i$ correlation coefficients
- We need a long data period in order to have enough degrees of freedom to estimate a reliable correlation matrix. There is however no guarantee that the resulting matrix satisfy the multivariate properties of the data - we might even encounter a correlation matrix that is not positive definit

These observations have inspired research to reduce the dimensionality - where a common suggested method is PCA - like the one we performed in section 4 and 5 in this paper. The nice property in this context is - as mentioned in section 5 - that even though the correlation coefficients are highly unstable the factor loadings are extremely stable[14].

Barone-Adesi, Bourgoin and Giannopoulos (1997) has suggested an interesting approach to worst case scenarios. This method is based on historical returns combined with a GARCH approach (actually a AGARCH-model) for forecasting purposes. The procedure does not employ the correlation matrix directly as the correlation is embedded indirectly in the historical simulation procedure - which is a very neat property of their method.

7.1 VaR contra ETL

Artzner, Delbaen, Eber and Heath (1999) showed that VaR is not what is termed a coherent risk-measure. Where the definition of a coherent risk-measure is a risk-measure that satisfies:

---

[14] The Factor ARCH approach of Engle, Ng and Rothschild (1990) is also in this spirit - see Christiansen (1998) on Danish data.

1: Monotonicity: $Y \geq X \Rightarrow m(Y) \leq m(X)$
2: Subadditivity: $m(X+Y) \leq m(X) + m(Y)$
3: Positive homogeneity: $m(hX) = hm(X)$ for $h > 0$
4: Translational invariance: $m(X+n) = m(X) - n$, for some certain amount n

Properties 1, 3 and 4 are well-behavedness properties to rule out strange outcomes. The most important property is the sub-additivity property.

The sub-additivity property tell us that a portfolio made up of subportfolios will have a risk that is not more than the sum of the risks of the subportfolios:

- This reflects an expection that when we aggregrate individual risk, they diversify or, at worst, do not increase – the risk of the sum is always less than or equal to the sum of the risks

This sub-addivity property spells trouble for VaR as VaR is not sub-additive - except by posing particulary restrictions on the profit/loss distribution.

One risk-measure candidate which is coherent is ETL (expected tail loss)[15].

ETL is defined as the average of the worst 100(1-α)% of loses, that is:

$$ETL_{\alpha} = \frac{1}{1-\alpha} \int_{\alpha}^{1} q_p \, dp$$

In case the profit/loss distribution is discrete ETL can be calculated as:

$$ETL_{\alpha} = \frac{1}{1-\alpha} \sum_{p=0}^{\alpha} [\text{pth highest loss}] \text{ x } [\text{probability of pth highest loss}]$$

From this we can deduce that ETL is the probability-weighted average of the tail losses – which suggest that ETL can be estimated as an average of ”tail VaRs”.

For the above reason we will be both considering VaR and ETL when calculating risk-measures.

For our implementation of a Risk model we will limit ourselves to domestic interest rate data - but extending to multiple currencies is of course possible, this will however not be treated in this paper.

---

[15] This is sometimes also referred to as Tail-Loss or CvaR (Conditional Value-at-Risk), we however prefer ETL.

Our Risk approach is based on the following basic assumptions:

- We will build our VaR/ETL approach on top of our empirical yield curve dynamics model from section 4 and 5, for elaboration see section 7.1
- We will estimate the volatility structure using the GARCH approach, see section 7.3
- Our approach is a simulation based method which follows along the lines of Jamshidian and Zhu (1997)[16], see section 7.2
- One could say that we combine the methodology of Barone-Adesi, Bourgoin and Giannopoulos (BABG) (1997)[17], with the simulation based approach from Jamshidian and Zhu (JZ) (1997)
- We will impose the restriction that the expected price change in a discount bond is equal to the price change with arises through the combined effect of a shortening of maturity and a movement down the yield-curve. This means that we will assume that the expected yield-curve will be equal to the initial yield-curve

7.1 An Empirical Multi-Factor Model

Before continuing it is worth mentioning that we have the following relationship between dW_i (the Wiener process) and dw_i (the factor scores), for i = [1,2,3,..,P], and P = number of key-rates (maturities) (in our case 12)[18]:

$$dw_i = \frac{1}{\lambda_i}\sum_{p=1}^{P} v_{pi} dW_p \Rightarrow dW_p = \sum_{i=1}^{P} v_{pi} dw_i \qquad \textbf{(24)}$$

this equation arises because $dW_i \approx N(0,1)$ and as we have normalized the correlation matrix[19] then $|v_i|^2 = \lambda_i$ and therefore $dw_i \approx N(0,1)$. That is, the factor scores are assumed to be standardised normal distributed variables.

The equal sign in the last equality in formula 24 becomes an approximation sign for P > 4 - where 4 is the number of factors. As follows we have limited ourselves to a 4-factor model as we are using Kaiser's theorem to select the number of factors, see section 4. Furthermore we assume that the approximation error this gives rise to in equation 24 is negligible.

---

[16] Se also Frey (1998).

[17] One might add that the approach from Hull and White (1998) is quite similar to the BAGS model.

[18] For these derivations, we note that if we consider the factor loadings before the varimax rotation - then the correlation matrix is given by KK^T and $K^TK = \text{diag}(\lambda)$ - if K is the factor loading matrix. It should here be stressed that if we consider the factor loadings after varimax rotation then K^TK will only be equal to $\text{diag}(\lambda)$ if the vector of eigenvalues are being recalculated from the factor loading matrix after rotation.

[19] In the case that we have not normalized the correlation matrix then λ is the vector of eigenvalues for the correlation matrix. This is however only true if we do not perform varimax rotation of the factors.

As mentioned above the volatility parameters can be obtained directly from the estimated eigenvalues - that is however not the approach we will employ here. We will instead use a GARCH(1,1) model to estimate the volatility structure and use this model to derive volatility forecasts that is to be used in connection with our VaR/ETL calculations, for more see section 7.3 and section 7.4.

This means that our model setup will belong to the type of models that normally is referred to as unspanned stochastic volatility (USV) models (see for example Casassus, Collin-Dufresne and Goldstein 2005).

A number of papers have investigated whether the existing models can capture the joint dynamics of term structure and plain-vanilla derivatives such as caps and swaptions. Simple, model-independent evidence based on principal component analysis suggests that there are factors driving Cap and Swaption implied volatilities that do not drive the term structure. Collin-Dufresne and Goldstein 2002 call this feature unspanned stochastic volatility, simplified put, it appears that fixed-income derivatives such as Caps and Swaptions cannot be perfectly replicated by trading (even in a very large number of) the underlying bonds. That is, in contrast to the predictions of standard short-rate models - which assume that bonds does span the bond market.

A model that exhibits USV, is a model where bond prices are insensitive to volatility-risk, and hence cannot be used to hedge volatility-risk. In our case this is true because there is no correlation between our GARCH (1,1) model for volatility and our PCA-model for the yield-curve dynamic.

Using the results from section 3 we can now in abstract form express the dynamic in the yield-curve as follows:

$$dR(t,T) = \mu(T)dt + \sigma(t,T)S(t,T)dW \qquad \textbf{(25)}$$

where S(t,T) is the shift-function matrix which in our case for all practical purposes can be fixed to have 4 columns. Furthermore we have that σ(t,T) is a vector of volatilities. Where as mentioned above the vector of volatilities σ(t,T) will be derived from a GARCH-model.

The drift in formula 25 depends on the expectation hypothesis, which means that the process has been written under the original probability measure - the implication is that the model as it is formulated in formula 25 is not appropriate for pricing purposes.

Using formula 24 and because for our purpose it is more convenient (and appropriate) to assume a lognormal process, we can rewrite formula 25 as:

$$\frac{dR(t,T)}{R(t,T)} = \mu(T)dt + \sigma(t,T)v_1dw_1(t) + \sigma(t,T)v_2dw_2(t) + \sigma(t,T)v_3dw_3(t) + \sigma(t,T)v_4dw_4(t) \textbf{(26)}$$

or alternatively expressed in integral form:

$$R(t,T) = \mu(T)e^{\sigma(t,T)v_1dw_1 + \sigma(t,T)v_2dw_2 + \sigma(t,T)v_3dw_3 + \sigma(t,T)v_4dw_4} \qquad \textbf{(27)}$$

We have assumed that μ(T) (for all T > t) is equal to the initial yield-curve. That is we assume that the expected yield-curve is equal to the actual yield-curve.

This is a more intuitive assumption than using forward rates when our motive is to generate yield-curve scenarios for future dates that are to be used as inputs in various valuation models. The model setup is nevertheless arbitrage free as the uncertainty of each key rate is governed by its own market price of risk - which implicitly are determined through the expectation hypothesis. Again we stress that we are here working under the actual probability measure and not under the risk-neutral probability measure.

For longer time-horizons, the above formulation is not appropriate, due to the fact that we have omitted mean-reversion in the drift specification - extension to a model which has mean-reversion is however straightforward, but for our purpose the above formulation is sufficient.

7.2 Scenario Simulation

In order to simulate yield-curve movements, we can apply the Monte Carlo method to equation 27. If we now assume that we perform N simulations for each of the 4-factors - then the total number of yield curves at a given point in time will be N^4. For N = 100 we will therefore have 100,000,000 future yield-curves which is (more than) adequate in order to generate a probability distribution of the future yield curves. Another approach is of course to perform a Monte Carlo simulation directly in our 4-factor model, however still a fairly high number of simulations is required in a 4-factor model in order to obtain a "precise" probability distribution - especially as we are interested in the tails of the distribution.

In general, it can be said that "brute force" Monte Carlo will require a substantial amount of portfolio evaluation. This approach will for that reason not be of practical use, if firstly the portfolio is fairly large and secondly if the portfolio contains a fair number of non-linear instruments of moderate complexity, like for example Bermudan Swaptions, flexi-caps and/or Mortgage-Backed-Bonds.

The interesting question is - how can we reduce the computational burden while at the same time having a reasonable specification of the future yield-curves probability distribution?

As shown by Jamshidian and Zhu (1997) this is possible. The argumentation is as follows:

Let us suppose that x is a random variable with distribution P(x). In a Monte Carlo method, we simulate N possible outcomes for the random variable x - where each simulated x has the same probability.

However, we have that numbers x_i's, for i = [1,2,...,N], that falls between x_l and x_u is proportional to the probability $P(x_l < x_i \leq x_u)$. From this it follows that we can select a regio (x_l, x_u] and assign a given probability for all numbers that fall inside this region. If we utilize this procedure, it is possible for example to perform 100 simulations for each factor but we only need to perform the portfolio evaluation at a limited number of states - more precisely at the number of states that equal the number of predefined regions.

A good candidate for a probability distribution is the multinomial distribution[20]. If k + 1 states (ordered from 0 to k) are selected then the probability for a given state i is given by the binomial distribution and can be expressed as:

$$Prob(i) = \frac{2^{-k}k!}{i!(k-i)!} \quad for\ i\ \varepsilon\ [0,1,2,...,k] \tag{28}$$

For this distribution we have that there is a distance of $\frac{2}{\sqrt{k}}$ between two adjacent states, and furthermore is the furthest state $\sqrt{k}$ standard deviation away from the center.

Let us now assume that we only need to select nine (9) states for factor 1, five (5) states for factor 2 and three (3) states for factor 3[21]. This gives us all in all 9 x 5 x 3 = 135 scenarios in which we have to evaluate our portfolios - which is independent of the number of Monte Carlo simulations.

Under this assumption we get the following probabilities for each of the states for the 3-factors:

$$\begin{gathered} Factor\ 1: \frac{1}{252}, \frac{1}{32}, \frac{7}{64}, \frac{7}{32}, \frac{35}{128}, \frac{7}{32}, \frac{7}{64}, \frac{1}{32}, \frac{1}{252} \\ Factor\ 2: \frac{1}{16}, \frac{1}{4}, \frac{3}{8}, \frac{1}{4}, \frac{1}{16} \\ Factor\ 3: \frac{1}{4}, \frac{1}{2}, \frac{1}{4} \end{gathered} \tag{29}$$

As the 3-factors are independent we have that their joint probability is defined by the products of the three marginal probabilities.

There has been discussing in the litterature about the robustness of the JZ approcimation, for example Abken (2000), he states the following: "*The outcomes for the nonlinear test portfolios demonstrate that scenario simulation using low- and moderate dimensional discretizations can give "poor" estimates of VaR. Although the discrete distributions used in scenario simulation converge to their continuous distributions, convergence appears to be slow, with irregular oscillations that depend on portfolio characteristics and the correlation structure of the risk factors*." This observation is in line which the oscillations seen in the pricing of long dated options using the Cox, Rox and Rubenstein (1979) model, and something one should be aware of....we will address that issue later.

---

[20] See Cox and Rubenstein (1985).

[21] In Jamshidian and Zhu (1997) they select 7,5 and 3 - but as they mention (their footnote 7) the determination of the appropriate number of states for each factor is an empirical question.

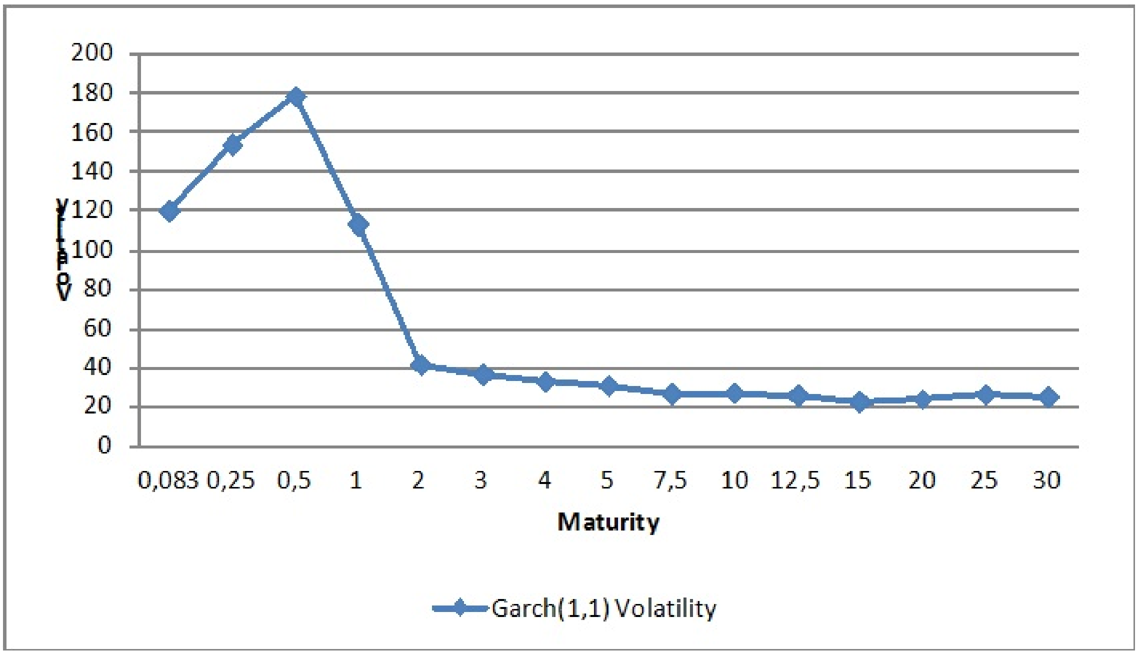
Gibson and Pritsker (2000) have the following comment: "*We show that their method has two serious shortcomings which imply it cannot accurately estimate VaR for some fixed-income portfolios. First, risk factors chosen using principal components analysis will explain the variation in the yield curve, but they may not explain the variation in the portfolio's value. This will be especially problematic for portfolios that are hedged. Second, their discrete distribution of portfolio value can be a poor approximation to the true continuous distribution*." They instead propose a Grid-based Monte Carlo approach on the (derived) risk-factors for the portfolio in question. A similar approach is advocated by Chishti (1999).

We will look into these issues in section 7.4.

7.3 Stochastic Volatility

As mentioned earlier we are using the GARCH(1,1) model to specify the volatility in our multi-factor yield-curve model. To examplify this we have below in figure 4 shown the volatility structure as of the 2 January 2012 that has been determined using the GARCH(1,1) model:

Figure 4 **GARCH(1,1) Volatility Structure (2. January 2012)**



For the interested reader we have in Appendix C briefly discussed the idea behind GARCH and also explained how we have constructed the volatility structure.

7.4 An Illustrative Example

In order to show how the approach works we have constructed the following example:

The portfolio we have selected is the following:

Table 2: **Government Bond Portfolio as of 2. January 2012**

Isin	Instrument-Name	Position	Clean Price	Accrued IR	Dirty Price	Value
DK0009920894	5 Ink st 2013	2500000	109,084	0,697	109,781	2.744.518,03
DK0009902728	4 Ink st 2017	1500000	108,000	2,230	110,230	1.653.442,62
DK0009922320	4.5 Ink st 2039	4000000	149,520	0,627	150,147	6.005.881,97
					Total Value:	10.403.842,62

From table 2 it follows that the value of the portfolio as at 2 January 2012 is 10.403.842,62 kr.

We now wish to determine the distribution of the portfolio for a 10-day horizon using the approach we have specified above for our 4-factor model. Given this distribution we can then derive our VaR estimate give our model setup.

As comparison benchmark we have used Monte Carlo to simulate 250.000 yield-curves for our 10-days horizon for our 4 factor-model and calculated the joint probability distribution.

The purpose of this is in order to investigate the approximation error in the Jamshidian and Zhu (1997) model for calculation VaR - and in that connection address some of the issues that has been stated in litterature, as explained by the end of section 7.2.

Before we show the results we need to address one more issue - namely the building of a yield-curve given the price of zero-coupon bonds at the following fixed maturity dates: [0.083,0.25,0.5,1,2,3,4,5,7.5,10,12.5,15,20,25,30].

For that purpose we are using the maximum smoothness approach of Adams and Deventer (1994), see Appendix C for an elaboration.

The results are stated in the table 3 below:

Table 3: Comparison of scenario simulation vs Monte Carlo simulation

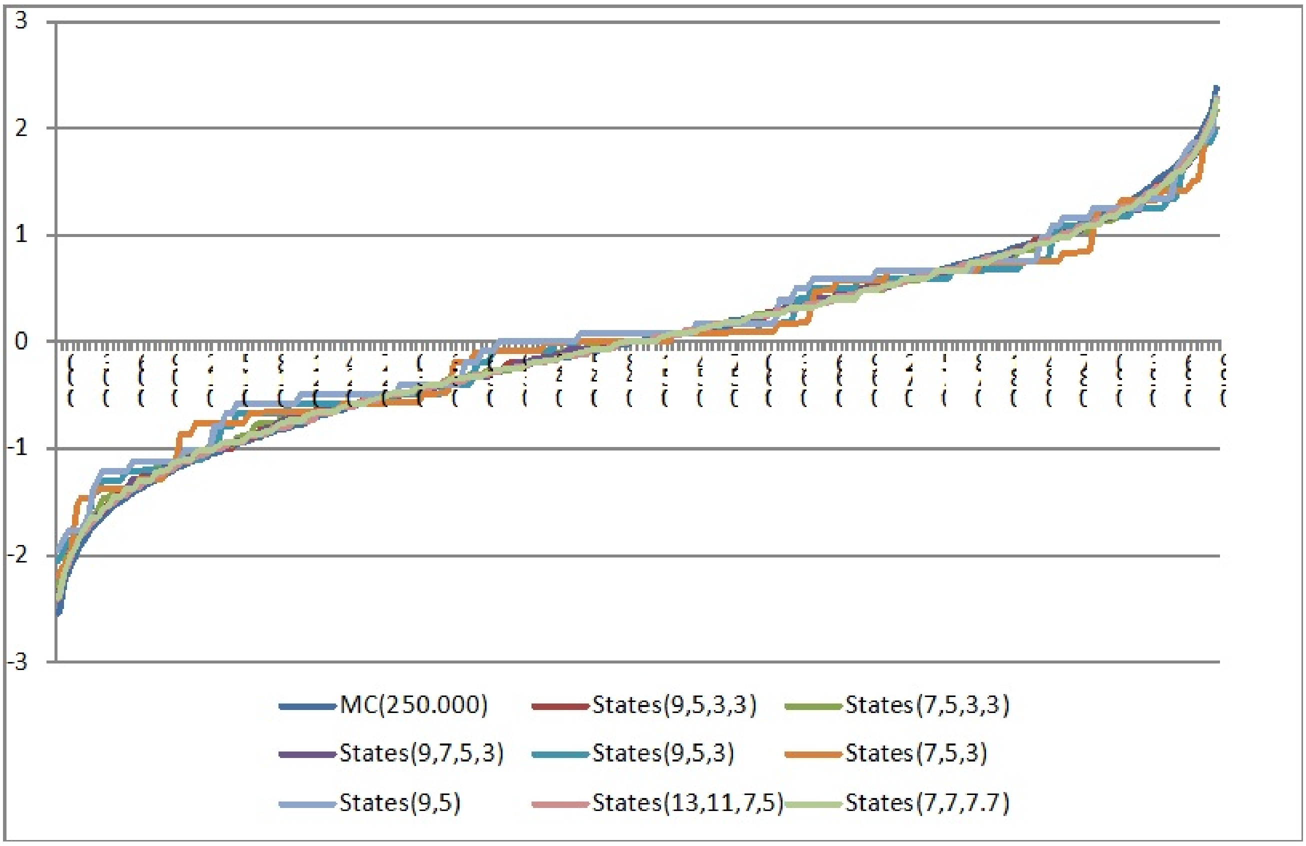
	MC(250.000)	States(9,5,3,3)	States(7,5,3,3)	States(9,7,5,3)	States(9,5,3)
Sample	250000	405	315	945	135
Mean	-0,012	-0,013	-0,013	-0,013	-0,011
Standard Deviation	0,956	1,651	1,450	1,652	1,622
Maximum	4,068	3,005	2,622	3,053	2,645
Minimum	-4,519	-3,299	-2,851	-3,371	-2,858
	States(7,5,3)	States(9,5)	States(13,11,7,5)	States(7,7,7.7)	
Sample	105	45	5005	2401	
Mean	-0,011	-0,011	-0,014	-0,014	
Standard Deviation	1,417	1,634	2,021	1,517	
Maximum	2,257	2,642	3,967	3,062	
Minimum	-2,415	-2,853	-4,539	-3,420	

From the results in table 3 we conclude the following:

- The scenario simulation procedure generates a mean that is very close to the mean derived from the Monte Carlo simulation - this is the case for all the 8 scenario simulations
- The standard deviation in the scenario simulations is in all cases higher than the standard deviation obtained in the Monte Carlo simulation
- We also observe that the range of obtained values (the distance between maximum and minimum) is higher in the Monte Carlo simulation compared to the scenario simulations

Before we start investigating the VaR numbers we have below in figure 5 make a comparison between the True CDF (as obtained from Monte Carlo) and JZ Approximation.

Figure 5: Comparison of the True CDF and the JZ Approximation



The main conclusion from this figure is the following:

- The less number of factors we select the more oscillation we see
- The less number of states for the factors the more oscillation we see
- We can also deduce that the scenario simulation will under estimate the VaR number (this is in line with the results from Gibson and Pristker (2000)) - the question is how much?

Knowing the probability distribution makes it straightforward to calculate VaR/ETL. This is done below in table 4.

Table 4: A comparison of Risk-Measures

	MC(250.000)	States(9,5,3,3)	States(7,5,3,3)	States(9,7,5,3)	States(9,5,3)
VaR(99) 10-Day	-2,297	-2,200	-2,107	-2,215	-1,965
ETL(99) 10-Day	-2,647	-2,665	-2,654	-2,554	-2,978
	States(7,5,3)	States(9,5)	States(13,11,7,5)	States(7,7,7.7)	
VaR(99) 10-Day	-2,114	-1,873	-2,210	-2,191	
ETL(99) 10-Day	-2,631	-5,173	-2,551	-2,453	

From table 4 we find:

- Yes, VaR will be underestimated using the JZ simulation technique - in our case with an error that lies between 3-17%
- A big surprise is that, as a rule of thumb the ETL estimate is significant better than the VaR estimate. The only exception is the much higher ETL for the case: States(9,5)

In interesting question is how can that be?

In the MC method each simulation has the same probability for occuring, that is not the case in the JZ approach - here the probability for any given observation is defined as the product of being in a given state for each of the factors.

What happens with the JZ in case there is not "enough" states is that some of the observations in the tails will be derived from products of probabilities for being in states for one or more factors where the state is not "far enough" away - in terms of standard deviations - from the centre. In connection to this is that we are bound to get quite a few observations in the tails and that the profit loss distribution will oscillate[22].

To conclude how many states we need to get reliable ETL estimates is real hard to say, but it seems like if we; 1) select a suitable number of states for each factor; 2) that all factors are taken into account and 3) that our sample size is significantly larger than 1/(1-quantile) then we will get fairly reliable ETL estimates[23].

To perform a detailed analysis of this interesting subject lies however outside the boundaries of this paper, we will for the current analysis assume this and will for the rest of the paper be using the JZ model with States(7,5,3,3).

The analysis performed in this section gives rise to the following conclusion:

- The scenario simulation procedure is computationally very superior to the

---

[22] We did some more calculations in order to investigate this phenomenon (like for example States(7,5)) and we see the same problem in all the cases - namely "much" to high ETL number.

[23] A similar observation was done in connection with figure 5.

Monte Carlo simulation procedure

- It is easier with scenario simulation to obtain good estimates for ETL than for the point estimate - VaR. Which is good news as ETL also is a better risk-measure than VaR - see section 7
- There is also evidence that it is appropriate to select 7 x 5 x 3 x 3 states - at least in our case

Some closing remarks.

The scenario simulation procedure is so fast for simple (nearly) linear instruments that it rivals the parametric approach.

Actually to derive the return distribution using the scenario simulation for the case States(7,5,3,3) takes less than 0.1 second all in all, where this includes:

- Determining the states
- Calculating the probability for realising a particular yield curve
- Calculating the 315 yield curves for the 15 maturities[24]
- Estimating 315 yield curves using the maximum smoothness approach of Adams and Deventer (1994)
- Calculating the price distribution of each of the 3-bonds in the portfolio
- Calculation the mean, standard deviation etc
- Calculating VaR/ETL for a given confidence interval

From this we conclude that there is no reason why we should be satisfied with a single number from a risk-management point of view - when we can construct the whole probability distribution with very little effort from a computational perspective.

7.5 Scenario Simulation for Non-linear Portfolios

Of course scenario simulation is even more important for non-linear securities - or more precisely to limit the number of calculations.

We will for that reason turn our attention to some interesting non-linear securities - namely Danish Mortgage-Backed-Bonds[25].

Here we will not focus on a portfolio but instead select a few different MBBs and then use the scenario simulation procedure to calculate the probability distribution for these securities.

---

[24] One might note here, in case we had not been looking at such an extreme time-period as we did in the analysis here, the general result would have been that we only would need somewhat around 3-factors and the most appropriate number of states being 9 x 5 x 3 - these results is because of lack of space omitted here.

[25] For a good introduction to Danish Mortgage-Backed-Bonds, see Karner, Kelstrup and Schelde (1998) and Kelstrup, Madsen and Rom-Poulsen (1999).

However, expanding to a portfolio approach is straightforward as the probability distribution of a portfolio is just defined as the value-weighted sum of the probability distributions for each of the securities in the portfolio.

For the purpose of the following discussion we assume that a pricing model is available. We will however here mention a few of the ingredients that are an integrated part of a pricing model for Danish MBBs (and for that matter American MBBs):

- A yield-curve
- A prepayment model for the behaviour of the debtors - we are here using the Madsen (2005) model
- A yield-curve model, we will here be using the same swap-curve model as introduced in section 5
- A model for the volatility structure, we are implying the volatility parameters from traded caps/floors
- A pricing model that uses the first four ingredients to calculate the price - we are here using the Black and Karisinski (1991) model (Log-Normality assumption) , though calibrated using the trinomial approach of Hull and White (1994)
- An OAS-Model, that is a model that takes into account that OAS is a function interest rate changes[26]

As far as we know VaR for Danish MBBs has only been considered by Jacobsen (1996), his method is however completely different than the methodology we employ in this paper.

Let me for that reason briefly mention a few stylized facts about Jacobsens approach:

The main idea in Jacobsen is to construct a delta equivalent hedge portfolio of zero-coupon bonds - that is a 1 order approximation. For that purpose he selects 17 key-rates[27] and using the triangular method of Ho (1992) it is possibly to construct 17 different shifts to the initial yield-curve - where these shifts by definition are constructed in such a manner that the sum of the shifts equals an additive shift to the yield-curve. This delta equivalent hedge portfolio is now being used to calculate VaR measures using the parametric approach suggested in RiskMetric^{TM} for linear securities - that is the delta equivalent cash-flow (DECF) is being treated as a straight bond.

This approach is very simple and the computational burden is acceptable - “just” 17 price calculations - but there are some undesirable features in the methodology:

- Firstly, the DECF is not constant, it depends on all the information that is related to the pricing of MBBs (this is also pointed out by Jacobsen). This means that it is only useful for calculating VaR over short periods and/or limited yield-curve changes

---

[26] It lies outside the boundary of this paper to come into details about our MBB pricing model, and associated models, like our OAS-Model.

[27] The reason being that this is the number of maturities that RiskMetric^{TM} operates with.

- Secondly, it does not produce a probability distribution - which especially for non-linear securities is of importance. Ofcourse we can construct a probability distribution using the DECF approach - but because the DECF is not constant - this probability distribution will not be of limited use

It could of course be argued that in the scenario simulation procedure advocated here we need to perform 315 price calculations in order to derive the probability distribution - which however still could be quit time-consuming for large portfolios of MBBs - at least relatively to the DECF approach from Jacobsen.

8. Backtesting of our Model Setup

In this section we will perform a backtest of out model setup for the period 2. January 2008 - 2. January 2009.

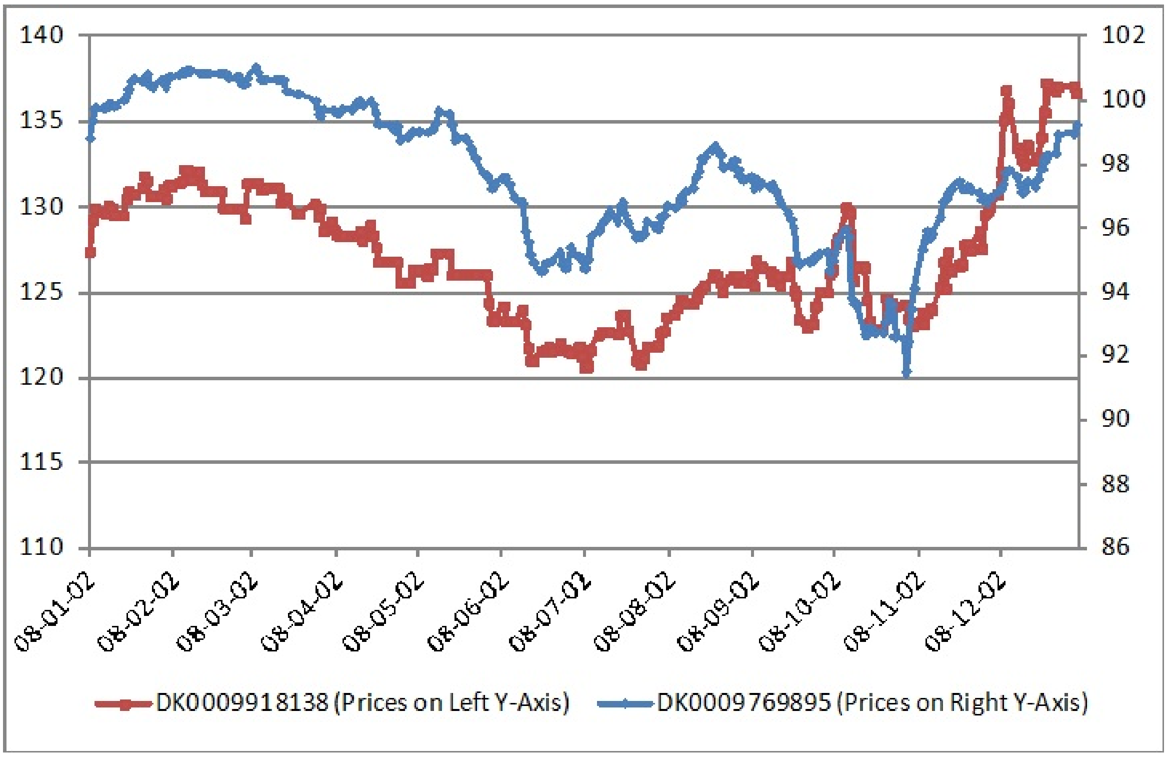
We will test the abilities of our model setup to capture the “true” risk for the following 2 bonds.

Table 6: Backtest Portfolio (Date: 2. January 2008)

Isin	Instrument-Name	Clean Price	Accrued IR	Dirty Price
DK0009769895	6 NYK MBB 2041	98,800	0,099	98,899
DK0009918138	7 Stat Ink 2024	127,312	1,109	128,421

In figure 6 below we have shown the price pattern for the 2 above bonds over our backtest period.

Figure 6: Price evolution for the period 2. January 2008 - 2. January 2009



We have performed the following calculations:

Starting with the 2. January 2008 we have every 10-days for the whole year 2008 done the following:

- Re-Estimated the PCA model, assuming 4-factors are to be selected and rotated the factor loadings using the Varimax-Method
- Re-Estimated our Garch(1,1) model and forecastet the volatility structure for the next 10-days
- Generate our 315 Yield-Curves using the JZ Approach with states(7,5,3,3) (as described in the previous sections)
- Given the simuleted 31 Yield-Curves we will now price our 2 Bonds 10-day in the future.
    - For the Government Bond that is straightforward as the value is only a function of the yield-curve
    - For our Mortgage Backed Bond, things are somewhat more complicated. The procedure is as follows:
        - We will first estimate the volatility parameters for our Log-Normal Hull-White model using at-the-money Caps, and assume the estimated volatility is constant over the next 10-days
        - Using the estimated volatility parameters and each of the 315 (simulated) yield-curves as inputs, we will now price the MBB 10-day in the future, using our one-factor Log-Normal Hull-White model combined with our

Prepayment-Model

- Knowing the 315 simulated values (prices) of our 2-Bonds 10-day in the future, will allows us to calculate the Profit-Loss distribution, as we know the value of the Bonds at the beginning of the 10-day period (=today)
- From the Profit-Loss Distributions determine VaR and ETL

These calculated risk-measures we wish to compare with the real-changes in the value of these bonds - especially will it be interesting to see how the model has performed during the Financial Crisis - notable autumn of 2008.

These data are shown in the following 2 figures.

Figure 7: Profit-Loss, VaR and ETL (DK0009918138)
(2. January 2008 - 2. January 2009)

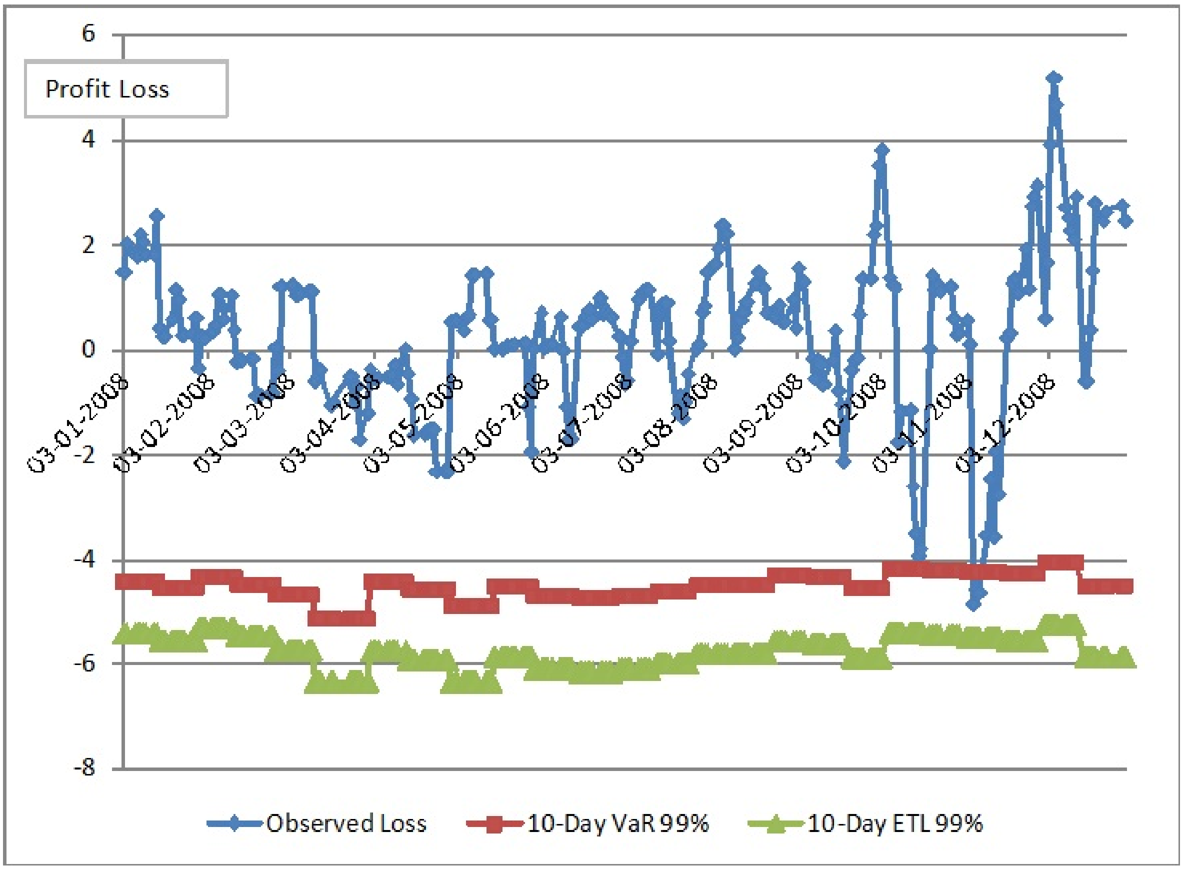
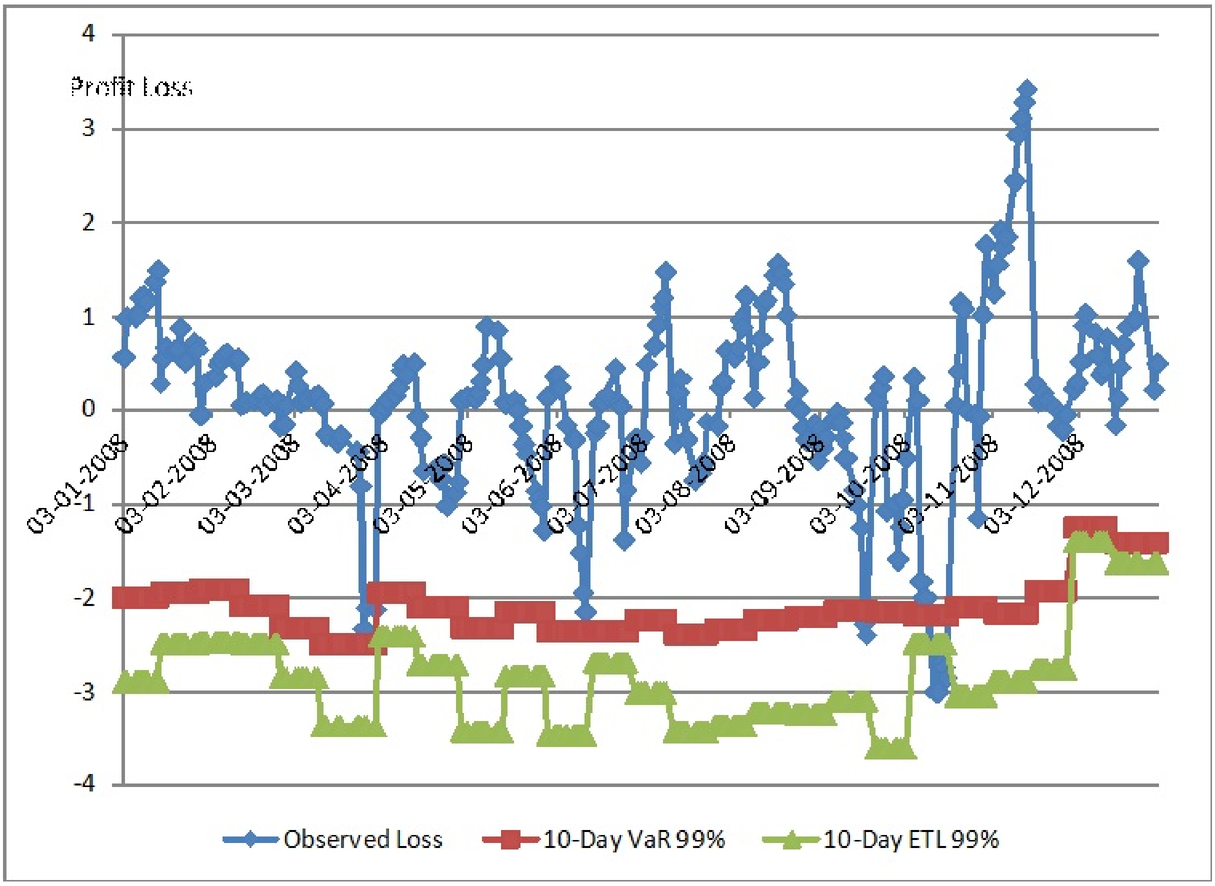


Figure 8: **Profit-Loss, VaR and ETL (DK0009769895) (2. January 2008 - 2. January 2009)**



From the above 2 figures we can draw the following conclusion:

- Number of breeches of our VaR number for the Government-Bond is 3 - expected breeches is approximately 2.5
  - The 3 breeches occur on the following 3 dates:
    - 5. November, 6. November and 7. November
  - The highest breech is less then 15% worse than anticipated by our VaR number
- There are no breeches of our ETL number for the Government-Bond
- Number of breeches of our VaR number for the Mortgage-Bond is 7 - expected breeches is approximately 2.5
  - The breeches occur on the following 7 dates:
    - 18. September, 19. September, 13-17 of October
    - The highest breech is approximately 37% worse than anticipated by our VaR number
- There are 5 breeches of our ETL number for the Mortgage-Bond
  - The breeches occur on the following 5 dates:
    - 13-17 of October
    - The highest breech is approximately 20% worse than anticipated by our ETL number

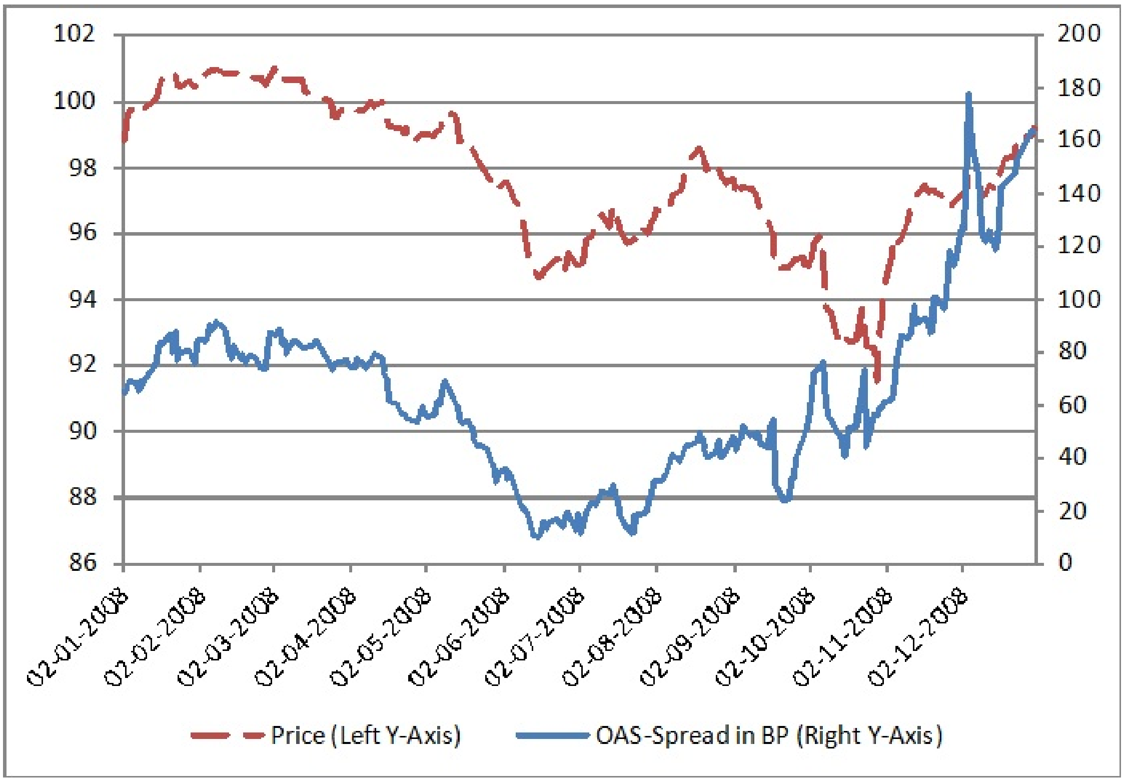
If we consider the results for our Government-Bond then we obtain very interesting results, one might even say that it is incredible that our model setup is so robust that under such extreme price movements we were witness to during the autumn of 2008.

The result obtained here is a strong indication that the model setup as defined in this paper is of practical interest due to its ability to capture the risk under extreme market conditions.

When we turn our attention to the results for the Mortgage-Bond, the results are not nearly as strong as for the Government-Bond. However, considering the fact that there are a lot more in play in modelling a Mortgage-Bond, it is not really that surprising - where the most notable problem during the financial crises was the following 2 facts:

- The market-marker arrangements for Mortgage-Bonds was suspended on several occasions during this period, which means that no real trading was done
- The was a huge changes in OAS-Spread - which no OAS-Model was able to capture, this can be seen in figure 9 below

Figure 9: Evolution in OAS-Spread for DK0009769895



If one takes that into account, we believe the results we obtain in our model setup also is very promising in connection with Danish MBB's.

8. Conclusion

The first result in this paper was the construction of a general model for the variation in the term structure of interest rates - that is we defined a model for the shift function. In this connection we showed - using the Heath, Jarrow and Morton (1991) framework - that the shift function could be understood as a volatility structure - more precisely the spot rate volatility structure.

The class of shift functions considered in this paper was of the linear type, with independence between the individual factors; the model was therefore comparable to the Ross (1976) APT model.

Using PCA we showed that it took a 4-factor model to explain the variation in the term structure of interest rates over the period from the 2. January 2002 to the 2. January 2012. These 4 factors can be called a Slope factor, a Short-Curvature Factor, a Short Factor and a Curvature Factor.

Because of the relationship between the volatility structure in the Heath, Jarrow and Morton framework and the shift function implied from our empirical analysis of the evolution in the yield-curve we concluded that PCA could be used to determine the volatility structure in the Heath, Jarrow and Morton framework - a non-parametric approach.

We then extended this model setup by adding stochastic volatility. This we did by assuming that the stochastic volatility was driven by a GARCH(1,1) model. Due to the fact that our stochastic volatility process was independent of our yield-curve factor-model, means that the extended model belonged to the class of USV (unspanned stochastic volatility) models.

In the last part of the paper we turned our attention to Risk model. Our approach to the calculation of VaR/ETL was a scenario simulation based methodology with relied on the framework of Jamshidian and Zhu (1997). This scenario simulation procedure builds on the factor loadings derived from a PCA of the same kind we used in our analysis of the empirical dynamics in the yield-curve.

The general idea behind the scenario simulation procedure is to limit the number of portfolio evaluations by using the factor loadings derived in the first part of paper and then specify particular intervals for the Monte Carlo simulated random numbers and assign appropriate probabilities to these intervals (states).

From our analyses of both straight bonds and Danish MBBs using the scenario simulation procedure we conclude the following:

- The scenario simulation procedure is a computational effective alternative to Monte Carlo simulation
- The scenario simulation procedure is capable of producing reasonable good approximations of the probability distributions with a limited number of states

- There is much better control over the extreme values in the scenario simulation than in Monte Carlo simulation
- The scenario simulation procedure is more efficient than the parametric approach for “linear” securities and it rivals the parametric approach for “linear” securities because of the speed of calculation
- We suggested using 7 x 5 x 3 x 3 states for the scenario simulation. Because we “only” need to perform 315 re-evaluations of the portfolio the scenario simulation procedure is feasible for large portfolios consisting of highly complex non-linear securities
- One last important feature is that we in the spirit of Barone-Adesi, Bourgoin and Giannopoulos (1997) extended the scenario simulation procedure to include stochastic volatility - in our case we showed that a Garch(1,1) was appropriate

We also backtested our Risk-Model setup for the year 2008 on 2 bonds - a Government-Bond and a Mortgage-Bond. The overall conclusion for this was:

- Our Risk-Model setup was able to capture the extreme movements influenced by the financial crisis for our Government-Bond case - notable the autumn of 2008
- Our Risk-Model setup did not manage to completely capture the extreme price movements for our Mortgage-Bond. Due to the fact that the reason for that was mostly (if not entirely) because of the enormous changes on level of OAS, we found the model indeed very promising - both of term of efficient from a calculation point of view, to its ability to capture the Risk of Mortgage-Bonds

More work is of course necessary - both with respect to backtesting the model and with respect to the determination of the input to the scenario simulation procedure.

In connection with the inputs to the scenario simulation method we especially need to address the following:

- How to forecast the volatility?
  - In this paper we argued that GARCH is a logical method to utilize
- How to select the expected yield-curve at a given time-horizon?
  - In this paper we suggested using the initial yield-curve as the expected yield-curve
- Would it be possible to make our OAS-Model even more robust?
- More examination on the appropriate number of states that is used in the scenario simulation method

Literature

Adams and Deventer (1994) “Fitting Yield Curves and Forward Rates with Maximum Smoothness”, Journal of Fixed Income vol. 4, no. 1, page 52-62

Alexander (1996) (editor) “Risk Management and Analysis”, John Wiley and Sons 1996

Amin and Morton (1993) "Implied Volatility Functions in Arbitrage Free Term Structure Models", Working Paper University of Michigan, May 1993

Barrett, Heuson and Gosnell (1992) "Yield Curve Shifts: An Empirical Solution to a Theoretical Dilemma", Working Paper, University of Miami, October 1992

Baron (1989) "Time Variation in the Modes of Fluctuation of the Treasury Yield Curve", Bankers Trust, October 1989

Barone-Adesi, Bourgoin and Giannopoulos (1997) “A Probabilistic Approach to Worst Case Scenarios”, working paper University of Westminster, August 1997

Beaglehole and Tenney (1991) "General Solutions of some Interest Rate-Contingent Claim Pricing Equations", Journal of Fixed Income 3, pp. 69-83

Berndt, Hall, Hall and Hausman (1974) “Estimation and Inference in Nonlinear Structural Models”, Annals of Economic and Social Management 3/4, page 653-665

Bierwag, Kaufmann and Toevs (1983) "Innovation in Bond Portfolio Management", JAI Press Inc.

Bierwag (1987a) "Duration Analysis Managing Interest Rate Risk", Balling Publishing Company 1987

Bierwag (1987b) "Bond Returns, Discrete Stochastic Processes, and Duration", Journal of Financial Research 10, pp. 191-209

Bierwag, Kaufmann and Latta (1987) "Bond Portfolio Immunization: Tests of Maturity, One- and 2-factor Duration Matching Strategies", Financial Review 22, pp. 203-219

Black and Karasinski (1991) "Bond and Option pricing when short rates are lognormal", Financial Analyst Journal, July-August 1991, page 52-69

Bollerslev (1986) "Generalized Autoregressive Conditional Heteroskedasticity", Journal of Economtrics 31, page 307-327

Boudoukh. Richardson and Whitelaw (1997) “Investigation of a Class of Volatility Estimators”, Journal of Derivatives vol. 4, no. 3, page 63-71

Casassus, Collin-Dufresne and Goldstein (2005) “Unspanned stochastic volatility and fixed income derivatives pricing”, Journal of Banking & Finance 29 (2005) 2723–2749

Caverhill and Strickland (1992) "Money Market Term Structure Dynamics and Volatility Expectations", Working Paper, University of Warwick, June 1992

Chambers, Carleton and Waldman (1984) "A New Approach to Estimation of the Term Structure of Interest Rates", Journal of Financial and Quantitative Analysis, Vol. 19, pp. 233-252

Chambers and Carleton (1988) "A Generalized Approach to Duration", Research in Finance, Vol. 7, pp. 163-181

Chambers, Carleton and Mcenally (1988) "Immunizing Default-Free Bond Portfolios with a Duration Vector", Journal of Financial and Quantitative Analysis, pp. 89-104

Christiansen (1998) “Value at Risk Using the Factor-ARCH Model”, working paper D 98-6, Department of Finance, The Aarhus School of Business

Cox, Ingersoll and Ross (1985) "A Theory of the Term Structure of Interest Rates", Econometrica 53, pp. 363-385

Cox and Rubinstein (1985) “Options Markets”, Englewood Cliffs: Prentice Hall 1985

Dahl (1989) "Variations in the term structure of interest rates and controlling interest-rate risks" [in Danish], Finans Invest, February 1989

Elton, Gruber and Nabar (1988) "Bond Returns, Immunization and The Return Generating Process", Studies in Banking and Finance 5, pp. 125-154

Engle (1982) "Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation", Econometrica vol 50, no. 4, page 987-1007

Engle and Ng (1991) “Measuring and Testing the Impact of News on Volatility”, working paper University of California

Engle, Ng and Rothschild (1990) “Asset Pricing with a Factor-ARCH covariance structure”, Journal of Econometrics 45, page 213-237

Garbade (1986) "Modes of Fluctuations in Bonds Yields - an Analysis of Principal Components", Bankers Trust, Money Market, June

Garbade (1989) "Polynomial Representations of The Yield Curve and its Modes of Fluctuation", Bankers Trust, July 1989

Giannopoulos (1995) “The Theory and Practice of ARCH Models”, working paper University of Westminster, 27 November 1995

Glosten, Jagannathan and Runkle (1991) “Relationship Between the Expected Value and the Volatility of the Nominal Excess Return on Stocks”, working paper Northwestern University

Golub and Van Loan (1993) “Matrix Computations”, The John Hopkins University Press, second edition 1993

Greene (1993) "Econometric Analysis", 2nd Edition, Macmillan Publishing Company 1993

Hald (1979) “Updating formulas in Quasi-Newton methods” (in Danish), Department for Numerical Analysis, report no. NI-79-03

Harman (1967) "Modern Factor Analysis", University of Chicago Press, 1966

Harvey, Ruiz and Shepard (1994) “Multivariate Stochastic Variance Models”, Review of Economic Studies, 61, page 247-264

Heath, Jarrow and Morton (1990) "Contingent Claim Valuation with a Random Evolution of Interest Rates", The Review of Futures Markets 9, pp. 54-76

Heath, Jarrow and Morton (1991) "Bond Pricing and the Term Structure of Interest Rates: A new Methodology for Contingent Claim Valuation", Working Paper Conell University

Ho (1992) “Key Rate Durations: Measures of Interest Rate Risks”, Journal of Fixed Income, September, page 29-44

Hull and White (1994) "Numerical Procedures for Implementing Term Structure Models", Working paper University of Toronto, January 1994

Ingersoll, Skelton and Weil (1978) "Duration forty years later", Journal of Financial and Quantitative Analysis, November, pp. 627-652

Jacobsen (1996) “Value-at-Risk for Danish Mortgage-Backed-Bonds” (in Danish), Finans/Invest 1/96, page 16-21

Jamshidian and Zhu (1997) “Scenario Simulation: Theory and methodology”, Finance and Stochastic vol. 1, no. 1, page 43-67

J. P. Morgan (1995) “RiskMetric^{TM} - Technical Document, third edition”, New York 26 May

J. P. Morgan (1997) “RiskMetric^{TM} - Monitor, fourth quarter 1997”, New York 15 December

Karner, Kelstrup and Schelde (1998) “A Guide to the Danish Bond and Money Market”, February 1998, Handelsbanken Markets, Reference Library no. 10

Kelstrup, Madsen and Rom-Poulsen (1999) “A Guide to the Danish Mortgage Bond Market”, Primo 1999, Handelsbanken Markets, Reference Library no. 18

LeBaron (1994) “Chaos and Nonlinear forecastability in Economic and Finance”, working paper University of Wisconson, February 1994

Litterman and Scheinkman (1988) "Common factors affecting bond Returns", Goldmann, Sachs & Co., Financial Strategies Group, September 1988

Longstaff and Schwartz (1991) "Interest-rate volatility and the term structure: A two-factor general equilibrium model", Working Paper, Ohio State University, November 1990

Lord and Pelsser (2007) “Level-Slope-Curvature - Fact or Artefact?”, Applied Mathematical Finance, Vol. 14, No. 2, 2007

Madsen (1997) “Determination of the implicit credit spread on Danish Mortgage-Backed Securities (MBS)”, Risk Magazine December 1997

Madsen (1998) “The Modelling of Debtor Behaviour - Mortgage Backed Securities I”, February 1998, Handelsbanken Markets, Reference Library no. 13

Madsen (1995) “Danish Mortgage Modeling in FinE”, working paper 17. June 2005, http://www.fineanalytics.com/images/stories/pdf/fine_mbb_modelling.pdf

Mandelbrot (1997) “Fractals and Scaling in Finance - Discontinuity, Concentration, Risk”, Springer-Verlag New York Berlin Heidelberg

McDonald (1996) “Probability Distributions for Financial Models”, in Maddala and Rao, eds., Handbook of Statistics vol. 14, page 427-461

Nelson (1992) “Filtering and Forecasting with Miss-specified ARCH Models: Getting the Right Variance with the Wrong Model”, Journal of Econometrics, vol. 52, page 61-90

Nelson and Siegel (1987) "Parsimonious modelling of Yield Curves", Journal of Business, October, pp. 473-489

Pedersen (1996) “Forecasting of interest rate volatility (3) - an empirical comparison of selected methods” (in Danish), Finans/Invest 6/96, page 22-27

Prisman and Shores (1988) "Duration Measures for Specific Term Structure Estimations and Applications to Bond Portfolio Immunization", Journal of Banking and Finance 12, pp. 493-504

Reitano (1992) "Non-Parallel Yield Curve Shifts and Immunization", Journal of Portfolio Management, spring, pp. 36-43

Repplinger (2008) “Pricing of Bond Options - Unspanned Stochastic Volatility and Random Field Models”, Springer-Verlag, 615 Lecture Notes in Economics and Mathematical Systems

Risk Books - editor Jarrow (1998) “Volatility - New Estimation Technique for Pricing Derivatives”, Risk Publications

Ross (1976) "The Arbitrage Theory of Capital Asset Pricing", Journal of Economic Theory 13, pp. 341-360

Rubenstein (1984) "A simple Formula for the Expected Rate of Return of an Option over a Finite Holding Period", Journal of Finance, Vol. 39, pp. 1503-1509.

Sentana (1991) “Quadratic ARCH Models: A Potential Re-Interpretation of ARCH Models”, working paper London School of Economics

Shea (1985) “Term Structure Estimation with Exponential Splines”, Journal of Finance vol. 40, page 319-325

Steeley (1990) "Modelling the Dynamics of the Term Structure of Interest Rates", The Economic and Social Review, Vol. 21, no. 4, pp. 337-361

Svensson (1994) “Estimating and Interpreting Forward Interest Rates: Sweden 1992-1994, IMF Working Paper, WP/94/114 (September 1994), 1-49

Tanggaard (1991) "A model for bond portfolio choice" [in Danish], Working Paper, the Århus Business School, January 1991

Tanggaard (1997) “Nonparametric smoothing of yield curves”. Review of Quantitative Finance and Accounting 9, page 251-267

Trolle and Schwartz (2008) “Unspanned stochastic volatility and the pricing of commodity derivatives”, The Review of Financial Studies / v 22 n 11 2009

Trzcinka (1986) "On the Number of Factors in the Arbitrage Pricing Model", Journal of Finance, Vol. XLI, pp. 347-368

Vasicek (1977) "An Equilibrium Characterization of the Term Structure", Journal of Financial Economics 5, pp. 177-188

Webber and James (2000) “Interest Rate Modelling” Wiley, 2000

Wilson (1994) “Plugging the Gap”, Risk Magazine, vol. 7, no. 10, October

Zakoian (1991) “Threshold Heteroskedastic Model”, working paper, INSEE, Paris

Appendix A

In principle, the literature contains three different ways in which it has been proposed to extend the traditional 1-factor duration model so that not only additive shifts in the term structure of interest rates are considered. In this connection, we have decided not to include the literature on stochastic term structure models.

The first class of models consists of the so-called multi-factor duration models, where they can be formulated as follows if we consider a 2-factor duration model:

$$PR_i = a_i + b_{i1}F_1 + b_{i2}F_2 + e_i \qquad \textbf{(30)}$$

Where PR_i is the return of the i'th bond, a_i is the risk-free return, b_{i1} and b_{i2} could be defined as the modified duration of a short bond and a long bond, respectively. In addition, F_1 and F_2 are defined as factors that relate to the first and second risk factors, which means that ε_i is the i'th bond residual. An alternative formulation could be to let b_{i2} be defined as a spread duration. With respect to duration models, the situation is that these factor values (F_1 and F_2) each represent the relation between changes in these basic rates (or basis spreads) and the term structure of interest rates itself; see, for instance, Ingersoll (1983).

Models that can be classified as being of this type have been defined by Elton, Gruber and Naber (1988), Reitano (1992), Bierwag (1987b) and Bierwag, Kaufmann and Latta (1987), and moreover Bierwag, Kaufman and Toevs (1983).

The other type of model is based on the functional form used in the estimation of the term structure of interest rates. The first model introduced in this class was Chambers, Carleton and McEnally (1988)[28]. The fundamental condition underlying these models is the fact that the functional form of the term structure of interest rates has been chosen in such a way that the unknown parameters that determine the term structure estimation are independent. One obvious functional form of the term structure that complies with this requirement is polynomial models, as follows:

$$P(t,T) = e^{-\sum_{j=0}^{n} a_j \tau^{j+1}} \qquad \textbf{(31)}$$

Where j is the polynomial degree, n is the maximum degree, a_j is the j'th coefficient and τ is the term to maturity, i.e. T - t[29]. From this can be clearly seen that there is independence between the individual unknown parameters, each single a_j. If a second-degree polynomial is considered, then the impacts on the term structure of interest rates can be seen to be defined by an additive factor, a term structure slope factor and a term structure curvature factor.

---

[28] In addition, reference can be made to Chambers and Carleton (1988), Prisman and Shores (1988), Steeley (1990) and Barret, Heuson and Gosnell (1992).

[29] This formulation can be seen to be identical to Chambers, Carleton and Waldmanns (1984) model, and was in fact tested on the Danish market by Tanggard and Jacobsen in a number of working papers in the late 1980's.

The risks of interest rate contingent claims as a function of changes in these coefficients can then to be measured, as these parameters uniquely determine the term structure of interest rates. At this point, it should be mentioned that Steeley (1990) used a cubic-spline model as his starting point, whereas Barret, Heuson and Gosnell (1992) use Nelson and Siegel's (1987) model. The fact that Nelson and Siegel's model also fulfils the condition regarding independence between the individual unknown parameters is not entirely the case, only to the extent that the τ-parameter is determined explicitly.

However, note 2 from Chambers, Carleton and McEnally (1988) demonstrated, in connection with the duration vector model, that the underlying principles of this model are only dependent on the fact that changes in the term structure of interest rates can be described by a polynomial of the n'th degree.

Where this leads us on to the third class of models, namely models that take the term structure of interest rates for given and explicitly specify some sort of functional form to describe the dynamics. At this point Garbade (1989) can be mentioned, which assumes that impacts on the term structure of interest rates are defined by a polynomium of the n'th degree. If this is formulated on the basis of the shift function, then this model can be expressed as follows:

$$S(t,T) = \sum_{j=0}^{n} a_j \tau^{j+1} \qquad \mathbf{(32)}$$

Or to put it differently, the spot rate volatility structure can be described by a polynomial of the n'th order. Where it is this last formulation of the shift function that is closest to the model constructed in section 3 in the text. On the basis of formula 3, it can be seen that the traditional 1-factor duration model is achieved for n = 0, i.e. the impacts on the term structure of interest rates are of the order a_0.

Appendix B

Some Statistics for the Relative Log-Returns

Table 1: **(period: 2 January 2002 - 2 January 2012)**[30]

Maturity	0,083	1	2	5	10	20	30
Centered Mean	-0,04752	-0,04909	-0,05439	-0,04504	-0,03327	-0,03019	-0,03272
Centered Variance	24,39939	12,14947	3,42270	3,09164	1,43235	1,47645	1,75291
Centered Skewness	0,06600	-0,10737	0,41372	-0,53219	-0,25574	-0,38302	-0,77399
Centered Kurtosis	23,86390	34,53680	22,97702	26,91020	7,49355	13,17215	12,65991
Auto-Correlation 1	-0,39468	-0,39681	-0,05862	-0,14179	-0,00966	-0,02977	-0,00274
Box-Ljung 1	387,56906	391,74695	8,55047	50,02209	0,23223	2,20483	0,01865
Chi-Square 1	3,84315	3,84315	3,84315	3,84315	3,84315	3,84315	3,84315
Auto-Correlation 5	-0,01558	-0,06142	0,01129	0,07892	0,04331	0,01012	0,01490
Box-Ljung 5	399,11255	423,48431	29,10601	98,69074	17,54937	15,51930	8,62700
Chi-Square 5	11,07050	11,07050	11,07050	11,07050	11,07050	11,07050	11,07050
Auto-Correlation 10	0,01408	0,00389	-0,04445	-0,01647	0,01907	0,01418	-0,01201
Box-Ljung 10	417,93811	456,36793	55,14677	125,32822	27,87941	19,98903	14,43744
Chi-Square 10	18,30705	18,30705	18,30705	18,30705	18,30705	18,30705	18,30705
Auto-Correlation 20	-0,03383	0,02147	0,01308	-0,00758	-0,02019	-0,03006	-0,00193
Box-Ljung 20	429,77804	495,70763	91,63903	195,42823	48,20522	32,54401	31,63084
Chi-Square 20	31,41044	31,41044	31,41044	31,41044	31,41044	31,41044	31,41044
Auto-Correlation 50	-0,02594	0,01039	-0,01013	-0,01163	-0,00842	0,01712	0,01810
Box-Ljung 50	575,22287	616,44750	157,82031	255,73414	86,35404	83,97849	92,86378
Chi-Square 50	67,50061	67,50061	67,50061	67,50061	67,50061	67,50061	67,50061

In general we observe that there is (significant) evidence of GARCH effect especially in the short end of the yield-curve - for short maturities.

In the next section we will briefly introduce the GARCH technique and in that connection specify how we have chosen to forecast the volatility in this paper.

B.1 GARCH Models, A Survey[31]

In theory there is no problem in finding useful estimations for the volatility, since volatility estimates are in principle a reflection of how the variable in question is expected to fluctuate. The direction of movements in the variable is in fact of no importance. The important thing , however, is how the variable fluctuates around its average value.

In practice, the volatility estimation is made complicated by a number of circumstances.

---

[30] Box-Ljung x is the Box-Ljung test statistics for the squared Auto-Correlation x - where x represents the number of lags.

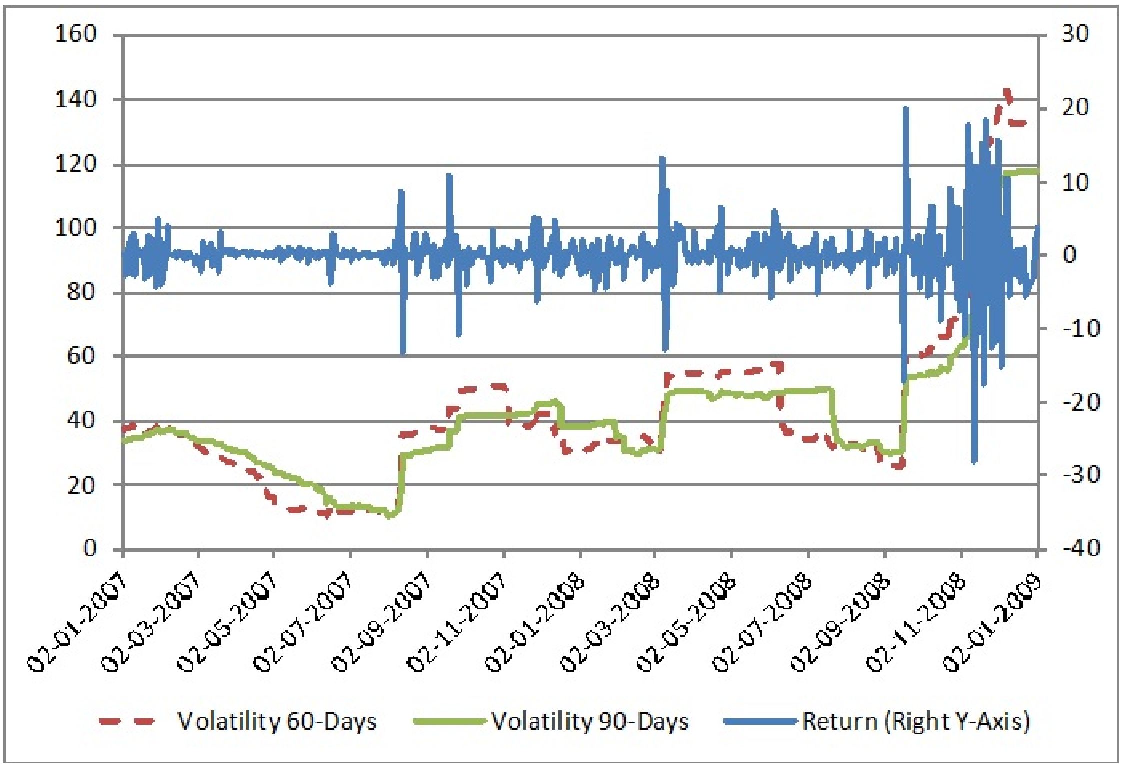
[31] We will not here try to reference the huge amount of litterature on the GARCH framework but instead refer to "Volatility - New Estimation Technique for Pricing Derivatives", Risk Books 1998 and RiskMetrics™ - Technical Document (1995,1997). Some particular references will however be given in the section whenever appropriate. On Danish data see in particular Pedersen (1996).

Probably the most significant problem is the instability of the volatility over short time-periods, i.e short estimation periods. On the other hand there is a certain degree of stability when considering longer estimation periods. This means of course that if the volatility changes to a relatively high degree, there is no reason to believe that volatility estimates achieved on the basis of previously observed interest rate movements will reflect future actual interest rate movements.

This has the implication that the assumption of constant volatility over the maturity spectrum is clearly not consistent with the observation that the volatilities which we have observed in the market are not constant.

This is illustrated in figure 1 below for the 1-year zero coupon rate[32]:

Figure 1 Historical Volatility (2. January 2007 - 2. January 2009)



---

[32] This data has been centered around the mean - that is we will determine volatility assuming a mean = 0 - so that the square root of the conditional second moment is in fact the volatility. This is a general assumption in the asset pricing literature, corresponding closely to the notion of market efficiency and the random walk hypothesis.

The maximum likelihood estimator is used to calculate the volatility, thus:

$$\sigma = \sqrt{\frac{1}{n-1}\sum_{t=1}^{n} y_t^2} \quad for \quad y_t = \ln\frac{r_{t+1}}{r_t} \tag{33}$$

where r_t is the interest rate at time t, n is the number of observations, σ is the volatility pr. day and y_t represents the return.

From the discussion above we have that if the volatility estimator obtained from formula 33 is to be useful for forecasting purpose, then:

- The interest rate volatility has to be constant
- The interest rates have to be lognormally distributed and for this reason the returns be normally distributed

More precisely we postulate the following model for the returns:

$$y_t = \varepsilon_t \qquad \varepsilon_t \approx N(0,\sigma^2) \tag{34}$$

That is the returns are treated as a time series of independent, normally distributed stochastic variables with a constant variance.

Clearly this assumption is not valid - as seen from figure 1. What we instead see is that large changes tend to be followed by large changes - of either sign - and small changes tend to be followed by small changes[33] - of either sign. This is often referred to as the clustering effect.

In general the following observation has been reported in the literature with respect to returns:

- Return distributions have fat tails and higher peak around the mean, that is we can observe excess kurtosis (normal distribution = 3)
- Returns are often negatively skewed (normal distribution = 0)
- Squared returns often have significant autocorrelation

For our data (see table 1) there is evidence of excess kurtosis, positive skewness and autocorrelation in the squared returns across the wbole maturity spectrum - though more pronounced for shorter maturities.

These findings indicates that a normal distribution assumption might not be appropriate when modelling returns.

---

[33] This was first observed by Mandelbrot 1963 - see Mandelbrot (1997).

The devotion to find more precise estimation techniques has inspired reseach in finance. Some of the most notable approaches can be categorized as follows:

- Introducing other probability distributions, for an overview, see McDonald (1996)
- GARCH[34]-type models, see Bollerslev (1986)
- Stochastic volatility models, see for example Harvey, Ruiz and Shepard (1994)
- Application of chaotic dynamics, see LeBaron (1994)
- MDE (Multivariate Density Estimation), see Boudoukh, Richardson and Whitelaw (1997)
- Jump Diffusion models, see for example J.P. Morgan (1997)

In the academic literature GARCH-models have been the most popular - which is due to the evidence that time series realizations of returns often exhibit time dependent volatility[35]. Because of that we will restrict myself to discussing GARCH-models.

Since the introduction of GARCH in the literature a number of new models in this framework have been developed. To name a few we could mention IGARCH, EGARCH and AGARCH.

Most financial studies conclude that a GARCH(1,1) is adequate, ie:

$$\begin{aligned} y_t &= \varepsilon_t \qquad \varepsilon_t \approx N(0,h_t) \\ h_t &= \omega + \alpha\varepsilon_{t-1}^2 + \beta h_{t-1} \end{aligned} \tag{35}$$

In order for a general GARCH process to be stationary we must have that the sum of the roots lies inside the unit circle, more precisely we have the following stationarity condition for the GARCH(1,1) model in formula 35; $1 > \alpha + \beta$.

Using the law of iterated expectation the unconditional variance for a GARCH(1,1) can be expressed as:

$$E_t[\varepsilon_t^2] = E[\varepsilon_t^2|F_{t-1}] = E[h_t] = \omega + [\alpha + \beta]E[\varepsilon_{t-1}^2] \tag{36}$$

which can be recognized as a linear difference equation for the sequence of variances.

Assuming the process began infinitely far in the past with a finite initial variance - then the sequence of variances will converge to a constant:

---

[34] GARCH mean Generalized Autoregressive Conditional Heteroskedasticity - and it represents an extension of the ARCH-model from Engle (1982).

[35] It is of course possible to model time dependent volatility directly using stochastic volatility models and the MDE-approach - but as of now limited investigation is available. It is nevertheless an interesting alternative, which however is left for future research.

$$E_t[\varepsilon_t^2] = h = \frac{\omega}{1-\alpha+\beta} \tag{37}$$

which implies $1 > \alpha + \beta$ in order for h to be finite.

From this we have that even though the conditional distribution of the error is normal then the unconditional distribution is non-normal - which is an very attractive feature of GARCH models.

An important property arises by inspecting equation 37, namely: shocks to the volatility decay at a speed that is measured by $\alpha + \beta$. Furthermore the closer to one (1) $\alpha + \beta$ is, the higher is the persistence of shocks to the current volatility.

For $\alpha + \beta = 1$ the shocks to volatility will persist for ever and in this case the unconditional variance is not determined. A process with such a property is known as an IGARCH - Integrated GARCH. It is of special interest to focus a little on the IGARCH since as will become apparent in a moment it is identical to the EVMA model[36] (apart from a constant) that is used in RiskMetric™.

The EVMA (Exponentially Weighted Moving Average) places more weight on recent observation - which will have the effect of diminishning the "ghost-features" which are apparent in figure 1. This exponential weighting is done by using a smoothing constant λ - where we have that the larger the value of λ the more weight is placed on past observations.

The infinite EVMA model can be expressed:

$$h_t = (1-\lambda)\sum_{i=1}^{\infty}\lambda^{i-1}\varepsilon_{t-i}^2 \tag{38}$$

To show the equivalence between the IGARCH and the EVMA model represented by formula 38 we first recall that the IGARCH(1,1) can be written as:

$$h_t = \omega + (1-\lambda)\varepsilon_{t-1}^2 + \lambda h_{t-1} \tag{39}$$

A repeated substitution in formula 39 yields:

$$\begin{aligned} h_t &= \omega + (1-\lambda)\varepsilon_{t-1}^2 + \lambda[\omega + (1-\lambda)\varepsilon_{t-2}^2 \\ &\quad + \lambda[\omega + (1-\lambda)\varepsilon_{t-3}^2 + \lambda[\ldots\ldots \\ &= \frac{\omega}{1-\lambda} + (1-\lambda)[\varepsilon_{t-1}^2 + \lambda\varepsilon_{t-2}^2 + \lambda^2\varepsilon_{t-3}^2 + \ldots.] \end{aligned} \tag{40}$$

---

[36] It is not equal to the EVMA in general but the IGARCH is equal to a particular EVMA model - which happen to be the one used in RiskMetric™.

If we compare formula 40 with formula 38 it follows that they are identical apart from a constant.

One very powerful result with respect to GARCH is that GARCH models are not so sensitive (as ARCH-models) to misspecification - because as Nelson (1992) shows, then even if the conditional variance in a linear GARCH model has been misspecified - then the parameters of the model are still consistent. Furthermore the GARCH model is not sensitive to the distribution assumption - that is the parameters are still consistent if there is evidence of non-normality in the squared normalized residuals[37], ie: $\hat{\varepsilon}_t^2 = \frac{\varepsilon_t^2}{h_t} \approx N(0,1)$. In this connection i however worth mentioning that if we decide to estimate a GARCH model under a normal distribution assumption - then we (normally) have to do a (< 99%) fractile adjustment of the original data - this fractile adjustment is however not necessary if we assume for example a t-distribution.

Before returning to the data in table 1, let me briefly mention a few of the non-linear GARCH models and their implications[38]:

Engle and Ng (1991) propose the AGARCH[39] (asymmetric GARCH) which can be expressed as:

$$\begin{aligned} y_t &= \varepsilon_t \qquad \varepsilon_t \approx N(0,h_t) \\ h_t &= \omega + \alpha[\varepsilon_{t-1} + \gamma]^2 + \beta h_{t-1} \end{aligned} \qquad \textbf{(41)}$$

which has similar properties as the EGARCH from Nelson (1990) - but from an implementaion point of view much simpler. The AGARCH has similar properties as the GARCH model but unlike the GARCH model which explores the magnitude of one-period lag-errors, the AGARCH model allows for the past error to have an asymmetric effect on the variance. For example if γ is negative then the conditional variance will be higher when ε_{t-1} is negative than when it is positive. For $\gamma = 0$ the AGARCH model degenerates to the GARCH model. Because of the γ-parameter the AGARCH model can - like the EGARCH model - capture the leverage effect which has been observed in the stock market. The unconditional variance for the AGARCH model has the following form:

---

[37] An investigation of the autocorrelation in the squared normalized residuals may also reveal model failure. Pagan and Schwert (1990) furthermore proposed to regress the squared normalized residuals against a constant and h_t, ie: $\hat{\varepsilon}_t^2 = a + bh_t + error$. If the forecast is unbiased, a = 0 and b = 1. In addition if a obtained by performing the regression - this will indicate that the model has a high forecasting power (for the variance).

[38] This section relies mostly on Giannopoulos (1995, section 4.2).

[39] The QGARCH from Sentana (1991) is similar to the AGARCH model except for the sign of γ.

$$E_t[\varepsilon_t^2] = h = \frac{\omega + \gamma^2\alpha}{1 - \alpha + \beta} \tag{42}$$

which indicates that the stationarity condition for the AGARCH model is identical with the condition for the GARCH model.

Lastly let me list a few other asymmetric GARCH specifications:

- The Threshold GARCH from Glosten, Jagannathan and Runkle (1991) and Zakoian (1991), ie:

$$h_t = \omega + \alpha\varepsilon_{t-1}^2 + \beta h_{t-1} + \gamma I_t \varepsilon_{t-1}^2 \tag{43}$$

where I_t is an indicator function defined as: $I_t = \begin{pmatrix} \varepsilon_t < 0 \Rightarrow 1 \\ \varepsilon_t \geq 0 \Rightarrow 0 \end{pmatrix}$. The reason why this mode

referred to as a threshold model is because when γ is positive then negative values of ε_{t-1} will have an additive impact on the conditional variance. As in the AGARCH model then negative errors have a greater impact on the variance if $\gamma < 0$.

- The non-linear AGARCH from Engle and Ng (1991), ie:

$$h_t = \omega + \alpha[\varepsilon_{t-1} + \gamma\sqrt{h_{t-1}}]^2 + \beta h_{t-1} \tag{44}$$

- The VGARCH model from Engle and Ng (1991), ie:

$$h_t = \omega + \alpha\left[\frac{\varepsilon_{t-1}}{\sqrt{h_{t-1}}} + \gamma\right]^2 + \beta h_{t-1} \tag{45}$$

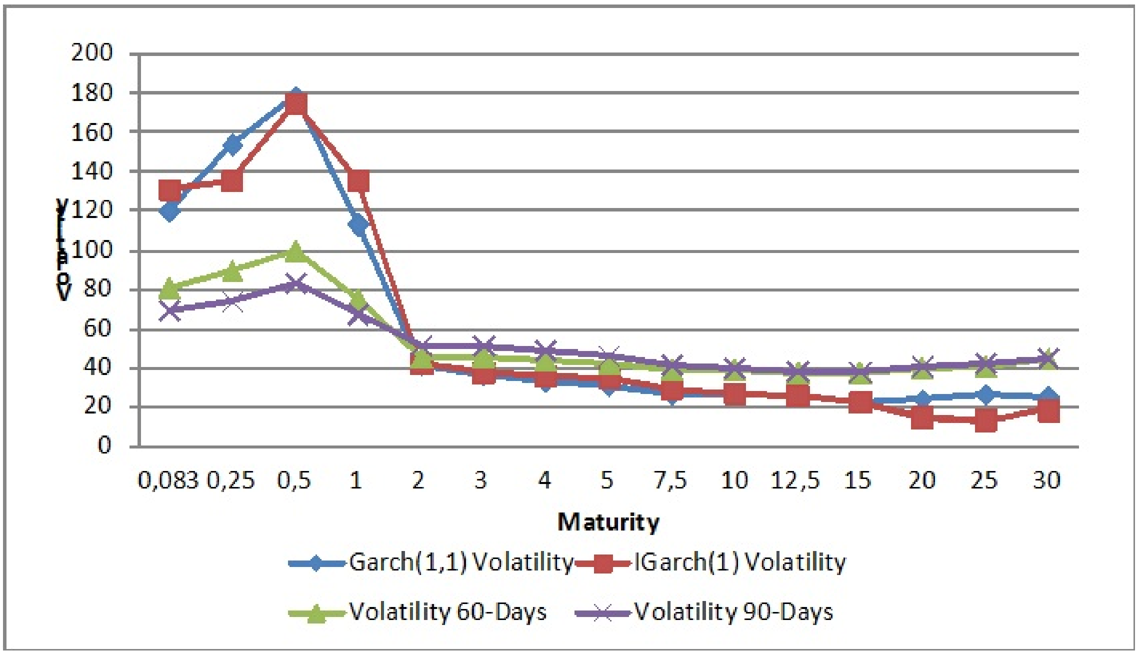
B.2 Modelling the Volatility structure

We will now return to the data in table 1.

For all the date series we have estimated the GARCH(1,1) model and the IGARCH(1) model.

The volatility structure as at 2 January 2012 for each of the estimation approaches compared to the 60-and 90 day rolling window technique, is shown below in figure 2:

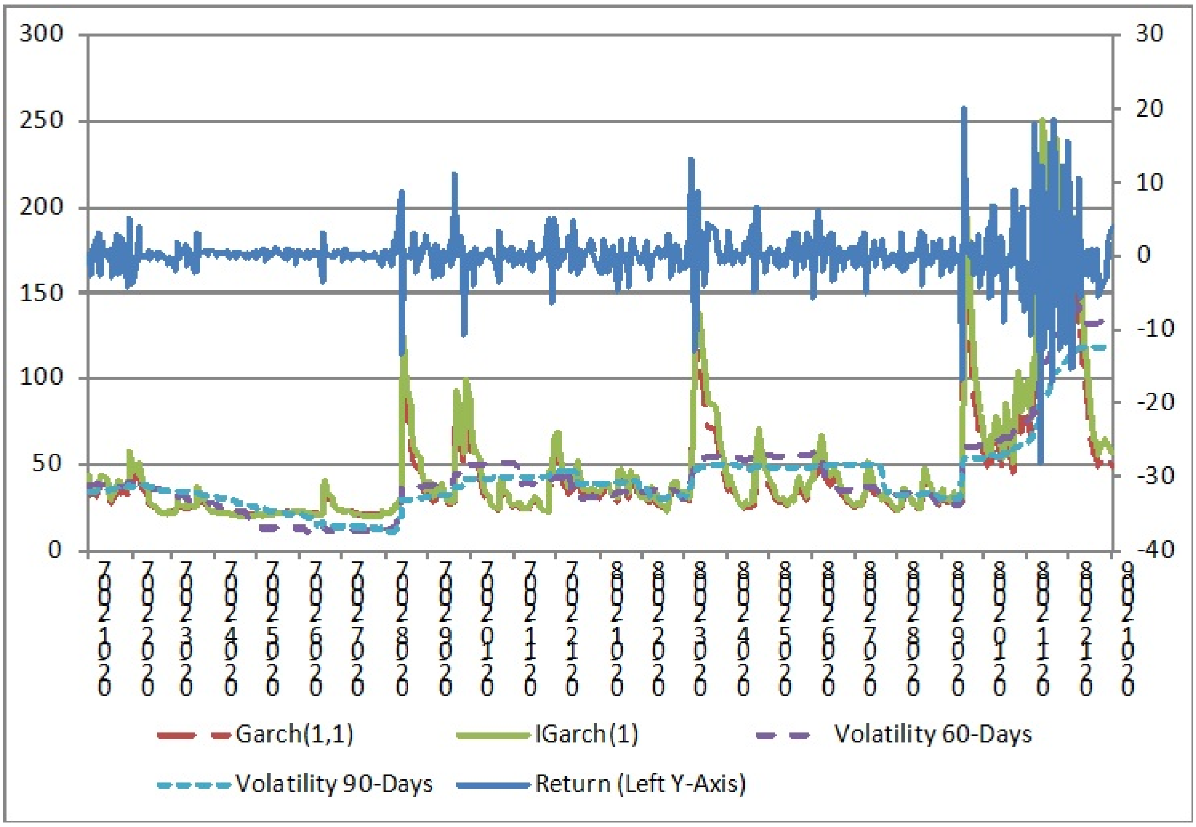
Figure 2 **Comparison of Volatility Structures**



As seen from the figure the differences in volatility are more pronounced for short maturities.

To illustrate this we have in figure 3 shown the volatility pattern for the 1-year rate over the period 2. January 2007 - 2 January 2009:

Figure 3 Evolution in the Volatility for the 1-year Interest Rate



In connection with the estimation results we can report the following[40]:

- The persistence in volatility is comparable to what is usually reported in the literature. Furthermore, our estimate of β is largest for the 1-month rate (0.92) and in general is declining with maturity - however it never falls below 0.81
- All the time-series are stationary - $1 >> \alpha + \beta$
- In most of the cases R^2 is higher in the GARCH model than in the IGARCH model - allthough little difference is observed. In general we have that R^2 is between 11-27% - with the highest R^2 for shorter maturities
- For both the GARCH model and IGARCH model we observe that we can accept the hypothesis of no autocorrelation in the squared normalized residuals
- In general we have that the normalized residuals exhibit lesser excess kurtosis and less skewness - that is we are close to the normal distribution (standardized normal distribution)

At the end of this Appendix we have shown a detailed description of our estimation results for both the GARCH model and the IGARCH model for the 1-year interest rate.

---

[40] For reasons of space we have omitted the tables here - except for the 1-year rate - but they can be obtained from the author.

From our estimation results we have decided to use the GARCH model instead of the IGARCH model, because of the following two observations:

- In most of the cases R^2 is higher in the GARCH model than in the IGARCH model
- In all cases Akaike's Information Criterion (AIC) was in favor of the GARCH model

Construction of a term structure of GARCH forecast for any time horizon can now be derived from the estimated model. This is in general straightforward[41], and is performed iteratively using the appropriate variance specifications - though taking into account that for the j-step ahead forecast for $j > 1$ the squared return is to be set equal to the variance for the j-1 step - this is what we do in our calculations.

Remark:

As mentioned by Alexander (1996) insufficient GARCH effect in data may lead to convergence problems in the optimization procedure. This is true if the GARCH model is being estimated using a variant of the BHHH-algoritm[42] - which seems to be the preferable optimization procedure suggested in the literature, see for example Greene (1993) and Bollerslev (1986). We howerver did not encounter any convergence problems, probably for the following reasons:

- There is significant evidence of GARCH effect in data
- We used the BFGS[43] as our main optimization procedure, which is much more robust than the BHHH

Lastly it is worth mentioning that the BFGS method was initialized with starting values obtained from the downhill-simplex procedure.

---

[41] Except for the EGARCH-model, see Alexander (1996).

[42] See Berndt, Hall, Hall and Hausman (1974).

[43] See Hald (1979).

Appendix C

In this appendix we will briefly explain how we have designed our yield-curve interpolation.

Why this is of importance can be formulated as:

- From the swap market it is possible to build a yield-curve using a bootstrapping procedure - which ensures that all swaps and any included forward-and spot-rates are priced perfectly. However, between observable data points, some yield-curve smoothing techniques are neccesary
- If we use Monte Carlo simulation procedure (or any other sampling technique) to generate interest rate paths from a subset of maturity points - then a smoothing technique is needed to generate a continuous yield-curve at every time-step

These are probably the two most important reasons why a flexible, robust and appropriate smoothing procedure has to be designed[44].

We have here selected the maximum smoothness approach of Adams and Deventer (1994). Other methods exists, such as for example spline-methods - but the problems outlined in Shea (1985) still remain. Shea has the following comments to forward-rates obtained by polynomial splines - namely that they are unstable, fluctuate widely and often drift off to very large positive or even negative values.

The yield-curve can either be formulated in terms of prices, spot-rates or forward-rates, ie:

$$\begin{aligned} P(0,T) &= e^{-R(0,T)T} \\ P(0,T) &= e^{-\int_0^T f(0,s)ds} \\ &\textit{where} \\ f(0,T) &= R(0,T) + T\frac{\partial R(0,T)}{\partial T} \end{aligned} \qquad \textbf{(46)}$$

where P(0,T), R(0,T) and f(0,T) are respectively the bond-price, the spot-rate and the forward-rate.

The idea of Adams and Deventer is to determine the maximum smoothness term structure within all possible functional forms.

If the maximum smoothness criteria is defined as the forward rate curve on an interval (0,T) that minimizes the functional:

---

[44] Of course smoothing techniques can also be used directly to obtain the yield-curve from prices of coupon bonds - this issue will however not be addressed here - see instead Tanggaard (1997).

$$\min \int_0^T \left[\frac{\partial^2 f(0,s)}{\partial s^2} \right]^2 ds \tag{47}$$

then Adams and Deventer show that the forward-rate model that satisfies the maximum smoothness criteria from formula 47 while fitting the observed prices is a fourth-order spline of the following special kind:

$$\begin{gathered} f(0,t) = a_i + b_i t + c_i t^4 \\ \text{for } t_{i-1} < t \le t_i \\ \text{and} \\ i = 1,2,3,....,m + 1 \end{gathered} \tag{48}$$

where m is the number of observed bond-prices.

The smoothness criteria and the requirement that we want to price the m bonds without measurement error gives rise to the following 3m + 3 system of equations for the coefficients a_i, b_i and c_i, for all i[45]:

$$\begin{gathered} a_i + b_i t_i + c_i t_i^4 = a_{i+1} + b_{i+1} t_i \quad \text{for } i = 1,2,3,...,m \\ b_i + 4c_i t_i^2 = b_{i+1} + 4c_{i+1} t_i^3 \quad \text{for } i = 1,2,3,...,m \\ a_i(t_i - t_{i-1}) + \frac{1}{2} b_i (t_i^2 - t_{i-1}^2) + \frac{1}{5} c_i (t_i^5 - t_{i-1}^5) = -\ln\left(\frac{P_i}{P_{i-1}} \right) \quad \text{for } i = 1,2,3,...,m \\ c_{m+1} = 0 \\ b_{m+1} = 0 \\ a_1 = r_1 \end{gathered} \tag{49}$$

It should be mentioned that the last two criteria have been selected for practical reasons - they are not associated with the smoothness criteria or the zero (0) measurement requirement for the observable bond-prices.

The last criteria ensures that the shortest observable spot-rate is equal to the shortest forward-rate - a logical feature. The second last criteria makes restrictions on the asymptotic behaviour of the yield-curve - more precisely it ensures that the slope of the forward-rate curve is zero (0) at the endpoint of the last interval.

As a final remark we might note that the algebraic linear system leads to a banded symmetric diagonally dominant and positive definite linear system which can be easily solved using special algorithms, see Golub and Van Loan (1993, chapter 5).

---

[45] See Adams and Deventer (1994).

Let us finally illustrate the method with a simple example:

For that purpose we assume we know the prices for the following maturity dates:

Period	Interest Rate	Bond Price
3-Month	3,75	99,067
1-Year	4,25	95,839
2-Year	4,35	91,668
5-Year	4,56	79,612
7.5-Year	4,20	72,979
10-Year	4,75	62,189
15-Year	4,31	52,388
30-Year	5,65	18,360

In figure 1 below we have shown the spot-rate curve and the forward-rate curve using the Adams and Deventer procedure.

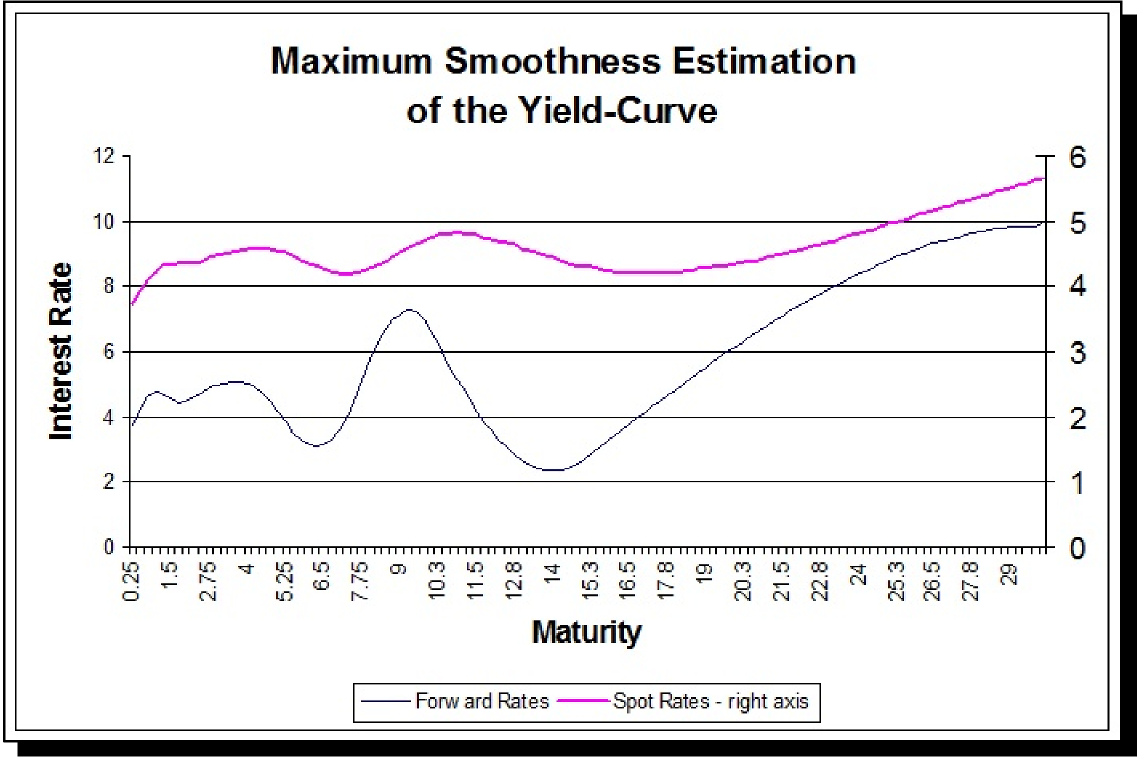


Figure 1

The data might be a bit extreme but it serves to illustrate the maximum smoothness procedure.

Part II

Chapter 6

Can Danish households benefit from Stochastic Programming models? – An empirical study of mortgage refinancing in Denmark

Can Danish households benefit from Stochastic Programming models? - An empirical study of mortgage refinancing in Denmark

Kourosh M. Rasmussen* Claus A. Madsen† , Rolf Poulsen‡

August 2012

Abstract

A number of stochastic programming (SP) models on mortgage choice and refinancing for Danish households have been introduced during recent years (Nielsen & Poulsen (2004), Rasmussen & Clausen (2004), Rasmussen & Zenios (2007)). A major Danish mortgage bank has adapted this model framework. Whereas most mortgage banks in Denmark today advise private home owners to finance their property with one loan only, our models suggest that most households are better off with two loans. With regards to refinancing, our models suggest a higher level of refinancing activity than what is observed today. The empirical study, which is the subject of this paper, is designed to perform a historical ex-ante test of the advice generated by our model framework for the period 1995-2010. We will compare SP-based advice with current practice which is based on rules of thumb and short-sighted forecasts. The study is tailored for the Danish market, and it demonstrates the advantages that the Danish mortgage system offers Danish household.[1]

---

*Technical University of Denmark, Lyngby, Denmark, email: kmra@man.dtu.dk
†Technical University of Denmark, Lyngby, Denmark, email: cmads@man.dtu.dk
‡University of Copenhagen, Copenhagen, Denmark, email: rolf@math.ku.dk

[1] The authors wish to thank the Danish knowledge center of housing economics for funding this project.

1 Introduction

The global financial crisis of 2008 which was triggered by the burst of the house price bubble in the US in 2007, shows the need for:

1. More regulation in the US mortgage market.
2. More robust evaluation methods for borrowers creditworthiness.
3. Better advising of individual borrowers on their choice of mortgage loan and refinancing.

The Danish housing market experienced a similar bubble and burst, though with a slight delay compared to the US market. However all Danish mortgage banks survived the crisis, since the foreclosures in Denmark were not nearly as bad as those observed in the US. See Figure 1.

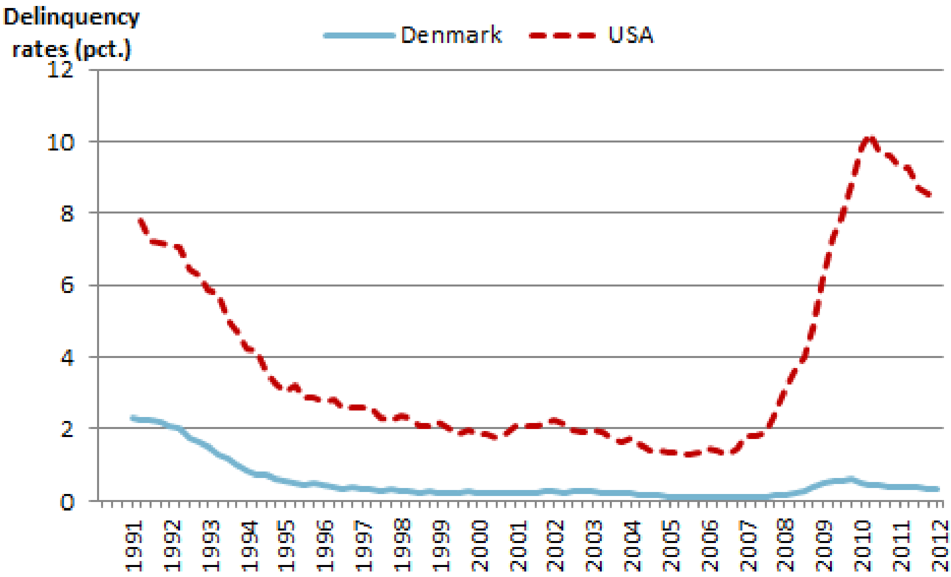


Figure 1: Delinquency rates for the period 1991 to 2012 in the USA and Denmark. The graphs are based on delinquency data from the homepages of Realkreditrådet, for Denmark, and the Federal Reserve Board for the USA.

There are a few considerable regulatory differences in between the Danish and the US mortgage systems, and these differences have certainly contributed to the huge difference in between the number of foreclosures in the two countries.

One major difference is that the Danish borrowers are liable to repay their debt, regardless of the value of the collateral, so Danish banks need not worry about the so called jingle mails[2], as soon as the value of the house falls below the outstanding debt.

Another major difference is that, despite introduction of several new mortgage products in Denmark in the period 1997 to 2004, some of the most risky loan products, such as balloon loan (with no initial payments for the first few years) were never introduced in Denmark.

Finally the Danish system offers complete transparency due to the practice of the principle of balance. Principle of balance, is a cornerstone of the Danish mortgage system, as practiced and enforced by the mortgage loan act since the introduction of the Danish mortgage system in 1797. The principle of balance is no longer enforced by law due to the covered bond legislation of 2004 imposed by the EU. However Danish mortgage banks still use the principle of balance for transparency, competition and risk management reasons. This has practically been the case for almost all Danish mortgage loans, also after the regulatory changes of 2004.

According to this principle there has to be an overall match between the funding side and the loan side at any given time. This is why the Danish mortgage system is also characterized as a match funding or a pass-through system. The few specialized mortgage banks in Denmark are intermediaries in between the investors in the one side and the borrowers in the other side. Danish borrowers essentially issue mortgage bonds via the mortgage banks, and the bonds are sold to investors. The borrowers pay interest and principle payments on the issued bond plus a markup of around 0.5 - 0.7 % as a risk premium and administration fee to the mortgage bank. The mortgage bank assumes only the default risk on the side of the borrowers. This is, however, covered mainly by the value of collateral. With maximum Loan to Value (LTV) levels of 80 % for private houses, the value of the collateral should fall by more than 20%, before the mortgage banks encounter any losses in cases of foreclosures. Keeping in mind, that the borrowers, are still held accountable for the debt they might have, after the foreclosure is settled, the mortgage banks are running a relatively small risk in Denmark. And that explains the relatively small markups.

The Danish mortgage market served the borrowers only with fixed rate mortgages during the first 200 years of its history. Several new mortgage loan products have been introduced since 1997 (See Fig-

---

[2] Jingle mails refer to envelopes containing house keys sent back to the mortgage providers.

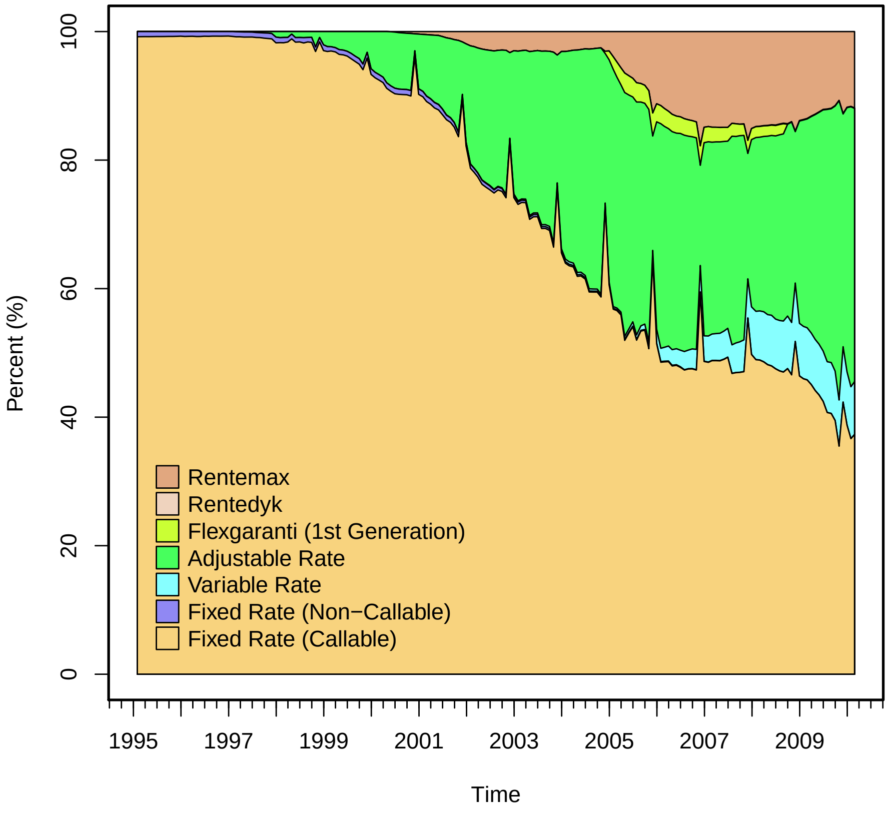


Figure 2: Development of the distribution of the underlying bonds used as funding instruments behind the Danish mortgage loans in the period 1995 to 2010.

ure 2 on distribution of mortgage loans in Denmark since 1995). The figure shows, however, that two products dominate the market, the adjustable rate mortgage (mostly with annual adjustments) and the traditional long term fixed rate mortgages.

Another innovation was the introduction of Interest Only (IO) versions of all types of mortgage loans in 2004. The IO loans contributed positively to the growth of the house price bubble in the period 2004 to 2007 in Denmark. Some central market players, such as the Danish National Bank, have therefor proposed banning or partially banning IO loans. IO loans, however, once used as a last resort, may prove to be a useful instrument to avoid a foreclosure in times of financial distress. This requires that IO loans are considered as part of an individual debt management strategy with high focus on assessing the risks associated with them.

Despite a very robust and well functioning market, there is room for improvement, in particular when it comes to advising home owners on their choice of mortgage and the following refinancing opportunities. The advisory practice has not followed the rapid product development seen in the period 1997-2007. In particular there is no consistent practice when it comes to risk management (on the borrower side) of the very popular short term adjustable rate mortgages. The post crisis extreme low interest rate regime makes development of such methods into an immensely urgent matter, since the new loans issued in Denmark are by far of the type adjustable rate mortgages with annual adjustments (F1 loans). If interest rates increase sharply within the next few years, there is a serious risk that the foreclosure rates in Denmark increase sharply as well, contributing to another wave of house price declines and the consequences it will have on the rest of the economy.

This paper focusses on improving advisory standards for mortgage choice and refinancing for the Danish market. In particular the focus is on improving and extending the rules of thumbs used for refinancing. Refinancing rules of thumbs have traditionally only dealt with refinancing among fixed rate mortgages, but given that most borrowers today opt for F1 loans there is a need for refinancing rules which reduce the risk of holding F1 loans by looking into strategies which allow movement from one type of loan to the other.

The contribution of this paper is the following:

1. Showing that the existing rules of thumb for refinancing are of little or no use.
2. Introducing a model-based refinancing framework ensuring significant refinancing gains without adding risk.
3. Reducing the risk of holding a F1 loan by introducing strategies which allow mixing fixed and adjustable rate loans.

The rest of this paper is organized in the following way. Section 2 explains the existing rules of thumb for refinancing in Denmark.

Section 3 introduces the model framework used in this paper for individual mortgage choice and refinancing.

Section 4 compares the performance of the model-based advising framework compared to the existing rules of thumb.

Finally section 5 concludes the paper suggesting that future innovations in the Danish mortgage market are likely to have special focus on risk management and advisory services. The existing mortgage products offer considerable potential for diversification and individualization in terms of risk and cost characteristics of the loans. This potential is however not taken advantage of due to lack of effective and personal consultancy tools and services. The framework suggested in this paper suggests a foundation for such tools and services.

2 Refinancing rules of thumb

Based on interviewing key analysts with overall responsibility for advising home owners on their choice of mortgage loan and refinancing, the rules of thumb for refinancing of mortgage loans across the Danish mortgage banks may be generalized as follows:

Refinancing from a fixed rate loan to another fixed rate loan with lower coupon:

- The coupon of the new loan should be 2 percent lower than the coupon of the existing loan.
- The quarterly payments of the new loan should be at least 5 percent lower than the ones from the existing loan.
- The outstanding debt of the existing loan should be more than DKK 500.000.
- The maturity of the existing loan should be greater than 10 years.
- The price of the new loan should be over 95.

Refinancing from a fixed rate loan to another fixed rate loan with higher coupon:

- At least 10 percent reduction in outstanding debt.
- The outstanding debt of the existing loan should be more than DKK 500.000.
- The maturity of the existing loan should be greater than 10 years.
- The price of the new loan should be over 98.

This kind of refinancing (to a higher coupon) makes sense in Denmark, since the borrower has the right to redeem the mortgage loan by buying back the underlying bonds at the prevailing market price.

This is referred to as the "buy back delivery option", even though this is not really an option in a financial sense, but rather the access to the features of the underlying bond.

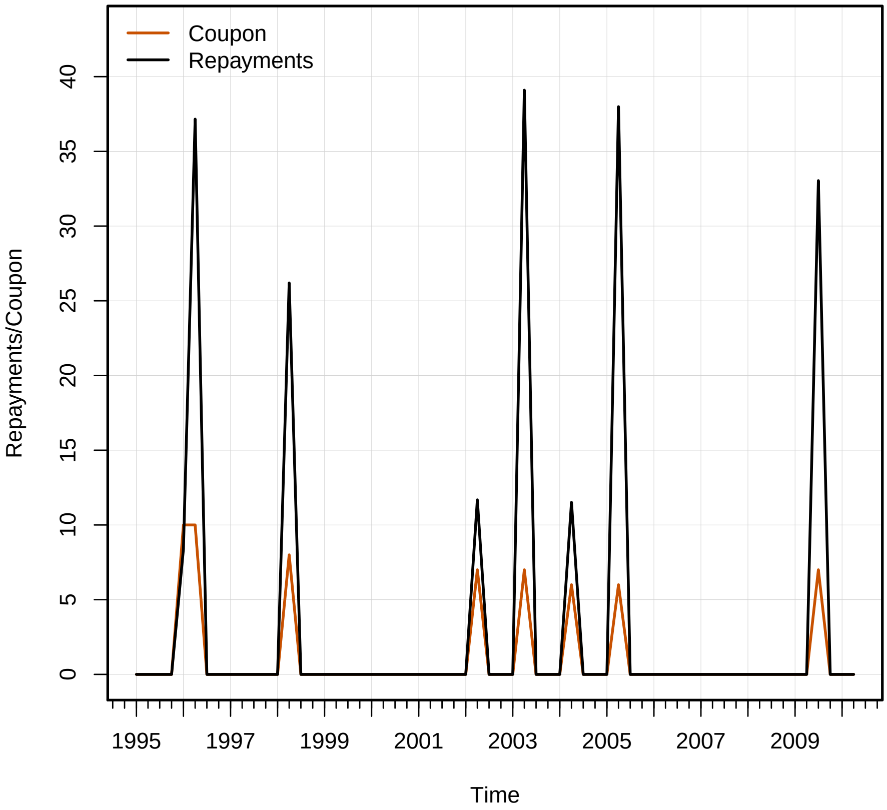


Figure 3: The graph shows how many percents of a given loan have been refinance on the dates that the refinancing rules of thumb dictated that one holding such loans should do so.

Figure 3 shows that historically observed refinancing activities have been extraordinarily high at around the dates for which the rules of thumb have suggested that borrowers should refinance. The graph indicates that the refinancing rules of thumb (at least the one for refinancing from a higher to a lower coupons) are indeed known to borrowers and that they act upon them.

Refinancing from fixed rated to adjustable rate:

This is normally done as a means of reducing the payments or as an alternative to refinancing to higher coupons. There is however no rules of thumb governing this type of refinancing. The advisors are reluctant to advise borrowers to refinance to adjustable rate mortgages since the risk profile of the borrower is thereby changed. Regardless of this fact over half of all mortgage loans in Denmark are adjustable rate mortgages with annual adjustments - F1 loans.

Next section describes a model-based alternative for generating personal refinancing strategies.

3 Optimization Model

The main purpose of this paper is to test the performance of the refinancing rules of thumb described in the last section. In order to do that a few alternatives to the rules of thumbs are considered. One obvious alternative is to do nothing all, a so called issue and hold strategy, where the borrower just chooses a loan to begin with and stay with the loan for a the whole period until the loan is redeemed due to other reasons than refinancing, for example selling the house.

Another, very unrealistic, alternative is to refinance at all the right times, so that the maximum refinancing potential is gained. Obviously this alternative is very unlikely to be realized in real life, since it requires full foresight on future information on coupon and price movements of the underlying bonds. Nevertheless this is a useful benchmark which provides an upper bound for the refinancing potential. If this potential is not much more than what is gained by following the rules of thumb or by not refinancing at all, there will be then no reason to try to develop more efficient methods for refinancing.

In this section therefore, a deterministic optimization model with full foresight is developed first. This model is referred to as the crystal ball benchmark. After that, the full foresight assumption is relaxed and a stochastic model is developed to generate realistic ex-ante refinancing strategies.

3.1 The crystal ball benchmark model

The model explained in this section finds the optimal refinancing strategy in the presence of full information about the future. Basically historical data on bonds are used to find the strategy that minimizes the borrowers total borrowing costs for a given period.

The reason why this is not a trivial problem is the costs associated with refinancing. Especially the existence of the fixed transaction costs (over 1000 EURO per refinancing) makes this into a combinatorial optimization problem.

The crystal ball benchmark model is described in the following. The model is inspired by the formulations given in Nielsen & Poulsen (2004), Rasmussen & Clausen (2004) and Rasmussen & Zenios (2007).

The crystal ball information is given on a time line with time points t, where refinancing is allowed. It is assumed that the model may only refinance 4 times a year (1/1, 1/4, 1/7 and 1/10).

There are two horizons to consider in the problem, one is the loans maturity T and the other is the analysis period τ. It is always true that $\tau \leq T$.

Input data:

At any given time point t, the following information on the any loan i is given:

d_t, discount factor at time t,

r_{ti}, coupon rate for loan i at time t,

P_{ti}, origination price for loan i at time t,

K_{ti}, redemption price for loan i at time t. The distinction between the origination price P_{ti} and the redemption price K_{ti} is due to the call option embedded in the Danish mortgage bonds. The following relation holds between the two prices: $K_{ti} = \min\{1, P_{ti}\}$ for the fixed rate callable bonds and $K_{ti} = P_{ti}$ for the adjustable non-callable bonds,

γ_t, tax reduction rate, given as a percentage of negative capital income,

c_a, administration costs, given as a percentage of the loans outstanding debt,

c, variable transaction cost, given as a percentage of the loans outstanding debt,

c_f, fixed transaction costs associated with a refinancing,

M, the big M constant, used when a binary variable is set to 1 in the case of refinancing, ensuring that the fixed costs are incurred.

Variables:

z_{ti}, outstanding debt for loan i at time t,

y_{ti}, amount of originated bonds for funding loan i at time t,
x_{ti}, amount of redeemed bonds for repaying loan i at time t,
$Z_{ti} = \begin{cases} 1, & \text{If loan } i \text{ should be originated or redeemed at time } t, \\ 0 & \text{otherwise,} \end{cases}$
A_{ti}, principal payments for loan i at time t,
F_t, total loan payment at time t,
R_τ, value of outstanding debt (adjusted for bond price) at horizon τ.

The crystal ball model is then defined as follows:

Object function:

The objective is to minimize the present value of the total loan costs in the given analysis period:

$$\min \quad \Big[\sum_{t=1}^{\tau} d_t F_t\Big] + d_\tau R_\tau \tag{1}$$

Initialization:

To begin with there should be originated enough bonds to fund the needed cash V_0 by the borrower and to pay the costs associated with the origination:

$$\sum_{i \in U} P_{0i} y_{0i} \geq V_0 + \sum_{i \in U} \Big(c y_{0i} + c_f Z_{0i}\Big), \tag{2}$$

In equation (3) the outstanding debt is set to be equal to originated amount of bonds at the beginning:

$$z_{0i} = y_{0i}, \qquad \text{for all } i \in U. \tag{3}$$

Balance equations:

Equation (4) is a balance equation, which sets the outstanding debt for loan i at time t equal to the outstanding debt from time

$t-1$ less the principal payment in case of no refinancing. If a refinancing takes place the equation initializes the outstanding debt in the new loan, and deduces the redeemed amount from the old loan:

$$z_{ti} = z_{t-1,i} - A_{ti} - x_{ti} + y_{ti}, \qquad \text{for all } i \in U, t = 1, \ldots, \tau. \tag{4}$$

Equation (5) ensures that the redeemed amount and the costs of refinancing come from originating new loan:

$$\sum_{i \in U} (P_{ti} y_{ti}) = \sum_{i \in U} \big(K_{ti} x_{ti} + c(y_{ti} + x_{ti}) + c_f Z_{ti}\big), \qquad t = 1, \ldots, \tau - 1. \tag{5}$$

Definition of loan payments (annuities):

Principal payments are defined in equation (6):

$$A_{ti} = z_{t-1,i} \left[\frac{r_{t-1,i}(1 + r_{t-1,i})^{-T+t-1}}{1 - (1 + r_{t-1,i})^{-T+t-1}} \right], \quad \text{for all } i \in U, t = 1, \ldots, \tau. \tag{6}$$

The total loan payment at time t is then defined in equation (7) as the sum of principal payments as well as interest payments and administration cost. The last two payments deduced by the amount which is given as a tax rebate. The total repayment cost at the end of the analysis period is given in equation (8).

$$F_t = \sum_{i \in U} \big(A_{ti} + (1 - \gamma)(r_{t-1,i} + c_a) z_{t-1,i}\big), \qquad t = 1, \ldots, \tau. \tag{7}$$

$$R_\tau = \sum_{i \in U} (z_{\tau i} K_{\tau i}) \tag{8}$$

Fixed costs:
The following equation ensures that the binary variable Z_{ti} is set to 1 when there is refinancing. If $Z_{ti} = 1$ then the fixed costs are counted as part of the total cost of refinancing.

$$MZ_{ti} - y_{ti} \geq 0, \qquad \text{for all } i \in U, t = 0, \ldots, \tau. \tag{9}$$

And finally the non-negativity constraints are defined as follows:

$$z_{ti},\ y_{ti},\ x_{ti} \geq 0\ , Z_{ti} \in \{0, 1\} \qquad \text{for all } i \in U, t = 0, \ldots, \tau. \tag{10}$$

The results of the above model are studied and compared to the other strategies in section 4, but before that a realistic model for generating personal refinancing strategies is introduced in the following.

3.2 Model-based refinancing strategies

The optimization model which is used in this section is a variation of the model introduced in Rasmussen & Zenios (2007). It is not the model, however, which is the focus of this paper, but rather the test framework in which the model is used.

If the results of any model, in this case refinancing strategies, are to be used in real world, the creators of the model need to substantiate that the strategies indeed will work better than the existing practice, in this case the refinancing rules of thumb. It is not justifiable to advise borrowers to follow a new set of refinancing strategies, before these strategies are tested in a simulated framework and before enough insight is gained to justify trying the strategies in the real world. This paper describes the simulation framework for testing the model-based refinancing strategies and discusses the insights gained.

3.2.1 The idea behind the model

The optimal refinancing model resembles, in its overall structure, the crystal ball model. There is one main difference though, no information about future is given, instead a number of scenarios are generated.

The optimization model is equipped with the following information:

1. One loan or a portfolio of loans at the beginning of the period.
2. Information about costs related to refinancing.
3. A number of alternative loans which the borrower may refinance the existing loan with.
4. Information about cost of the alternative loans in 200 different scenarios.

The scenario generation process is not the topic of this paper. It is described in Rasmussen, Madsen & Poulsen (2011). The focus here is rather on the differences between model-based and rule of thumb based refinancing advice. The superiority of the model-based refinancing advice on a conceptual level may be summarized as follows:

- Time horizon is explicitly taken into account.
- Risk of refinancing is both measurable and manageable. Risk is defined as the average of the α % highest total period costs across the given scenarios.
- Mixing loans is allowed.
- The strategies are personal - borrower specific input is taken into account.

The simulation framework is made of two sets of scenarios. The first set of scenarios are to be considered as alternative historical data. These are coherent alterations to historical data for a given period - 2000 to 2010 for this paper. There are 250 of them analyzed in this paper corresponding to performing back tests on 250 sets of historical data with a length of 10 years each. The reason for simulating historical data is to avoid only performing back tests on one single period. So alternative likely "historical" data are generated for interest rates and bond prices.

The second set of scenarios are those used for the optimization model. They are generated along each of the 250 "historical" scenarios in steps of a quarter of a year, allowing the optimization model to come up with refinancing suggestions along the way. There are 200 of these scenarios at each step and they all have a horizon at 2010. Figure 4 demonstrates the analysis setup by depicting only one of the "historical" scenarios and three of the optimization scenario sets for three randomly picked quarters.

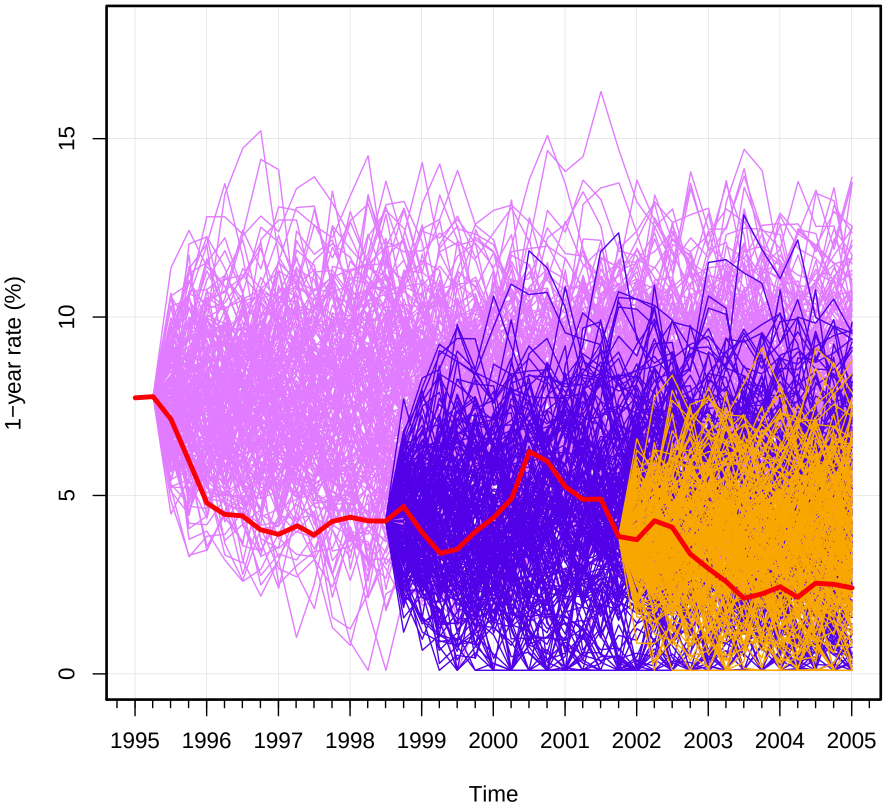


Figure 4: In each quarter there is a scenario tree of interest rates and bond prices. Three examples (out of 40) of such scenario trees are shown here for a simulation over the 10 year rate for the period 1995- 2005. The graphs is meant to illustrate the point with the simulation optimization framework only.

3.2.2 Measuring risk

The risk of refinancing is modeled here as the average of the 5 % highest total period costs for any given optimization scenario set.

This risk measure is known as Conditional Value at Risk (CVaR) illustrated in Figure 5. CVaR is a coherent risk measure (see

Artzner et al. (1999)) and in an optimization setting can be formulated linearly (see Rockafellar & Uryasev (2000)). Both properties are violated by other well known risk measures such as Value at Risk and Variance. Besides CVaR measure the tail risk - the idea is that if total costs after a given refinancing are lower than total costs of staying in the existing loan for the tail of the scenario set, the refinancing may be considered as robust - beneficial almost no matter what happens. That is obviously a function of how reasonable the scenarios are built - a topic which will be studied in section 4.

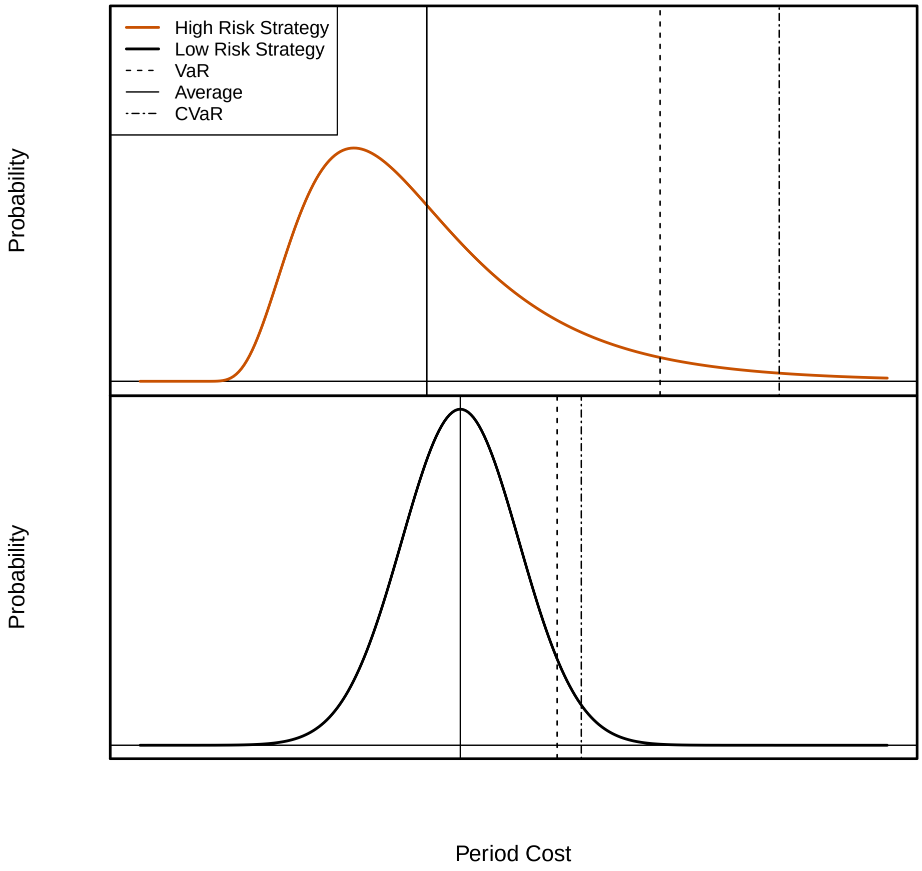


Figure 5: An illustration of the two risk measures CVaR and VaR. VaR is the biggest loss that can happen with a given probability. CVaR expresses the average loss for the scenarios where a loss higher than VaR is materialized.

3.3 The mathematical formulation

The model for generating model-based refinancing strategies is described in the following:

The model is a single period (two-stage) stochastic program where the scenario structure is defined on a starting point t_0, and a set of scenarios s. At any given starting point t_0 (start of a new quarter) there are set of loans i. The model is given an existing loan that it might either maintain or refinance to another loan.

Scenario data:

λ, risk weight, has values between 0 and 1. For $\lambda = 0$, the risk neutral model, the average of total costs is minimized. For $\lambda = 1$ the CVaR of period costs is minimized.

$P_{t_0,i}$, origination price for loan i at time t_0,

$K_{t_0,i}$, redemption price for loan i at time t_0,

$I_{t_0,i}$, outstanding debt for the existing loan i at time t_0.

p^s, probability of scenario s,

O_i^s, total period costs for loan i at scenario s, for an initial outstanding debt of 1,

c, variable transaction costs as a percentage of the loans outstanding debt,

c_f, fixed transaction cost in connection with a refinancing,

M, The Big M constant.

Variables:

$z_{t_0,i}$, outstanding debt for loan i at the starting point t_0,

$y_{t_0,i}$, amount of originated bonds for funding loan i at the starting point t_0,

$x_{t_0,i}$, amount of redeemed bonds for repaying loan i at the starting point t_0,

$$Z_{t_0,i} = \begin{cases} 1, & \text{If loan } i \text{ should be issued at the starting point } t_0, \\ 0 & \text{otherwise,} \end{cases}$$

ξ, Value-at-Risk (VaR) for a $100\alpha\%$ confidence level,

$CVaR(y;\alpha)$, Conditional Value-at-Risk for a portfolio of loans $y = (y_{t_0,i})$ for a $100\alpha\%$ confidence level,

ξ_+^s, the amount of the total costs on top of ξ for a scenario s. If the total costs do not exceed ξ, the value of ξ_+^s will be zero.

The optimization model is then defined as follows:

Object function:

The objective of the optimization model is to minimize the average period cost of the loan for a predetermined risk level, defined as a weight on CVaR:

$$\min \quad (1-\lambda)\Big[\sum_i \sum_s p^s y_{t_0,i} O_i^s\Big] + \lambda CVaR(y;\alpha) \tag{11}$$

Refinancing equation:
Equation (12) and (13) together ensure that refinancing can take place at time t_0:

$$z_{t_0,i} = I_{t_0,i} - x_{t_0,i} + y_{t_0,i}, \qquad \text{for all } i \in U. \tag{12}$$

$$\sum_{i \in U} (P_{t_0,i} y_{t_0,i}) = \sum_{i \in U} \big(K_{t_0,i} x_{t_0,i} + c(y_{t_0,i} + x_{t_0,i}) + c_f Z_{t_0,i}\big). \tag{13}$$

Equation (12) sets the outstanding debt for loan i equal to the amount of existing debt in case of no refinancing. If there is a refinancing the existing loan must be repaid. In this case equation (13) ensures that the repayment is funded by originating a new loan. The outstanding debt for the new loan is registered then in equation (12).

Fixed costs:
The following constraint makes sure, that the binary variable $Z_{t_0,i}$ is set to 1, if there is a refinancing. If $Z_{t_0,i} = 1$ then the fixed costs are entered as part of the total costs at the refinancing.

$$MZ_{t_0,i} - y_{t_0,i} \geq 0, \qquad \text{for all } i \in U. \tag{14}$$

CVaR constraints:
The two constraints 15 and 16 combined define CVaR for period costs. The linear formulation of CVaR in a CVaR minimizing context is due to Rockafellar & Uryasev (2000).

$$\xi_+^s \geq \Big[\sum_i y_{t_0,i} O_i^s\Big] - \xi, \qquad \text{for all } s \in S. \tag{15}$$

$$CVaR(y;\alpha) = \xi + \frac{\sum_s p^s \xi_+^s}{1-\alpha} \tag{16}$$

Non-negativity constraints:
Finally the non-negativity constraints are given as:

$$y_{t_0,i},\ x_{t_0,i},\ \xi_+^s \geq 0\ , Z_{t_0,i} \in \{0,1\} \ \text{ for all } i \in U. \tag{17}$$

With the model framework at hand the results of different strategies and benchmarks are compared in the next section.

4 Results

In this section, the results from applying the rules of thumb, the crystal ball benchmark and the model-based strategies are compared for the 250 simulated scenarios (including the actual historical data as one of the scenarios) for the period 2000-2010 are studied for a loan with an initial value of 2.000.000 DKK (approximately 270.000 EUR).

First these strategies are compared for the actual historical data. Figure 6 shows the development in outstanding debt, loan payments and the period's effective loan costs in percent (PEL) for the three different strategies in the top three graphs. The bottom graph shows the strategy suggested by the stochastic programming model of the last section for a risk averse borrower.

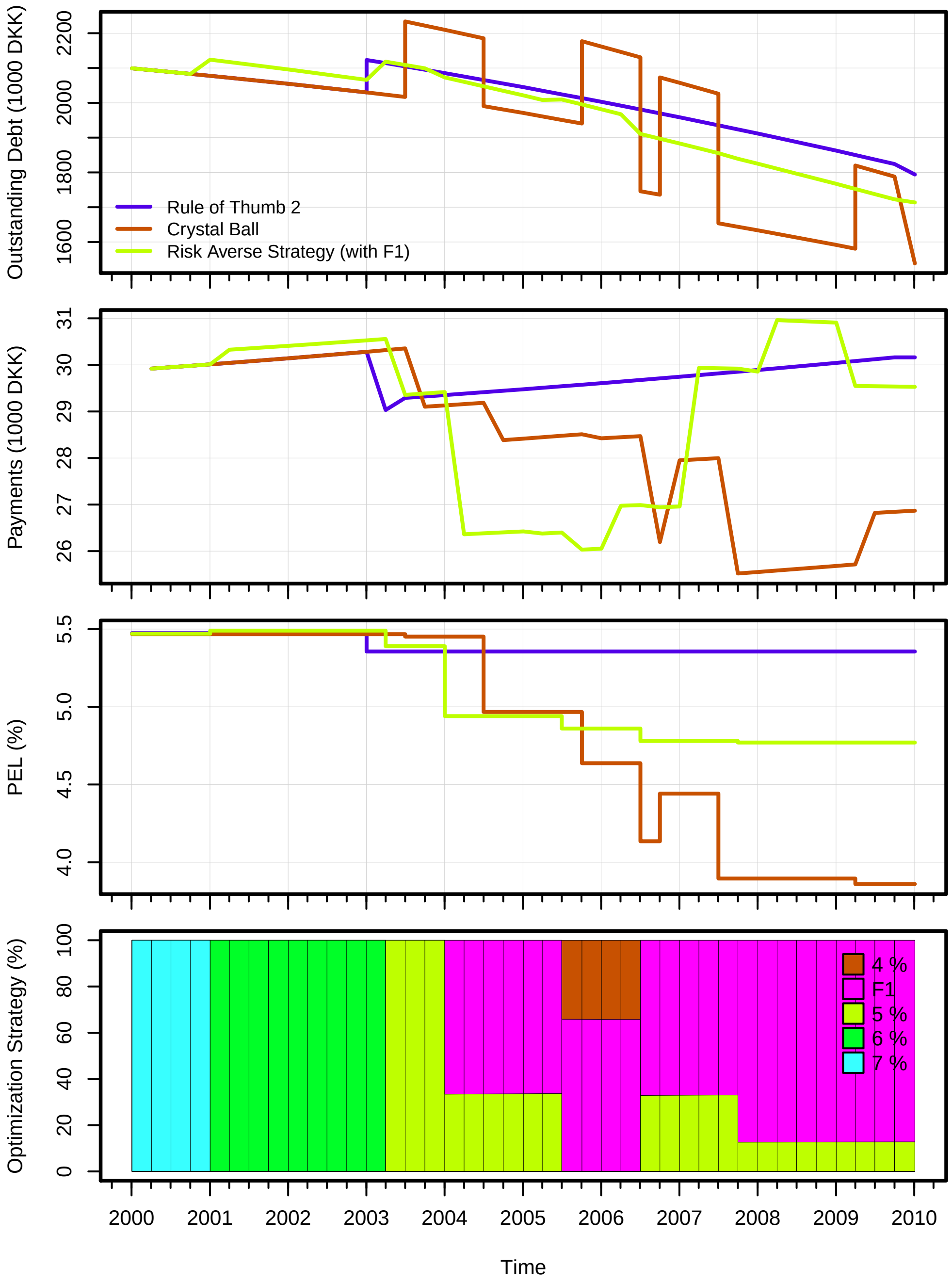


Figure 6: The development in outstanding debt, payments, and the period costs of the loan in percent are given here for the historical data from the period 2000-2010. The performance of the extended rule of thumb, the risk-averse model-based strategy and the crystal ball benchmark are compared.

Note that the rule of thumb is denoted "rule of thumb 2". The existing market rules as described in section 2 are extended here with an extra condition ensuring that the total cost of the new loan will not be higher than the total cost of the existing loan for the given period. This enhances the result of the rules of thumb for almost all simulated scenarios and makes the comparison with the model-based strategies more fair, since many advisor use one or other kind of break-even rule on top of the given rules of thumb. Such rules enforce a minimum upfront reduction in payments or outstanding debt to justify the refinancing.

Note also that the result from the crystal ball model does not consider F1 loan. The optimization is done among fixed rate mortgage loans only. The maximum potential for refinancing gain is thereby even bigger when both types of loans are considered.

The observations made based on these results are the following:

1. The potential refinancing gain is sizable.
2. A risk-averse (CVaR minimizing) model-based strategy seems to easily outperform the rules of thumb.
3. The model refinances more often. It both takes advantage of mixing loans as well as refinancing when the interest rates increase, in order to reduce outstanding debt

Next the above strategies are tested for all the simulated scenarios to see whether or not the results gained in Figure 6 are robust across scenarios with varying interest rates.

Figure 7 shows how the various strategies are performing compared with the crystal ball-benchmark and rule of thumb 2. The refinancing gain of any strategy is made relative to the no-refinancing strategy. The gain for the risk-neutral strategy (with F1 included in the loan universe), minimizing average total costs, is compared with that of the risk-averse strategy (not including F1 in the loan universe), minimizing the CVaR of total costs.

It is clear that the risk-averse strategy clearly outperforms rule of thumb 2 when it comes to generating risk-adjusted benefits to the borrowers. The left tails of the distributions represent the risk associated with the strategy. It appears that the risk-neutral strategy has a higher risk compared with rule of thumb 2. Based on that, and keeping in mind that the risk of refinancing should

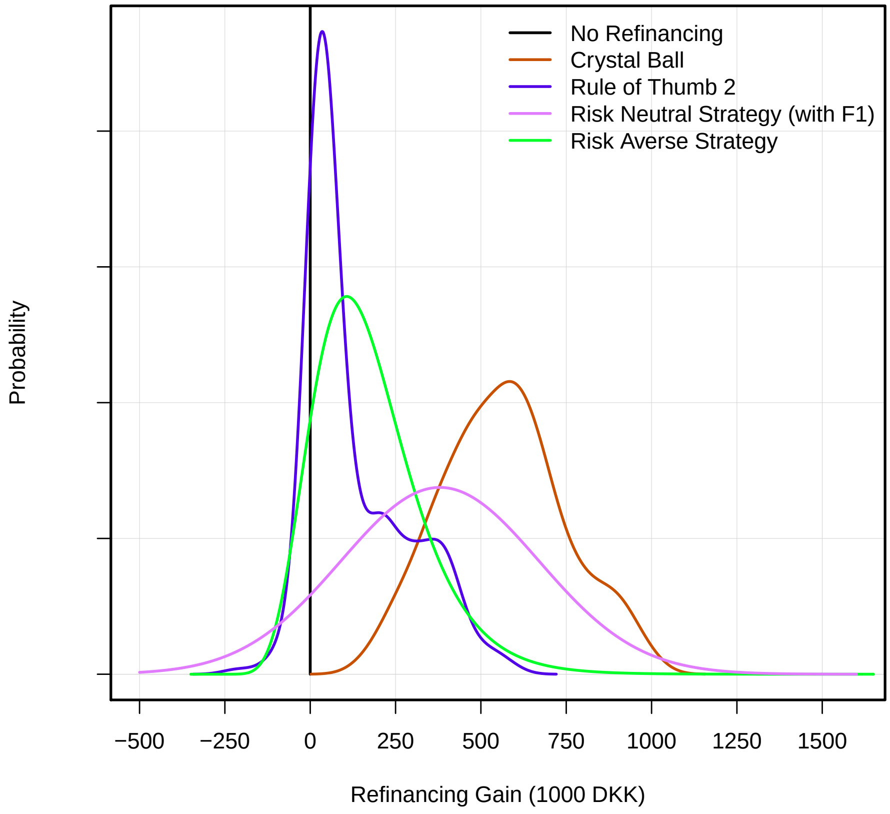


Figure 7: Comparison of different strategies across the simulated scenarios for the period 2000-2010.

not exceed that of the rules of thumb, the risk-neutral strategy is ruled out as a practical strategy for advising even though the strategy generates a nice average refinancing gain.

Now after having ruled out the risk-neutral strategy the two risk-averse strategies (with or without F1 in the loan universe) are compared in figure 8.

Due to diversification effects the risk-averse strategy including F1 in the loan universe, yields lower risk and higher return comparing to the risk-averse strategy not including F1.

The observations based on these results may be summarized

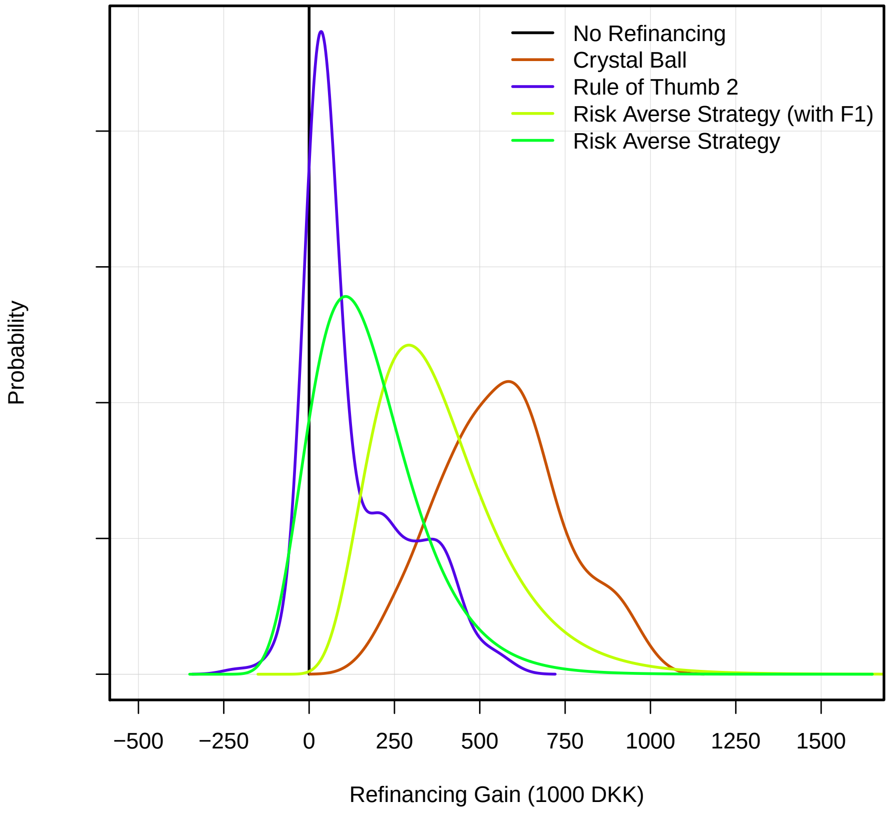


Figure 8: Comparison of the risk-averse strategies (with and without F1 in the loan universe) for the period 2000-2010.

as follows:

- Rules of thumb are very risk-averse and they outperform a no-refinancing strategy.
- The risk-averse model-based strategy clearly outperforms the existing refinancing rules of thumb.
- Using F1-loans, once used in a portfolio together with fixed rate loans, not only reduces the total costs, but also reduces the risk of refinancing.

5 Conclusion

Jørgensen et al. (1996) argued in their paper from 1996 that relatively simple claims on refinancing gain (such as those formulated in the refinancing rules of thumb) lead to close to optimal decisions.[3] They do not, however, take into account borrowers preferences with respect to horizon and risk aversion. What is more, several new types of loan have been introduced, most notably the adjustable rate loans, and that alone asks for new strategies and risk management in connection with the borrowers mortgage choice and refinancing.

New methods for refinancing should take horizon and risk explicitly into consideration. Once doing so, more refinancing particularly when the interest rates go up, as well as mixing fixed rate and adjustable rate loans in a portfolio, will enhance the result of refinancing without risk.

Based on the results of the optimization and simulation framework studied in this paper a new paradigm for mortgage choice and refinancing consultancy is introduced. The new paradigm may be summarized as follows:

- Risk - risk of refinancing should be measured and managed.
- Loan portfolios - instead of individual loans.
- Loan horizon and its inherent uncertainty should be considered.
- Focus on cost of borrowing - instead of presenting technical details of loans.
- Responsiveness to change - a refinancing strategy should react to the observed changes of interest rates and bond prices.

The method framework introduced in this paper has the potential to be used as the analytical backbone for this new consultancy paradigm. In particular the results suggest so far that:

- Risk-averse model-based refinancing strategies yield good and robust results.

---

[3] The word "optimal" should be understood in an option pricing setting in a complete market.

- Both fixed rate and adjustable rate mortgages should be used in a well diversified portfolio of loans. That increase the chance of successful refinancing.
- It is important to look at the long term consequences of a refinancing, in particular more than 5 years.
- Too many refinancings will drain the positive effect of refinancing. Timing is important. Few but well analyzed refinancings will almost certainly yield a better result than not refinancing at all.
- Rules of thumb are outdated and do not live up to the new consultancy paradigm.

There is room for much more analysis in the Danish as well as other mortgage markets with the aim of enhancing consultancy services when it comes to mortgage choice of refinancing. Many more products exist than the two studied in this paper. However the analysis framework of this paper may be reused when considering other loan types or market specific information.

Finally the purpose of using any model is gaining intuition. Once the intuitions behind a model are well understood and formulated in a coherent manner, the model itself may be left out and the intuitions applied directly to the problem. Such a study will be the subject of future work.

References

Artzner, P., Delbaen, F., Eber, J. & Heath, D. (1999), 'Coherent measures of risk', *Mathematical Finance* **9**, 203–228.

Jørgensen, P. L., Miltersen, K. R. & Sørensen, C. (1996), 'En sammenligning af konverteringsstrategier for konverterbare realkreditlån', *Finans Invest* (7), 22–29.

Nielsen, S. S. & Poulsen, R. (2004), 'A Two–Factor, Stochastic Programming Model of Danish Mortgage–Backed Securities', *Journal of Economic Dynamics and Control* **28**, 1267–1289.

Rasmussen, K. M. & Clausen, J. (2004), 'Mortgage Loan Portfolio Optimization Using Multi Stage Stochastic Programming', *Journal of Economic Dynamics and Control* **31**, 742–766.

Rasmussen, K. M., Madsen, C. A. & Poulsen, R. (2011), *Realkreditrådgivning - Et studie af danskernes valg af realkreditlån og konverteringspraksis*, Boligøkonomiks videncenter.

Rasmussen, K. M. & Zenios, S. A. (2007), 'Well ARMed and FiRM: Diversification of mortgage loans for homeowners', *The Journal of Risk* **Fall**, 1–18.

Rockafellar, R. T. & Uryasev, S. (2000), 'Optimization of conditional Value-at-Risk', *The Journal of Risk* **2**, 21–41.

Part III

Chapter 7

Conclusions

Chapter 7: Conclusions

Paper 1: The normal class of arbitrage-free spot-rate models

Two of the main results in this paper were: First we showed how to determine the T-forward adjusted risk-measure using the concept of fundamental solution to linear PDEs. That is, it turned out to be possible to derive the T-forward adjusted risk-measure without the use of Girsanovs theorem, which is the traditional approach, see for example Karoui, Myneni and Viswanathan (1993).

The other main result was that we were able to carry the analysis of the Quadratic interest rate model further than Jamshidian (1996) by relying on the Fourier transformation in order to obtain the fundamental solution for the PDE.

In that connection we showed how it was possible to derive the price of a discount bond and the price of an option on a discount bond. This was also as an example done for the extended Vasicek model. Using the idea from Hull and White (1990), we showed how it was possible to fit the model to the initial term structure.

We designed a special discrete time model for the Quadratic interest rate model, as it in some cases was not possible to use the trinomial approach from Hull and White (1994).

After that we focused on pricing techniques for path-dependent interest rate contingent claims. We focused in that connection on Monte Carlo simulation of spot-rate models with a time-dependent drift. The third main result in the paper was to introduce a forward-induction technique that made it possible to constrain the Monte Carlo simulation for the matching of the initial yield-curve.

When using Monte Carlo simulation for time-dependent spot-rate models, we encountered problems obtaining variance reduction - by methods that in the case of time-homogenous spot-rate models are very effective. Further research is obviously needed in connection with Monte Carlo simulation of time-dependent spot-rate models.

Finally we compared lattice-based pricing method to Monte Carlo simulation procedures for the pricing of European put-and call-options on zero-coupon bonds for both the Extended Vasicek model and the Quadratic interest rate model.

In that connection we concluded:

- Markovian spot-rate models are best represented by a method that takes into account the Markovian nature in the process - that is lattice-based methods
- Monte Carlo are not as efficient as lattice-based procedures for time-dependent Markovian spot-rate models, as Monte Carlo methods do not take into account the Markovian nature - Monte Carlo is doing precisely the opposite - as Monte Carlo methods by nature are non-Markovian.

With respect to the pricing of path-dependent claims, no clear conclusion was drawn. It was however suggested that before deciding on a method (Monte Carlo or lattice-based method), it was of importance to separate the problem, as follows:

- How do we most efficiently construct the time-space dimension of the interest rate process?.
- Given that, how do we most efficiently utilize the constructed time-space dimension to price a path-dependent contingent claim?

From this we deduce that, even though Monte Carlo is the natural method to use when pricing path-dependent interest rate contingent claims, it might not be the most efficient one - at least not when the spot-rate is Markovian.

As of now all work in connection with the multi-factor Quadratic interest model has been under the assumption of constant parameters in the SDEs, for example Laurance and Wang (2005) , so an interesting topic for future work (and a natural extension) is a multi-factor version of the model setup presented here and how to efficiently price derivatives in this more complex model setting.

Another line of research is how to construct – if possible – a market model, along the lines of Brace, Gatarek and Musiela (BGM) (1997).

Paper 2: Empirical Yield-Curve Dynamics, Scenario Símulation and Risk-Measures

The first result in this paper was the construction of a general model for the variation in the term structure of interest rates - that is we defined a model for the shift function. In this connection we showed - using the Heath, Jarrow and Morton (1991) framework - that the shift function could be understood as a volatility structure - more precisely the spot rate volatility structure.

The class of shift functions considered in this paper was of the linear type, with independence between the individual factors; the model was therefore comparable to the Ross (1976) APT model.

Using PCA we showed that it took a 4-factor model to explain the variation in the term structure of interest rates over the period from the 2. January 2002 to the 2. January 2012. These 4 factors can be called a Slope factor, a Short-Curvature Factor, a Short Factor and a Curvature Factor.

Because of the relationship between the volatility structure in the Heath, Jarrow and Morton framework and the shift function implied from our empirical analysis of the evolution in the yield-curve we concluded that PCA could be used to determine the volatility structure in the Heath, Jarrow and Morton framework - a non-parametric approach.

We then extended this model setup by adding stochastic volatility. This we did by assuming that the stochastic volatility was driven by a GARCH(1,1) model. Due to the fact that our stochastic volatility process was independent of our yield-curve factor-model, means that the extended model belonged to the class of USV (unspanned stochastic volatility) models.

In the last part of the paper we turned our attention to Risk model. Our approach to the calculation of VaR/ETL was a scenario simulation based methodology with relied on the framework of Jamshidian and Zhu (1997). This scenario simulation procedure builds on the factor loadings derived from a PCA of the same kind we used in our analysis of the empirical dynamics in the yield-curve.

The general idea behind the scenario simulation procedure is to limit the number of portfolio evaluations by using the factor loadings derived in the first part of paper and then specify particular intervals for the Monte Carlo simulated random numbers and assign appropriate probabilities to these intervals (states).

From our analyses of both straight bonds and Danish MBBs using the scenario simulation procedure we conclude the following:

- The scenario simulation procedure is a computational effective alternative to Monte Carlo simulation
- The scenario simulation procedure is capable of producing reasonable good approximations of the probability distributions with a limited number of states
- There is much better control over the extreme values in the scenario simulation than in Monte Carlo simulation
- The scenario simulation procedure is more efficient than the parametric approach for "linear" securities and it rivals the parametric approach for "linear" securities because of the speed of calculation
- We suggested using 7 x 5 x 3 x 3 states for the scenario simulation. Because we "only" need to perform 315 re-evaluations of the portfolio the scenario simulation procedure is feasible for large portfolios consisting of highly complex non-linear securities
- One last important feature is that we in the spirit of Barone-Adesi, Bourgoin and Giannopoulos (1997) extended the scenario simulation procedure to include stochastic volatility - in our case we showed that a Garch(1,1) was appropriate

We also backtested our Risk-Model setup for the year 2008 on 2 bonds - a Government-Bond and a Mortgage-Bond. The overall conclusion for this was:

- Our Risk-Model setup was able to capture the extreme movements influenced by the financial crisis for our Government-Bond case - notable the autumn of 2008
- Our Risk-Model setup did not manage to completely capture the extreme price movements for our Mortgage-Bond. Due to the fact that the reason for that was mostly (if not entirely) because of the enormous changes on level of OAS, we found the model indeed very promising - both of term of efficient from a calculation point of view, to its ability to capture the Risk of Mortgage-Bonds

More work is of course necessary - both with respect to backtesting the model and with respect to the determination of the input to the scenario simulation procedure.

In connection with the inputs to the scenario simulation method we especially need to address the following:

- How to forecast the volatility?
    - In this paper we argued that GARCH is a logical method to utilize
- How to select the expected yield-curve at a given time-horizon?
    - In this paper we suggested using the initial yield-curve as the expected yield-curve
- Would it be possible to make our OAS-Model even more robust?

    - More examination on the appropriate number of states that is used in the scenario simulation method

Paper 3: Can Danish households benefit from Stochastic Programming models? – An empirical study of mortgage refinancing in Denmark

Jørgensen et al. (1996) argued in their paper from 1996 that relatively simple claims on refinancing gain (such as those formulated in the refinancing rules of thumb) lead to close to optimal decisions. They do not, however, take into account borrower preferences with respect to horizon and risk aversion. Moreover, several new types of loan have been introduced, most notably the adjustable rate loans, and that alone asks for new strategies and risk management in connection with the borrowers mortgage choice and refinancing.

New methods for refinancing should take horizon and risk explicitly into consideration. Once doing so, more refinancing particularly when the interest rates go up, as well as mixing fixed rate and adjustable rate loans in a portfolio, will enhance the result of refinancing without risk.

Based on the results of the optimization and simulation framework studied in this paper a new paradigm for mortgage choice and refinancing consultancy is introduced. The new paradigm may be summarized as follows:

- Risk, that is the risk of refinancing should be measured and managed
- Loan portfolios, instead of individual loans. The key-word here is diversification
- Loan horizon and its inherent uncertainty should be considered
- Focus on cost of borrowing instead of presenting technical details of loans.
- Responsiveness to change. What we mean is that a refinancing strategy should react to the observed changes of interest rates and bond prices.

The method framework introduced in this paper has the potential to be used as the analytical backbone for this new consultancy paradigm. In particular the results suggest so far that:

- Risk-averse model-based refinancing strategies yield good and robust results
- Both fixed rate and adjustable rate mortgages should be used in a well diversified portfolio of loans. That increase the chance of successful refinancing
- It is important to look at the long term consequences of a refinancing, in particular more than 5 years
- Too many refinancings will drain the positive effect of refinancing. Timing is important. Few but well analyzed refinancings will almost certainly yield a better result than not refinancing at all
- Rules of thumb are outdated and do not live up to the new consultancy paradigm

There is room for much more analysis in the Danish as well as other mortgage markets with the aim of enhancing consultancy services when it comes to mortgage choice of refinancing. Many more products exist than the two studied in this paper. However the analysis framework of this paper may be reused when considering other loan types or market specific information.

The purpose of using any model is gaining intuition. Once the intuitions behind a model are well understood and formulated in a coherent manner, the model itself may be left out and the intuitions applied directly to the problem. Such a study will be the subject of future work.

Thesis Conclusion

In the 3 papers in this dissertation we have focused on 3 very interesting topics:

1. Modelling the yield-curve for the purpose of pricing interest rate derivatives
2. How to make “VaR” to work
3. Applying modern optimization technique to support refinancing decisions

All papers have been written, with the Financial Manifesto in mind, and interesting and useful results have been obtained:

- In paper 1, I managed to carry the analysis of the Quadratic interest rate model further than Jamshidian (1996) by relying on the Fourier transformation in order to obtain the fundamental solution for the PDE. That the model was nearly as tractable as the Hull and White (1990) model was also shown. Lastly the pricing of interest rate derivatives was analysed, both analytical, using discrete time trees and Monte-Carlo
- In paper 2, I showed a method for calculating Risk for portfolios containing Danish Mortgage Bonds, using a combination of GARCH-modelling and focused scenario generation procedures that it was possible to build a Risk-Model that does work – even in extreme cases! – it was tested during the financial crisis
- In paper 3, we showed that applying a combination of an optimizing and scenario generation technique to the financial decision – When to refinance? was very promising.

References

Andersen L and Andreasen J (2000) “Volatility skews and extension of the LIBOR market model” Appl. Math. Finance 7, page 1–32

Artzner, Delbaen, Jean-Marc Eber, David Heath (2001) “Coherent Measures of Risk”, Mathematical Finance Volume 9, Issue 3, pages 203–228, July 1999

Bachelier, Louis (1900) "Theorie de la speculation," Annales Scientifiques de I'Ecole Normale Superieure, 3, Paris, Gauthier-Villars, 21-86

Bernstein, Peter L. (1998) “Against the Gods: The Remarkable Story of Risk”, Wiley (August 31, 1998)

Bhar, R. and Chiarella, C. (1997) "Transformation of Heath-Jarrow-Morton models to Markovian systems", The European Journal of Finance, 3(1), page 1-26

Black, Fischer and Myron Scholes (1973) "The Pricing of Options and Corporate Liabilities," Journal of Political Economy 81:3 (May/June), page 637-654

Black, Fischer (1976) “The pricing of commodity contracts”, Journal of Financial Economics, 3, page 167-179

Brace, A., Gatarek, D. et Musiela, M. (1997) “The Market Model of Interest Rate Dynamics”, Mathematical Finance, 7(2), page 127-154

Cox, Ingersoll og Ross (1985) "A Theory of the Term Structure of Interest Rates", Econometrica 53, page 363-385

Duffie og Kan (1996) "A Yield-Factor Model of Interest Rates", Mathemetical Finance, Vol. 6. No. 4 (October 1996), page 379-406

Fama, Eugene (1965) "The Behavior of Stock Prices," Journal of Business 38:1, page 34-105

Fenton, Normal and Neil, Martin (2012) “1.Risk Assessment and Decision Analysis with Bayesian Networks”, CRC Press; 1 edition (4 Dec 2012)

Fisher, Lawrence (1966) "An Algorithm for Finding Exact Rates of Return," Journal of Business , Vol. 39, No. 1 (January, 1966), page 111-118

Grove, M. (1966) “A Model of the Maturity Profile of the Balance Sheet”, Metroeconomica 18: page 40-55

Hagan, P., Kumar, D., Lesniewski, A., and Woodward, D (2002) “Managing smile risk”, Wilmott magazine (2002), page 84–108

Harrison, J. Michael and David M. Kreps (1979) “Martingales and arbitrage in multiperiod securities markets”, Journal of Economic Theory, 20, page 381-408

Haug and Taleb (2011) “Option traders use (very) sophisticated heuristics, never the Black–Scholes–Merton formula”, Journal of Economic Behavior and Organization 2011

Harrison, J. Michael and Stanley R. Pliska (1981) “Martingales and stochastic integrals in the theory of continuous trading”, Stochastic Processes and their Applications, 11 ,page 215-260

Harrison, J. Michael and Stanley R. Pliska (1983) “A stochastic calculus model of continuous trading: Complete markets”, Stochastic Processes and their Applications, 15, page 313-316

Heath, Jarrow and Morton (1990) "Contingent Claim Valuation with a Random Evolution of Interest Rates", The Review of Futures Markets 9, page 54-76

Heath, Jarrow and Morton (1992) “Bond Pricing and the Term Structure of Interest Rates: A New Methodology for Contingent Claims Valuation”, Econometrica 60, page 77-105

J.R. Hicks (1939, 2nd ed. 1946) “Value and Capital: An Inquiry into Some Fundamental Principles of Economic Theory”, Oxford: Clarendon Press

Holten, Glyn A. (2002) “The history of Value-at-Risk: 1922-1998”, working paper, july 25, 2002

Hopewell, Michael H. and George G. Kaufman (1973) "Bond Price Volatility and Term to Maturity: A Generalized Specification", American Economic Review, Vol. 63, No. 4 (September, 1973), page 749-753

Hull og White (1990b) "Pricing Interest-Rate-Derivative Securities", The Review of Financial Studies 3, pqge 573-592

Jamshidian (1991) “Forward Induction and Construction of YieldCurve Diffusion Models”, Journal of Fixed Income 1, page 62-74

Jarrow and Protter (2004) “A short history of stochastic integration and mathematical finance: The early years, 1880-1970”, in IMS Lecture Notes, volume 45 of IMS Lecture Notes, pages 75–91. IMS, 2004

Mark S Joshi and Riccardo Rebonato (2003) “A displaced-diffusion stochastic volatility LIBOR market model: motivation, definition and implementation”, Quantitative Finance, 3:6, 458-469

Langetieg (1980) "A Multivariate Model of the Term Structure", Journal of Finance 35, page 71-97

Macaulay, F. (1938) “The Movements of Interest Rates. Bond Yields and Stock Prices in the United States since 1856”, New York: National Bureau of Economic Research

Madsen (2009) “The 2008 Financial Crisis and ERM”, PRMIA Chapter Meeting Copenhagen, March 2009

http://www.prmia.org/Chapter_Pages/Data/Files/2973_3340_2008_Crisis_and_ERM_presentation.pdf

Markowitz, Harry (1952) “Portfolio Selection”, Journal of Finance 7, page 77-91

Markowitz, Harry (1999) “The early history of portfolio theory: 1600-1960”, Financial Analyst Journal 55, page 5-16

Merton, Robert C. (1973) "Theory of Rational Option Pricing", Bell Journal of Economics and Management Science (The RAND Corporation) 4 (1), page 141–183

Miltersen, K., Sandmann, K. et Sondermann, D., (1997) “Closed Form Solutions for Term Structure Derivates with Log-Normal Interest Rates“, Journal of Finance, 52(1), page 409-430

Morton (1988) "Arbitrage and Martingales", Ph.d. Dissertation Cornell University

Robert O’Harrow (2010) “Derivatives made simple, if that’s really possible”, Washington Post, April 21 2010

Pallaviciniy and Tarenghi (2010) “Interest-Rate Modeling with Multiple Yield Curves”, working paper Cornell University Library, 2010

Rady, S. and K. Sandmann (1994) “The Direct Approach to debt option pricing”, The Review of Futures Markets 13(2), page 461-514

Redington, F. M (1952) "Review of the Principles of Life-office Valuations", Journal of the Institute of Actuaries , Vol. 78 (1952), page 286-340

Ritchken and Sankarasubramanian (1995) “The importance of forward rate volatility structures in pricing interest rate-sensitive claims”, Journal of Derivatives (Fall), page 25-41

Samuelson, Paul A. (1945) “The effect of interest increases on the banking system”, The American Economic Review Vol. 35, no. 1, page 16-27

Samuelson, Paul A. (1965a) "Rational Theory of Warrant Pricing," Industrial Management Review 6:2, page 13-31

Samuelson, Paul A. (1965b) "Proof That Properly Anticipated Prices Fluctuate Randomly," Industrial Management Review 6:2, page 41-49

Sharpe, William F. (1964) "Capital Asset Prices: A Theory of Market Equilibrium Under Conditions of Risk," Journal of Finance 19:3, page 425-442

Sidenius (1997) “Examining Multi-Factor Market Models”, Global Derivatives 97, proceedings of the ICBI Conference, Paris, April 24-25, 1997

The Economist (2009) “Over the counter, out of sight”, The Economist, pages 93–96, November 14 2009

Vasicek, Oldrich (1977) "An Equilibrium Characterisation of the Term Structure", Journal of Financial Economics 5 (2): page 177–188

Wallas, G. E. (1960) "Immunization", Journal of the Institute of Actuaries Students' Societies 15 (1960), page 345-57

Weil, R. (1973) “Macaulay’s Duration: An Appreciation”, Journal of Business 46: page 589-592

Wilmott (2009) “Revenge of the Nerd”, in Newsweek (2009), see the following link: http://www.thedailybeast.com/newsweek/2009/05/28/revenge-of-the-nerd.html

Peter Zangari (1996) “RiskMetrics Technical Document”, J.P. Morgan 1996