CHALMERS | GÖTEBORG UNIVERSITY

MASTER THESIS

On Stochastic Volatility in Interest Rate Dynamics

DANIEL POLLERYD
NUTVADEE WONGTOSRAD

Department of Mathematical Sciences
CHALMERS UNIVERSITY OF TECHNOLOGY
GÖTEBORG UNIVERSITY
Göteborg, Sweden 2012

Thesis for the Degree of Master of Science

On Stochastic Volatility in Interest Rate Dynamics

Daniel Polleryd
Nutvadee Wongtosrad

CHALMERS | GÖTEBORG UNIVERSITY

Department of Mathematical Sciences
Chalmers University of Technology and Göteborg University
SE – 412 96 Göteborg, Sweden

Göteborg, May 2012

Abstract

In this Master Thesis we investigate the presence of stochastic volatility in interest rate dynamics and its effect on pricing interest rate derivatives. Different stochastic volatility models are compared to each other and to models where the volatility is constant. The models are calibrated to the short rate with the EMM procedure. The models are also calibrated to market data like the yield curve and swaption prices with the Kalman filter method and by minimizing the sum of squares. We find that stochastic volatility is very much present in interest rate dynamics, and that models are improved by adding a stochastic volatility process. The pricing capabilities of the models are on the other hand not improved by adding a stochastic volatility factor but instead make the models considerably slower.

Acknowledgements

We would like to thank our supervisor Patrik Albin for his help and encouragement, as well as for the interesting discussions we have had. We would also like to thank the many people in the mathematics building that have helped us during our thesis work. Last but not least, we would also like to thank our friends and families for their support.

Contents

1 Introduction

The interest rate is a factor that affects almost every individual on an economic level. For instance, the interest you receive on deposits in your bank account or the interest that you need to pay when you borrow money. The interest rate causes the amount of money to grow as the time goes by. If you are able to lend money out, it usually comes back with a reward which is the interest. This concept can also be described as the time value of money: An amount of money received today is worth more than the same amount of money received in the future. This kind of simple interest rate process is called the money-market account, which represents a risk-less investment. This process follows the dynamics of the stochastic differential equation (SDE),

$$dB(t) = B(t)r(t)dt, \quad B(0) = B_0,$$

where $B(t)$ is the money-market account and $r(t)$ is short rate which change according to time t. The solution this equation is

$$B(t) = B_0 e^{\int_0^t r(s)ds},$$

and the interest rate equation can also be approximated as

$$\frac{B(t + \Delta t) - B(t)}{B(t)} = r(t)\Delta t.$$

This equation shows that money grows in every time step. Moreover, Figure 1 illustrates how the interest rate behaves as a stochastic process by showing the U.S. Treasury 3 month yield.

Even if you have not taken a mortgage loan to finance your home, the rent you pay is still affected by the level of the interest rate. Larger businesses and corporations can be extensively affected from the change in interest rate, as they are usually financed by debt to some extent. To them, the interest rate level can mean the difference between a healthy profit or a loss. The interest rate is thus the source of a lot of risk within the company and has to be managed so that the company does not make a big loss, or even worse, faces bankruptcy. This is often done by the use of interest rate derivatives such as interest rate swaps, caps, floors or swaptions that are traded over-the-counter (i.e. not in an exchange).

To manage interest rate efficiently, one needs to know how the interest rate behaves. How likely is it that the interest rate will increase by a certain amount in the future, and what is the expected future level of the interest rate? These are questions that need to be answered in order to manage interest rate risk, and which depend mainly on economic factors. These questions can often be answered to some extent by interest rate models, that are calibrated to either historical data or to current market prices.

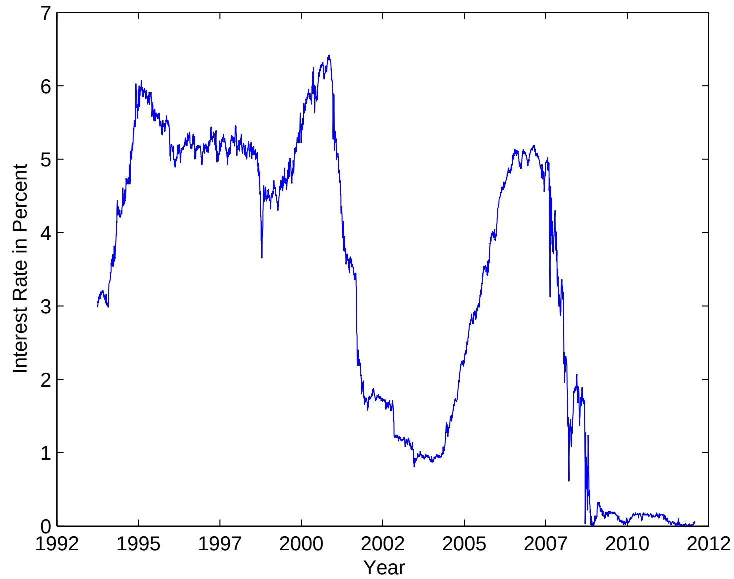


Figure 1: Three month U.S. Treasury yield from 1993 to 2012

The most common models for describing the interest rate dynamics are mathematical diffusion models. They assume that the interest rate is a stochastic diffusion process with a large diffusion component and a small (or sometimes non-existent) drift component. The most simple example of these models assume that the size of the diffusion component (commonly called volatility) is constant. Others assume that volatility is only related to the level of the interest rate, if the interest rate goes up so does volatility and vice versa, a so called level effect. These models are most often not accurate enough for interest rate modelling and pricing purposes. They often fail to capture different phenomena, like the extreme spike in volatility that occurred during the credit crunch, or the changing expectations about the future levels of the interest rate.

In this thesis we will aim to extend these models by modelling volatility as a stochastic process. By doing this we aim to increase the models accuracy in describing interest rate dynamics as well as their accuracy in pricing. The models we will look at are mainly stochastic volatility models for the short rate, but also the Heston model, which does not model the short rate but instead tries to model bonds and bond option dynamics directly.

2 Financial Theory

In this section we will discuss basic theory of interest rate derivatives and mathematical models of the interest rate. We will start with bonds and then move on to more complex products like swaps, caps, floors and swaptions.

2.1 Bonds

A bond is a financial debt security in which the issuer owes the the holder a debt. There are two main types of bonds, coupon bonds and zero coupon bonds. Coupon bonds pay the holder coupons of fixed size at specific times until maturity when the entire debt, known as the nominal or face value, is paid back to the holder. Zero coupon bonds as the name implies pay no coupons, the entire debt is repaid upon maturity. When a bond is issued and subsequently sold, it is traded like any other security and can change holders arbitrarily many times until maturity.

Bonds are issued by companies as well as governments in order to finance debt. Governments are by far the the largest group of issuers of bonds, the US as well as the Swedish government for example sell bonds every week. This makes government bonds very liquid which in turn means better price data.

The price of a zero coupon bond at time t with maturity at time T is $B(t,T) = e^{-Y(t,T)(T-t)}$ where $Y(t,T)$ is the yield to maturity. If we assume that there exists a bond with infinitesimal time to maturity, call this bond $B(t, t+\delta)$, then the limit $\lim\limits_{T\to t+} Y(t,T)$ exists, and is equal to the short rate, $r(t)$. The price of a bond can then be written as

$$B(t,T) = e^{-\int_t^T r(u)du}.$$

The problem here is that $r(u)$ is unknown for $u > t$ which means that $r(u)$ and the integral of $r(u)$ are random variables. To be able to price this Bond correctly we define

$$B(t) = B(0)e^{\int_0^t r(u)du},$$

as the value of a self financing portfolio that holds bonds $B(t, t+\delta)$ with infinitesimal time to maturity. Thus $B(t)$ also represents the value of a bank account that pays the short rate with continuous compounding. If under some probability measure Q the portfolio $X(t) = B(t,T)/B(t)$ is a Q-martingale, then according to the no arbitrage pricing formula (stated in the appendix), we have

$$\frac{B(t,T)}{B(t)} = E^Q\left[\frac{1}{B(T)}|\mathcal{F}_t\right],$$

which means that the price of a bond under the probability measure Q is

$$B(t,T) = E^Q\left[\frac{B(t)}{B(T)}|\mathcal{F}_t\right] = E^Q\left[e^{-\int_t^T r(u)du}|\mathcal{F}_t\right].$$

Here, $\mathcal{F}_t$ is a filtration at time t, which essentially can be seen as the amount of available information up until time t. Thus $E^Q[...|\mathcal{F}_t]$ means: The expectation of something given the information that is available at time t. To price a coupon bond, each coupon can be viewed as a separate zero coupon bond with maturity at the time of coupon payment. In this way the price of the coupon bond becomes the sum of the prices of zero coupon bonds plus the price of the notional end payment. In the above equations, the probability measure Q is very important, it is not the same as the real world measure. This is a different measure that is used to account for the risk averse behaviour of investors, and the current risk premium. Using the empirical probability measure would cause us to ignore the risk premium, yielding consistently higher prices than quoted in the market.

2.2 Interest Rate Caps and Floors

An interest rate cap is a financial product that puts an upper limit on the interest rate the buyer pays over a specific time period. At the end of each period, the buyer receives a payment such that the total interest paid on the notional is equal to or less than the strike rate, K. A floor is instead a lower limit on the interest rate received by the buyer. These caps and floors are useful in hedging when a borrower who is paying interest on a loan according to some interest rate, for example LIBOR (London Interbank Offered Rate) can protect himself against interest rate fluctuations by buying a cap at $x\%$ fixed interest rate, the strike rate. If the interest rate exceeds $x\%$ then he can use the amount he gained from the caps to pay for the interest as shown in Figure 2. Conversely, the interest rate floor is useful when a lender who is receiving LIBOR wanted to secure himself/herself if the interest happened to drop too low as shown in Figure 3.

Mathematically, the payments are received at specific times during the life of the cap, at which the holder receives

$$Y_{\text{caplet}}(t_i) = N(e^{\int_{t_{i-1}}^{t_i} r(t)dt} - e^{\int_{t_{i-1}}^{t_i} Kdt})^+,$$

where K is the fixed interest strike and N is the notional. The holder of the floor receives

$$Y_{\text{floorlet}}(t_i) = N(e^{\int_{t_{i-1}}^{t_i} Kdt} - e^{\int_{t_{i-1}}^{t_i} r(t)dt})^+.$$

The present value of each of these cash flows from the cap are

$$\begin{aligned}
\Pi_{\text{cap}}(t) &= NE^Q\left[e^{-\int_t^{t_i} r(t)dt}(e^{\int_{t_{i-1}}^{t_i} r(t)dt} - e^{\int_{t_{i-1}}^{t_i} Kdt})^+|\mathcal{F}_t\right], \\
&= Ne^{\int_{t_{i-1}}^{t_i} Kdt}E^Q\left[e^{-\int_t^{t_{i-1}} r(t)dt}(e^{-\int_{t_{i-1}}^{t_i} Kdt} - B(t_{i-1},t_i))^+|\mathcal{F}_t\right], \\
&= N_i^{'}P(t, B(t_{i-1},t_i), K_i^{'}, t_i),
\end{aligned}$$

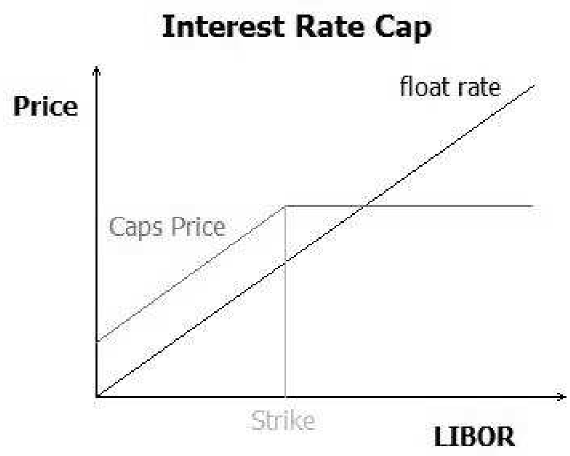


Figure 2: The hedging of interest rate caps with LIBOR.

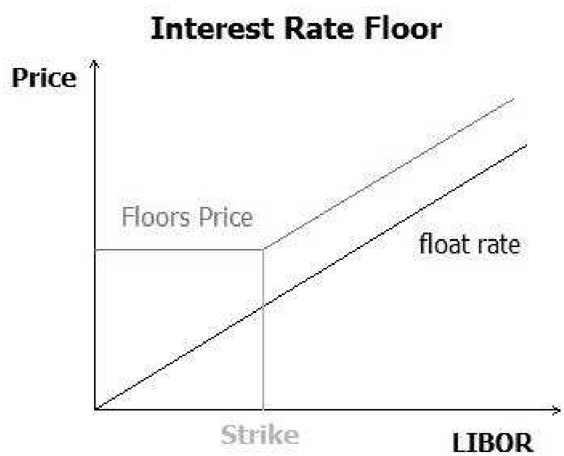


Figure 3: The hedging of interest rate floors with LIBOR.

where $P(t, B(t_{i-1}, t_i), K_i^{'}, t_i)$ denotes the value of a European put option on the bond $B(t_{i-1}, t_i)$. Hence, the discounted values of these cash flows are the same as the value of $N_i^{'}$ European put options with strike $K_i^{'}$ on the zero coupon bond $B(t_{i-1}, t_i)$. The value of a cap is thus

$$\Pi_{\text{cap}}(t) = \sum_{i=1}^{n} N_i^{'} P(t, B(t_{i-1}, t_i), K_i^{'}, t_i).$$

In the same way, the value of a floor is

$$\Pi_{\text{floor}}(t) = \sum_{i=1}^{n} N_i^{'} C(t, B(t_{i-1}, t_i), K_i^{'}, t_i),$$

where $C(t, B(t_{i-1}, t_i), K_i^{'}, t_i)$ denotes the value of a European call option.

2.3 Interest Rate Swaps

An interest rate swap is a financial contract in which two parties agree to exchange interest rate cash flows on a pre-specified value called the swaps notional or face value. The usual setup is that a fixed interest rate is exchanged for a floating rate, the party that receives the fixed rate from the swap is called the receiver and the party that receives the floating rate is called the payer as shown in Figure 4.

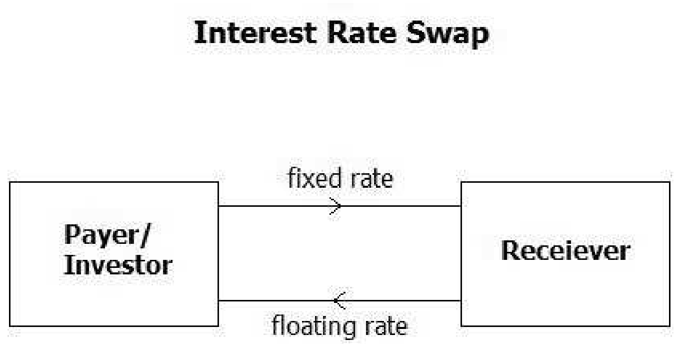


Figure 4: The swap between the method of payment between parties.

The parties exchange cash flows on agreed times during the life of the swap, this could be every three months, every six months or in unusual cases another interval length. This gives the following formula for the receiver cash flow for each payment,

$$Y(t_i) = N\Big(e^{\int_{t_{i-1}}^{t_i} [K-r(t)]dt} - 1\Big),$$

where N is the notional, K is the fixed rate and r(t) is the floating rate. Hence, the discounted payoff during the life of the swap, and thus its value is for the receiver,

$$\Pi_{\text{receive}}(t) = N \sum_{i=1}^{n} e^{-\int_t^{t_i} r(t)dt} \Big(e^{\int_{t_{i-1}}^{t_i} [K-r(t)]dt} - 1\Big).$$

The payer receives the opposite cash flow, which in discounted terms equals

$$\Pi_{\text{pay}}(t) = N \sum_{i=1}^{n} e^{-\int_t^{t_i} r(t)dt} \Big(e^{\int_{t_{i-1}}^{t_i} [r(t)-K]dt} - 1 \Big).$$

One can also take the options of the interest rate swap which is called swaption. Theoretically, buying swaption along with engaging in interest rate swap could help decrease the risk of making a loss. This is of course at the cost of buying a swaption.

2.4 Swaptions

Swaption is the industry term for a swap option, it is a financial derivative that gives the holder the right but not the obligation to enter an interest rate swap. Thus the price of a call option on a receiver swap with strike X at time t is

$$\begin{aligned} \Pi_{\text{call}} &= E^Q \Big[e^{-\int_t^T r(t)} (V_{\text{receive}}(\tau) - X)^+ | \mathcal{F}_t \Big], \\ &= E^Q \Big[e^{-\int_t^\tau r(t)} ((N \sum_{i=1}^{n} e^{-\int_\tau^{t_i} r(t)dt} (e^{\int_{t_{i-1}}^{t_i} [K-r(t)]dt} - 1)) - X)^+ | \mathcal{F}_t \Big]. \end{aligned}$$

Another over-the-counter traded derivative is the bond option.

2.5 Bond Options

An option is a financial contract that gives the holder the right but not the obligation to buy (or sell) an underlying asset in the future at a fixed strike price. A European option can only be exercised at maturity while an American option can be exercised at any time before or at maturity. A call is an option to buy the underlying asset while a put is an option to sell the underlying asset. The value of a put and a call at maturity are thus $P(X) = (0, K - X)^+$ and $C(X) = (0, X - K)^+$ respectively, where X is the asset and K is the strike price. The price of an option is according to the no arbitrage pricing formula is

$$\Pi(t, T) = E^Q \Big[e^{-\int_t^T r(u)du} h(X) | \mathcal{F}_t \Big],$$

where h(X) is the payoff at maturity, i.e. $(0, B(T_1, T_2) - K)^+$ for a European call with maturity T_1 on a bond with maturity T_2.

Investors usually buy a call bond option when he/she believes that the interest rate will fall as the bond price increases with the decreasing interest rate. In contrast, when it is believed that the interest rate will rise, people tends to buy a put bond option as the bond price decreases with the increasing interest rate.

All of mentioned financial instruments are risky and need to be carefully used in order to benefit from them. Hence, many economists and mathematicians came up with different methods in order to capture all the data process and express them in the models. People can, thus, forecast the future and understand the past events to be more cautious with the investment. In the next section, we will elaborate on how financial theory can be applied into the mathematical models.

2.6 Interest Rate Models

Interest rates are often modelled in continuous times by a stochastic differential equation (SDE), most often with an ordinary diffusion as the source of randomness. The diffusion process is scaled by a volatility factor that can be either constant, deterministically time dependent or stochastic. It is generally believed in finance that the volatility factor is stochastic, and in turn driven by another SDE.

One of the most important concepts when talking about interest rate models is the time-value of money which can be represented by using yield curve. The yield to maturity of a zero coupon bond is defined as the factor Y(t,T) that makes the bond price fit the equation

$$B(t,T) = B(T,T)e^{Y(t,T)(T-t)},$$

and is thus the average yield of the bond until its maturity. The yield to maturity is most often quoted as a discretely compounding yearly interest rate. The yield curve or term structure is a relationship between the time to maturity and the yield to maturity. When $Y(t,T)$ is graphed as a function of the time to maturity $(T-t)$ one obtains the yield curve. When fitting an interest rate model it is important that the model fits the initial term structure. As can be seen from the equation above, the term structure at time t can easily be inferred from the market prices of bonds by using the relationship

$$Y(t,T) = \frac{log(\frac{B(t,T)}{B(T,T)})}{(T-t)}.$$

There are many models that try to capture the dynamics of interest rates. For example, Vasicek, Dothan, CIR, Hull-White and many more but in this section, we will only discuss Vasicek [14] and CIR [6] in details.

$$\text{Vasicek: } dr(t) = (b - ar(t))dt + \sigma dW(t)$$
$$\text{CIR: } dr(t) = (b - ar(t))dt + \sigma\sqrt{r(t)}dW(t)$$

where $W(t)$ is a standard Wiener process and σ is a standard constant volatility. In these models the price of a bond is a function of time and the current short rate

$$B(t,T) = E\Big[e^{-\int_t^T r(t)dt}|\mathcal{F}_t\Big] = f(t, r(t)).$$

Because $D(t)B(t,T) = D(t)f(t,r(t))$ is a martingale, we can use Its' formula to obtain a PDE describing the bond price dynamics

$$\begin{aligned} d(D(t)f(t,r)) &= dD(t)f(t,r) + D(t)df(t,r), \\ &= D(t)[-rf + f_t^{'} + \beta f_r^{'} + \frac{\gamma^2}{2} f_{rr}^{''}]dt + D(t)\gamma f_r^{'} dW(t). \end{aligned}$$

As this process is a martingale, the dt term has to equal zero which by using the Feynman-Kac theorem gives us the PDE

$$f_t^{'}(t,r) + \beta f_r^{'}(t,r) + \frac{\gamma^2}{2} f_{rr}^{''}(t,r) = rf(t,r),$$

with the terminal condition

$$f(T,r) = 1 \text{ for all r.}$$

This PDE has the solution

$$f(t,r) = e^{-rC(t,T)-A(t,T)},$$

where

$$\begin{aligned} C(t,T) &= \frac{1}{a}\Big[1 - e^{-a(T-t)}\Big], \\ A(t,T) &= (\frac{b}{a} - \frac{\sigma^2}{2a^2})[C(t,T) - T + t] - \frac{\sigma^2}{4a}C^2(t,T). \end{aligned}$$

Thus there exists a closed-form for solution for the price of a bond in the Vasicek model. In the CIR model, we have the parameters $\beta = (b - ar(t))$ and $\gamma = \sigma\sqrt{r(t)}$ which gives us the PDE

$$f_t'(t,r) + (b - ar)f_r'(t,r) + \frac{\sigma^2 r}{2} f_r'' r(t,r) = rf(t,r).$$

This PDE has a solution of the same form as the Vasicek model, with

$$\begin{aligned} C(t,T) &= \frac{\sinh(\gamma(T-t))}{\gamma\cosh(\gamma(T-t)) + \frac{a}{2}\sinh(\gamma(T-t))} \\ A(t,T) &= -\frac{2b}{\sigma^2}\log\Big[\frac{\gamma e^{\frac{1}{2}a(T-t)}}{\gamma\cosh(\gamma(T-t)) + \frac{a}{2}\sinh(\gamma(T-t))}\Big] \end{aligned}$$

where $\gamma = \frac{1}{2}\sqrt{a^2 + 2\sigma^2}$.

To see how well the closed-form solution fits the actual data, we estimated parameters using the U.S. Treasury yield data in different interest rate models. They are plotted on Figure 5.

From Figure 5, the blue curve with the squares is the plot of the actual data while the red curve is the plot of the actual interpolated data. These two line resemble almost perfectly. The green curve is the estimation using Vasicek model while the black curve is the estimation using CIR process. Vasicek seems to follow more closely to the actual data than CIR

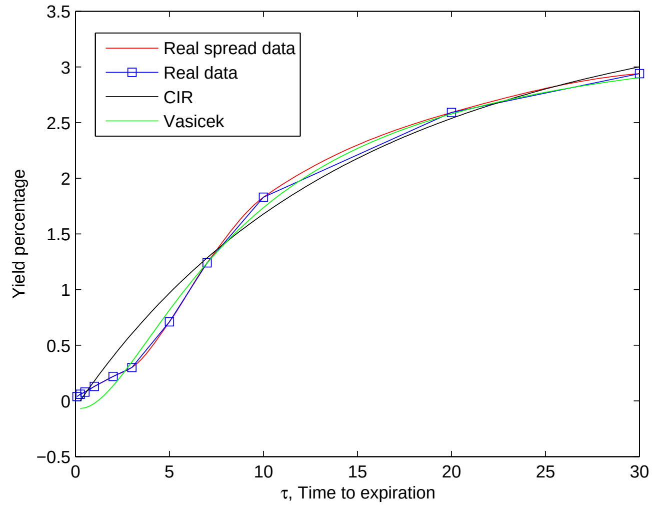


Figure 5: The plot of yield curve from different stochastic volatility model estimation

in this yield plot, especially at the first 10 years. However, the difference made in each model are not so critical since they are still going into the similar direction.

These models both capture the autoregressive tendencies of interest rates, which can be explained economically. For example, if the interest rate is low, people tend to increase their investments and borrowings. This leads to higher demand on capital investments and after some time, it raises the interest rate. On the other hand, if the interest rate is high, people tend to decrease their investments and borrowings which causes the production rate to fall and thus, the interest rate decreases. The Vasicek model however suffers from the drawback that it allows interest rates to become negative. The CIR model allows the interest rates to become zero though never negative but compared to the Vasicek model it suffers from less analytical tractability. There are still drawback with the models, like the fact that they do not fit the initial term structure of the interest rates.

Hull and White made improvements to both these models by letting the parameters a, b and σ be time dependent instead of constant. This leads to the Hull-White extended Vasicek model and the Hull-White extended CIR model:

$$\text{Extended Vasicek: } dr(t) = (b(t) - a(t)r(t))dt + \sigma(t)dW(t)$$
$$\text{Extended CIR: } dr(t) = (b(t) - a(t)r(t))dt + \sigma(t)\sqrt{r(t)}dW(t)$$

When the parameter $b(t)$ is time dependent it can be chosen such that the model fits the initial yield curve. In this thesis, we choose to keep the parameters a and b constant and instead aim to evaluate how interest rate models are affected by modelling $\sigma(t)$ as a stochastic process.

3 Stochastic Volatility

One of the most famous mathematical tools for option pricing is the Black-Scholes model. The basic idea behind this model is to set up a hedging portfolio in which the underlying asset is bought and sold to reflect the option dynamics in order to eliminate risk. It was the first to introduce the closed-form solution to the option pricing. Black-Scholes model can be applied not only to option pricing but also to other derivatives instrument. The solution for Black-Scholes partial differential equation is as follows:

$$c_t(t, S(t)) + rS(t)c_x(t, S(t)) + \frac{1}{2}\sigma^2 S^2(t)c_{xx}(t, S(t)) - rc(t, S(t)) = 0,$$

where $c(t, x)$ is value of call at time t and the stock price at that time is $S(t) = x$ while r and σ are the model parameters which is interest rate and stock volatility, respectively. By using terminal and boundary conditions, the Black-Scholes formula for Call options is formulated as:

$$C(t, S(t)) = S(t)\Phi(d_1) - Ke^{-r\tau}\Phi(d_2),$$

where

$$d_1 = \frac{ln(\frac{S(t)}{K}) + \tau(r + \frac{\sigma^2}{2})}{\sigma\sqrt{\tau}},$$
$$d_2 = d_1 - \sigma\sqrt{\tau},$$

and K is the strike price, τ is the time to maturity and Φ denotes the cumulative standard normal distribution.

However, the model is limited with some explicit assumptions such as that the model's stock price follows a geometric Brownian motion with constant drift and volatility. This is impossible to apply to stock option pricing in reality specifically after the 1987 crash. The volatility of stock prices tend to move more and more stochastically after this incident. This has lead to the development of several stochastic volatility models and another unique method of introducing white noise, ARCH/GARCH. The fact that volatility is not constant but behaves stochastically gives rise to certain unexpected effects. One of those is the volatility smile, and is related to the Black-Scholes formula.

When one uses the Black-Scholes formula to infer future volatility from the price of a call (the implied volatility), he or she finds that the volatility is a function of the strike price. A graph of future volatility as a function of the strike price will often be convex with a minimum around the ATM (At The Money) strike price and higher around the In The Money and Out of The Money as in Figure 6. This is known as the volatility smile. Having a model that fits the implied volatility structure is important both for pricing reasons and for hedging. Stochastic volatility models have the ability to fit the volatility smile and are extensively used in the equity area of finance. To see a clearer picture of option pricing,

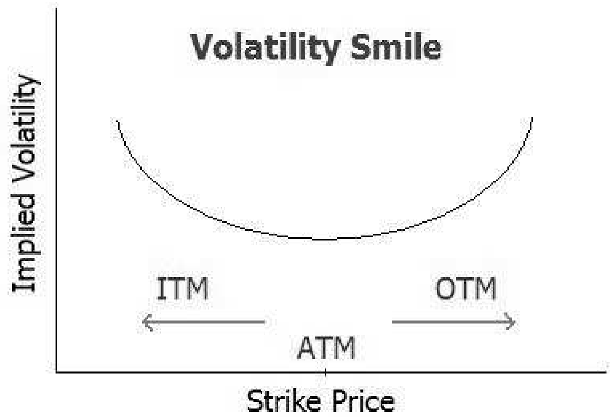


Figure 6: Volatility Smile

the implied volatility is considered an extrinsic value of the option's market price while the underlying stock contribute to the intrinsic value of the option's market price as shown in Figure 7.

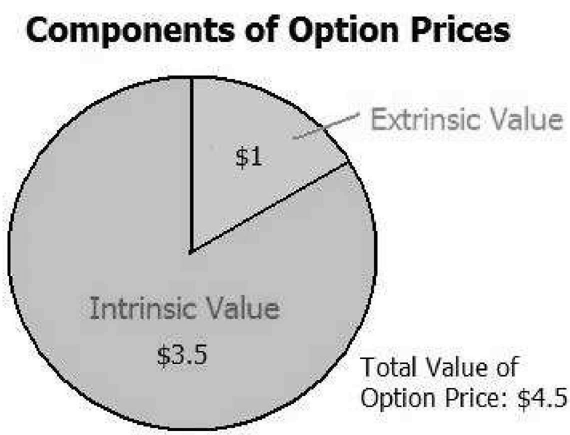


Figure 7: The ITM Option Prices Components

In the fixed income part of finance, implied volatility is the value of the volatility parameter that makes the Black caplet formula fit the market price. The Black caplet formula is based on assumptions about the dynamics of forward rates and is displayed below

$$\Pi_{\text{caplet}}(0) = B(0, T+\delta)[f(0,T)\Phi(d_+) - K\Phi(d_-)],$$

where

$$d_\pm = \frac{1}{\sqrt{\int_0^T \sigma^2(t,T)dt}}\Big[log\frac{f(0,T)}{K} \pm \frac{1}{2}\int_0^T \sigma^2(t,T)dt\Big].$$

Here, $f(0,T)$ is the current forward rate, $\sigma(t,T)$ is the forward volatility at time t, $B(0,T+\delta)$ is the price of a bond with maturity at time $T+\delta$ and K is the strike price [3].

3.1 Models for Stochastic Volatility

As mentioned earlier, volatility changes randomly according to some stochastic process which explains why options with different strikes and maturities have different implied volatilities. The Black-Scholes model uses a standard geometric Brownian motion which gives the SDE:

$$dS(t) = \mu S(t)dt + \sigma S(t)dW(t),$$

where μ is a constant drift, σ is a standard constant volatility and $W(t)$ is a standard Wiener process. Stochastic volatility models replace the constant volatility σ with the process $\sqrt{V(t)}$ yielding the SDE's:

$$\begin{aligned} dS(t) &= \mu S(t)dt + \sqrt{V(t)}S(t)dW_1(t), \\ dV(t) &= \alpha(s,t)dt + \beta(s,t)dW_2(t), \end{aligned}$$

where $\alpha(s,t)$ and $\beta(s,t)$ are function parameters which varies depending on each model. Furthermore, $dW_2(t)$ is the differential of a Brownian Motion that is correlated with $dW_1(t)$ through $\rho dt =< dW_1(t), dW_2(t) >$. In 1987, Hull & White [10] first introduce a Stochastic volatility model using the geometric Brownian motion for the variance process. However, this method causes the variance to grow exponentially instead of decaying over time. Stein & Stein [13], in 1991, incorporate the Ornstein-Uhlenbeck process to solve the mean-reverting behaviour. However, this Ornstein-Uhlenbeck has a main disadvantage that the variance can become negative which is impossible in reality. In 1993, Heston [8] comes up with a model assuming there is a correlation between the underlying assets process and variance process. The variance process is assumed to follow the square root process or Cox-Ingersoll-Ross (CIR) process. However, this model still unable to predict the extreme events that could happen.

In interest rate products, there are several ways to account for the volatility smile. One is to model the short rate and let the volatility of the interest rate process be a stochastic process. Another way is to model the interest rate as the sum of several stochastic processes, so called multi-factor models. If one chooses to model the implied volatility using the first alternative, one usually ends up with the following general equations

$$
\begin{aligned}
dr(t) &= \beta(v, r, t)dt + \gamma(v, r, t)dW_1(t), \\
dv(t) &= \alpha(v, r, t)dt + \phi(v, r, t)dW_2(t),
\end{aligned}
$$

The price of a bond or of an interest rate derivative is a function of the interest rate and follows the equation below

$$
d(D(t)f(v, r, t)) = D(t)df(v, r, t) + dD(t)f(v, r, t),
$$

where

$$
dD(t) = -r(t)D(t)dt,
$$

and

$$
\begin{aligned}
df(v, r, t) &= f'_v dv + f'_r dr + f'_t dt + \frac{1}{2} f''_{vv}(dv)^2 + \frac{1}{2} f''_{rr}(dr)^2 + f''_{vr} dv dr, \\
&= \left[\alpha f'_v + \beta f'_r + f'_t + \gamma\phi f''_{vr} + \frac{\phi^2}{2} f''_{vv} + \frac{\gamma^2}{2} f''_{rr}\right] dt + \phi f'_v dW_2(t) + \gamma f'_r dW_1(t).
\end{aligned}
$$

Using the Feynman-Kac theorem, this leads us to the general PDE for interest rate assets in stochastic volatility models

$$
\alpha f'_v + \beta f'_r + f'_t + \rho\gamma\phi f''_{vr} + \frac{\phi^2}{2} f''_{vv} + \frac{\gamma^2}{2} f''_{rr} = rf.
$$

3.2 The Heston Model

In 1993, Steven Heston has proposed a Stochastic volatility model that has a closed-form solution to the Black-Scholes model of pricing options [8]. It is assumed that the spot asset at time t follows:

$$
\begin{aligned}
dS(t) &= \mu S(t)dt + \sqrt{V(t)} S(t) dW_1(t), \\
dV(t) &= \kappa[\theta - V(t)]dt + \sigma\sqrt{V(t)} dW_2(t),
\end{aligned}
$$

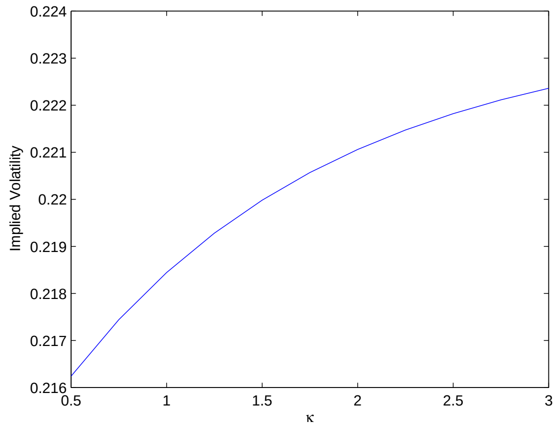
where $W_1(t)$ and $W_2(t)$ are Wiener processes and they are correlated by ρ. $S(t)$ and $V(t)$ are stock price and square root mean reverting volatility processes, respectively. θ is a long-run mean, κ is rate of reversion and σ is the volatility of volatility.

Using Ito's lemma, the Heston model must satisfy the partial differential equation, of any value U,

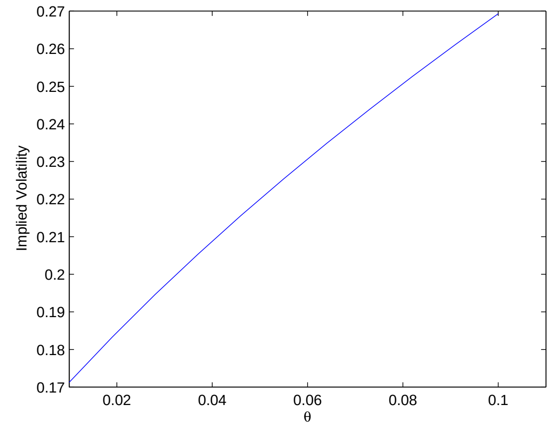
$$
\begin{aligned}
&\frac{1}{2} V S^2 \frac{\partial^2 U}{\partial S^2} + \rho\sigma V S \frac{\partial^2 U}{\partial S \partial V} + \frac{1}{2}\sigma^2 V \frac{\partial^2 U}{\partial V^2} + rS \frac{\partial U}{\partial S} \\
&\quad + [\kappa(\theta - V) - \lambda(S, V, t)] \frac{\partial U}{\partial V} - rU + \frac{\partial U}{\partial t} = 0,
\end{aligned}
$$

where $\lambda(S, V, t)$ represents the price of volatility risk and r is assumed to be a constant interest rate. The market price of volatility risk is defined as

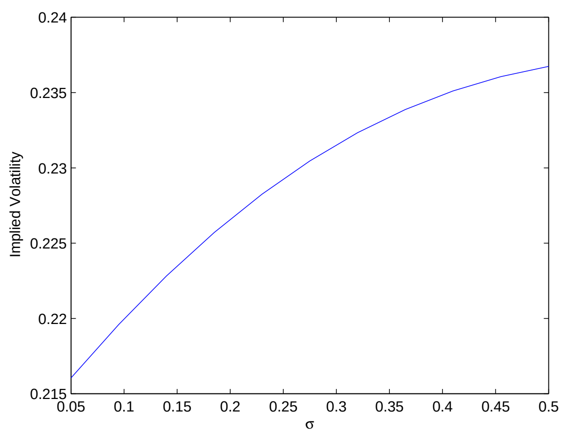
$$\lambda(S, V, t) = k\sigma V(t), \qquad \text{for some constant } k.$$



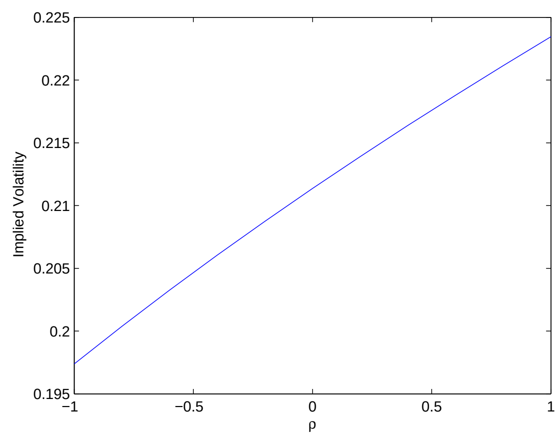
(a) κ values where $K = 1.2, \tau = 1, \theta = 0.05, \sigma = 0.1, \rho = 0.7, V_0 = 0.04, r = 0.03$ and $S_0 = 1$



(b) θ values where $K = 1.2, \tau = 1, \kappa = 1.5, \sigma = 0.1, \rho = 0.7, V_0 = 0.04, r = 0.03$ and $S_0 = 1$



(c) σ values where $K = 1.2, \tau = 1, \kappa = 1.5, \theta = 0.05, \rho = 0.7, V_0 = 0.04, r = 0.03$ and $S_0 = 1$



(d) ρ values where $K = 1.2, \tau = 1, \kappa = 1.5, \theta = 0.05, \sigma = 0.1, V_0 = 0.04, r = 0.03$ and $S_0 = 1$

Figure 8: The implied volatility simulated using Heston model

Under the risk-neutral measure, λ should be eliminated, which gives the exact PDE equation that can be derived from Heston formula. To solve the PDE equation for a European call option with strike price K and maturing time T. Analogously, with the Black-Scholes formula, it is guess that the solution is of the form

$$C(S(t), V(t), t) = S(t)P_1 - Ke^{-r\tau}P_2$$

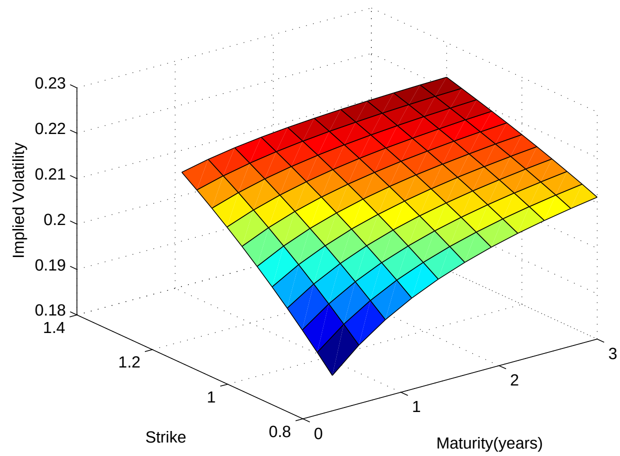


Figure 9: The implied volatility surface with different values of strike, K and maturity time, τ where $\kappa = 1.5, \theta = 0.05, \sigma = 0.1, \rho = 0.7, V_0 = 0.04, r = 0.03$ and $S_0 = 1$

which can be solved analytically using PDE and its proof is in the appendix. Hence, we obtained the solution:

$$f_j(x, V, t; \phi) = \exp(C(\tau; \phi) + D(\tau; \phi)V + i\phi x),$$

where

$$C(\tau; \phi) = r\phi i\tau + \frac{a}{\sigma^2}\left[(b_j - \rho\sigma\phi i + d)\tau - 2\ln\left(\frac{1 - ge^{dr}}{1 - g}\right)\right],$$
$$D(\tau; \phi) = \frac{b_j - \rho\sigma\phi i + d}{\sigma^2}\left(\frac{1 - e^{dr}}{1 - ge^{dr}}\right),$$

and

$$g = \frac{b_j - \rho\sigma\phi i + d}{b_j - \rho\sigma\phi i - d},$$
$$d = \sqrt{(\rho\sigma\phi i - b_j)^2 - \sigma^2(2u_j\phi i - \phi^2)},$$

by solving explicitly. To obtain the desired solution, we invert the characteristic functions,

$$P_j(x, V, T, K) = \frac{1}{2} + \frac{1}{\pi}\int_0^\infty \mathrm{Re}\left(\frac{e^{-i\phi\ln(K)}f_j(x, V, T; \phi)}{i\phi}\right) d\phi.$$

By using the closed-form solution, we can directly plot the implied volatility using the available option prices in the market. The example of implied volatility surface in different value of parameters.

3.2.1 The Heston Model with Fourier Transform

In estimating the price of the stock options, there exists several ways, for example, simulations, finite difference method and many more. However, according to empirical studies, one of the fastest computational method is to use Fourier transform. Another advantage is using the Fourier transform is that the characteristic functions and the FFT do not require the knowledge of the distribution of the underlying [8]. The characteristic function is defined as

$$F(\phi) = E[e^{i\phi x}] = \int_{-\infty}^{\infty} e^{i\phi x} f(x) dx.$$

According to [11], this is the Fourier transform of the function $f(x)$, $\mathcal{F}[f(x)]$ and the inverse of the Fourier transform is

$$\mathcal{F}[F(\phi)] = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i\phi x} F(\phi) d\phi = f(x).$$

Let x_t be the log price, $\ln S_t$, and denote k as the log of strike price K. Hence, the Call options value is

$$C_T(k) = e^{-rT} \int_{k}^{\infty} (e^{x_T} - e^{k}) f_T(x_T) dx_T,$$

and we let

$$c_T(k) = e^{\alpha k} C_T(k),$$

where α is a damping factor. By Carr and Madan [4], consider the Fourier transform of $c_T(k)$,

$$F_{c_T}(\phi) = \int_{-\infty}^{\infty} e^{i\phi k} c_T(k) dk,$$

and

$$c_T(k) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i\phi k} F_{c_T}(\phi) d\phi.$$

Thus, by substituting the inverse Fourier term into the Call price with damping factor, the Call options price is represented as

$$C_T(k) = e^{-\alpha k} \frac{1}{\pi} \int_{0}^{\infty} e^{-i\phi k} F_{C_T}(\phi) d\phi,$$

According to Hong [9], the forward characteristic function, F_{C_T} is

$$F_{C_T}(\phi) = e^{A(\phi)+B(\phi)+C(\phi)},$$

where

$$\begin{aligned}
A(\phi) &= i\phi(x_0 + rT), \\
B(\phi) &= \frac{2\zeta(\phi)(1-e^{-\psi(\phi)T})V_0}{2\psi(\phi) - (\psi(\phi)-\gamma(\phi))(1-e^{-\psi(\phi)T})}, \\
C(\phi) &= -\frac{\kappa\theta}{\sigma^2}\left[2\ln\left(\frac{2\psi(\phi) - (\psi(\phi)-\gamma(\phi))(1-e^{-\psi(\phi)T})}{2\psi(\phi)}\right) + (\psi(\phi)-\gamma(\phi))T\right], \\
\zeta(\phi) &= -\frac{1}{2}(\phi^2 + i\phi), \\
\psi(\phi) &= \sqrt{\gamma(\phi)^2 - 2\sigma^2\zeta(\phi)}, \\
\gamma(\phi) &= \kappa - \rho\sigma\phi i.
\end{aligned}$$

It is much more convenient to convert this kind of integral into summation and use FFT which is an efficient algorithm to calculate such summation:

$$w(k) = \sum_{j=1}^{N} e^{-i\frac{2\pi}{N}(j-1)(k-1)} x(j).$$

The Call option price can be approximated into similar form in order to solve using FFT method, by using Simpson's rule,

$$C_T(k_u) \approx \frac{e^{-\alpha k_u}}{\pi} \sum_{j=1}^{N} e^{-i\frac{2\pi}{N}(j-1)(u-1)} e^{ibv_j} F_{c_T}(v_j)\frac{\eta}{3}(3 + (-1)^j - \delta_{j-1}).$$

3.3 Complex Stochastic Volatility Models

In this section we describe stochastic volatility models, that we have chosen to denote "complex stochastic volatility models" because of the fact that none of them have analytical solutions to bond prices and prices of derivatives. They are models with two coupled SDE's, with one SDE describing the interest rate dynamics, while the other one describes the stochastic volatility dynamics. For simplicity, we chose only to look at the CIR SDE as a model for the interest rate dynamics, while varying the models for stochastic volatility. The first model we look at is the CIR model, which is defined by the SDE

$$dv(t) = (d - cv(t))dt + \eta\sqrt{v(t)}dW(t).$$

If we unlike the Heston model assume that the variance process has a quadratic variation proportional to the variance and not the volatility, we obtain the GARCH diffusion model with the SDE

$$dv(t) = (d - cv(t))dt + \eta v(t)dW(t).$$

In the 3/2 model the quadratic variation of the variance process is assumed to be proportional to the spot variance raised to 3/2, giving the SDE for the interest rate and variance process

$$dv(t) = (d - cv(t))dt + \xi v^{\frac{3}{2}}(t)dW(t).$$

This SDE experiences a phenomenon called volatility induced stationarity (VIS) where the volatility itself induces the stationarity in the dynamics. Basically, as the variance grows bigger so does the quadratic variation of this process. This increases the probability of reaching lower levels again, where the quadratic variation decreases due to the level effect. Thus at lower levels, the probability of reaching higher levels decreases, the process is kept at low levels and thus stationary.

The last model is the CKLS model, which is very similar to the previous models, with the exceptions that the level parameter, γ (1/2 for the CIR model, 1 for the GARCH model and 3/2 for the 3/2 model) is not fixed. It is instead allowed to vary, with the possibility of fitting the dynamics even better. The SDE for this model is

$$dv(t) = (d - cv(t))dt + \xi v^{\gamma}(t)dW(t).$$

As mentioned before, there exists no analytical solutions to the prices of financial contract using any of these models. Obtaining a price will thus require numerical solution of the price, either by using Monte Carlo simulations or by solving the PDE that is obtained by martingale pricing. These methods require a lot of computational power and are thus potentially very time consuming. Hopefully this will be outweighed by a better description of interest rate dynamics and more accurate pricing.

4 Fitting an Interest Rate Model

When fitting an interest rate model one needs to take into account the market information that is available. Depending on what is being modelled, one needs to take into account different kinds of market data. For example if it is only the interest rate that is being modelled as a stochastic process, then the interest rate can be modelled using historical data. If on the other hand it is financial assets that are of interest, on needs to take into account the current market price of risk as well as expectations about future levels of the interest rates and volatility, which can be inferred from the yield curve and traded interest rate products.

4.1 Parameter Estimation

To begin with, the parameters of the models are estimated under the empirical measure. The market price of interest rate and volatility risk is then calculated so that the model fits the yield curve and thus follows the risk neutral dynamics.

4.1.1 Discretization Approximation

When estimating the parameters under the empirical measure one needs only to look at the short rate. We choose to do this in the same way that is done by Nowman [12]. In the limit $\lim_{\Delta t \to 0} r(t)$ is Gaussian allowing us to use maximum likelihood on the normal probability function. From the SDE governing the interest rate process,

$$dr(t) = (b - ar(t))dt + \sqrt{v(t)}\sqrt{r(t)}dW(t),$$

we can obtain an approximation of the interest rate as a function of t,

$$r(t) = e^{-at}(r(0) + \frac{b}{a}(e^{at} - 1) + \int_0^t e^{at}\sqrt{v(t)}\sqrt{r(t)}dW(t)).$$

Because we are in the limit $\Delta t = \delta$ we can assume that v(t) and r(t) are approximately constant on the interval $t \in [t, t + \Delta t]$. We set $\sqrt{v(t)} = \sigma(t_i)$ and $r(t) = r(t_i)$ obtain

$$r(t) = e^{-a\Delta t}(r(t_i) + \frac{b}{a}(e^{a\Delta t} - 1) + \sigma(t_i)\sqrt{r(t_i)}\int_0^t e^{at}dW(t)).$$

We can now calculate the conditional expectation and variance of r(t).

$$E[r(t)] = e^{-a\Delta t}r(t_i) + \frac{b}{a}(1 - e^{-a\Delta t}),$$
$$m_{tt}^2 = Var[r(t)] = \frac{\sigma(t_i)^2 r(t_i)}{2a}(1 - e^{-2a\Delta t}).$$

Knowing that the increments are approximately Gaussian, we get the likelihood function

$$L(\theta) = \sum_{i=1}^{N} -2\log(m_{tt}) - \frac{(r(t_i) - e^{-a\Delta t} r(t_{i-1}) - \frac{b}{a}(1 - e^{-a\Delta t}))^2}{m_{tt}^2}.$$

We can thus calculate close approximations of the parameters in the interest rate model by maximizing this likelihood function. To later be able to model the variance process, we choose to initially model it as a GARCH(2,1)-process to obtain measurements of the underlying variance. In the GARCH(2,1)-model, the variance follows the discrete-time process

$$\sigma(t_i)^2 = p_0 + p_1\sigma(t_{i-1})^2 + p_2\sigma(t_{i-2})^2 + p_3\epsilon(t_{i-1})^2,$$

where $\epsilon(t_i)$ is the i^{th} residual of the interest rate model. This initial step will give us the parameters for the interest rate process and a close approximation of the variance at each discrete time, t_i.

Next we repeat the procedure to model the parameters for the variance process. Here, we obtain the equations

$$\begin{aligned}
v(t) &= e^{-c\Delta t}(v(t_i) + \frac{d}{c}(e^{c\Delta t} - 1) + \eta v(t_i)^\gamma \int_0^t e^{ct} dW(t)) \\
E[v(t)] &= e^{-c\Delta t} v(t_i) + \frac{d}{c}(1 - e^{-c\Delta t}) \\
m_{tt}^2 &= Var[v(t)] = \frac{\eta^2 v(t_i)^{2\gamma}}{2c}(1 - e^{-2c\Delta t}) \\
L(\theta) &= \sum_{i=1}^{N} -2\log(m_{tt}) - \frac{(v(t_i) - e^{-c\Delta t} v(t_{i-1}) - \frac{d}{c}(1 - e^{-c\Delta t}))^2}{m_{tt}^2}
\end{aligned}$$

The parameters for the variance process are thus obtained by maximizing the approximate likelihood function, $L(\theta)$.

MODEL	b	a	d	c	η	γ
CIR	0.001738	-0.000056	0.000034	0.0109501	0.039348	0.5
GARCH	0.001738	-0.000056	0.021058	95.062057	13.465162	1
3/2	0.001738	-0.000056	0.000012	-0.063316	35.049675	3/2
CKLS	0.001738	-0.000056	0.008792	42.983610	23.896824	1.134004

Table 1: Estimated parameters for the variance models by the discretization method.

As can be seen in the table, the interest rate experiences a positive drift from the parameter b which would be expected. What is not expected however is that the parameter a, which should account for a downward drift at high interest rates, is negative. This will cause the interest rate to diverge, which is unrealistic in economic terms. It is our belief that the parameters a and b are not very significant to the model performance and that they could be discarded.

If one takes a look at the parameters estimated for the volatility process one can see that all of them are plausible. Most models have drift parameters that make the variance mean reverting. This is true even for the 3/2 model, because it has elements of volatility induced stationarity (VIS) where the diffusion process generates mean reversion. It is therefore entirely possible that the 3/2 variance process has a positive drift parameter.

4.1.2 Efficient Method of Moments

The method described above is an approximation and thus might yield estimates that deviate from the true values. In an attempt to improve this weakness, the Efficient Method of Moments (EMM) of Galant & Tauchen was employed [7]. Instead of using the true transition probability, which in our case is unknown, the EMM method uses an auxiliary likelihood function. The derivatives of this auxiliary function serve as moment conditions to which the moments of the structural model i.e. our SDE's will be matched. The equation of moments is defined as

$$m(s,\hat{\eta}) = \int \frac{\partial \log\ f(r_t|r_{t-1},\hat{\eta})}{\eta} dP(r_t|r_{t-1},s) \approx \frac{1}{T}\sum_{t=1}^{T} \frac{\partial \log\ f(r_t(s)|r_{t-1}(s),\hat{\eta})}{\eta} = m_N(s,\hat{\eta}),$$

where s are the parameters of the structural model and η are the parameters of the auxiliary model. The true parameters of the structural model are given by the equation

$$\hat{s} = \arg\min_{s} m_N(s,\hat{\eta})' W m_N(s,\hat{\eta}),$$

where W is a weighting matrix. An optimal EMM estimator is obtained if the weighting matrix is a consistent estimator of the asymptotic covariance matrix for the moment conditions. The covariance matrix may be estimated through the 'outer product of the gradient' formula [7],

$$I(\hat{\eta}) = \frac{1}{T}\sum_{t=1}^{T} \frac{\partial \log\ f(r_t(s)|r_{t-1}(s),\hat{\eta})}{\eta} \frac{\partial \log\ f(r_t(s)|r_{t-1}(s),\hat{\eta})}{\eta'}.$$

Our auxiliary model is chosen to be an AR-GARCH-SNP model, as is chosen by Andersen & Lund [1]. In this auxiliary model, the interest rate process is modelled as an AR(5)-process, while the volatility is modelled as a GARCH(2,1)-Hermite(6) process. The Hermite polynomials are added to account for heavy tails and other non-Gaussian behaviour. Thus, our auxiliary likelihood function looks like

$$f(r_t|r_{t-1},\eta) = \frac{[P_K(z_t, r_{t-1})]^2 \varphi(z_t)}{\sigma_t^2 \int_{-\infty}^{\infty} [P_K(x, r_{t-1})]^2 \varphi(x) dx},$$

where

$$\begin{aligned}
z_t &= \frac{r_t - \mu_t}{\sigma_t}, \\
\mu_t &= p_0 + \sum_{i=1}^{5} p_i (r_{t-i} - p_0), \\
P_K(z_t, r_{t-1}) &= 1 + \sum_{i=1}^{6} a_i z_t^i, \\
\sigma_t^2 &= \alpha_0 + \alpha_1 \sigma_{t-1}^2 + \alpha_2 \sigma_{t-2}^2 + \alpha_3 z_{t-1}^2.
\end{aligned}$$

Here, the function $\varphi(z_t)$ denotes the standard normal probability density and $P_K(z_t, r_{t-1})$ denotes the Hermite polynomial.

The Auxiliary model as described above led the optimization to parameter values that were deemed infeasible and caused interest rates to diverge, just as in the previous section. It is probable that the maximum likelihood estimation of the auxiliary model points to this unnatural diverging behaviour. We believe that this estimate is the result of statistical uncertainty and that the parameter estimates of the drift are just noise. We therefore discard the mean parameter in the above model and set $\mu_t = 0$.

When calibrating the auxiliary model, parameters were estimated through maximum likelihood. This was done several times with different random starting points to ensure that the optimization did not exit in a local optimum. The EMM estimation was the started at the parameters obtained from the discretization approximation. The results from the EMM estimation are listed in Table 2 and 3.

MODEL	b	a	d	c	η	γ
CIR	0	0	0.00002469	0.02298355	0.00995850	0.5
GARCH	0	0	0.01158281	103.64424329	17.11198308	1
3/2	0	0	0.00001214	-0.06992139	32.57322946	3/2
CKLS	0	0	0.00581664	41.69924781	25.57455124	1.12722799

Table 2: Parameter estimates obtained by the EMM method.

Parameters	b_1	a_1	σ_1	b_2	a_2	σ_2
	0	0.00003546	0.00739025	0.00192166	0.00001646	0.02075589

Table 3: Parameter estimates for the Two factor CIR model obtained by EMM.

4.2 Market Calibration

In this section, we calibrate the models to a current market data. The CIR model with the Kalman filter is calibrated to present and historical value of the yield curve, whereas the other models are calibrated to the current yield curve.

4.2.1 Two-Factor CIR with Kalman Filter

The two factor CIR model is given by:

$$\begin{aligned} r(t) &= x_1(t) + x_2(t), \\ dx_i(t) &= \kappa_i(\theta_i - x_i(t))dt + \sigma_i\sqrt{x_i(t)}dW_i(t),\ i = 1, 2, \end{aligned}$$

where $W(t)$ is an independent Brownian motions. As derived previously and according to [5], the solution to this differential equation in the two-factor CIR is

$$P(\tau) = A_1(\tau)A_2(\tau)e^{-B_1(\tau)x_1 - B_2(\tau)x_2},$$

where

$$\begin{aligned} A_i(\tau) &= \left[\frac{2\gamma_i e^{(\kappa_i+\lambda_i+\gamma_i)\tau/2}}{(\kappa_i+\lambda_i+\gamma_i)(e^{\gamma_i\tau}-1)+2\gamma_i}\right]^{2\kappa_i\theta_i/\sigma_i^2}, \\ B_i(\tau) &= \frac{2(e^{\gamma_i\tau}-1)}{(\kappa_i+\lambda_i+\gamma_i)(e^{\gamma_i\tau}-1)+2\gamma_i}, \end{aligned}$$

where $\gamma_i = \sqrt{(\kappa_i+\lambda_i)^2 + 2\sigma_i^2}$. The parameters are defined as following: mean reversion of volatility κ, long-term volatility mean θ, volatility of volatility σ, price of volatility risk λ and the risk-free rate r. In order to eliminate risk using delta hedging, the no arbitrage principle should hold. However, for every unit of volatility risk, there is are λ units of extra return. On the other hand, it is argued that λ should be independent and could be determined by any volatility-dependent asset and used it to price other assets. The continuously compounded yield for discount bond for each different period, τ_i, is

$$Y(t,T) = \frac{B_1(\tau_i)x_1}{\tau_i} + \frac{B_2(\tau_i)x_2}{\tau_i} - \frac{\ln A_1(\tau_i)}{\tau_i} - \frac{\ln A_2(\tau_i)}{\tau_i}, \qquad \tau_i = T - t_i.$$

To estimate the model numerically, we use the state space model with Kalman filter adjustment for unobservable variables. The continuous time model is expressed as:

$$\begin{aligned} y_t &= C + Dy_{t-1} + \nu_t, \\ z_t &= A + By_t, \end{aligned}$$

where y_t is the mean value and

$$C = \begin{bmatrix} \theta_1(1 - e^{-\kappa_1 \Delta t}) \\ \theta_2(1 - e^{-\kappa_2 \Delta t}) \end{bmatrix},$$

$$D = \begin{bmatrix} e^{-\kappa_1 \Delta t} & 0 \\ 0 & e^{-\kappa_2 \Delta t} \end{bmatrix},$$

$$B_i = \begin{bmatrix} \frac{B_1(\tau_i)}{\tau_i} & \frac{B_2(\tau_i)}{\tau_i} \end{bmatrix},$$

$$A_i = -\frac{\ln A_1(\tau_i)}{\tau_i} - \frac{\ln A_2(\tau_i)}{\tau_i},$$

and variance of y_t is

$$Q = \begin{bmatrix} q_1 & 0 \\ 0 & q_2 \end{bmatrix},$$

where $q_i = \frac{\sigma_i^2}{\kappa_i}(e^{-\kappa_i \Delta t} - e^{-2\kappa_i \Delta t})y_t + \frac{\theta_i \sigma_i^2}{2\kappa_i}(1 - e^{-\kappa_i \Delta t})^2$. According to [5], the parameters are typically estimated by the method of maximum likelihood using the Kalman filter. The basic theory behind Kalman filter is the cycle of time updates that projects the current state estimate ahead in time or "predicts", and a measurement update that adjusts the projected estimate by an actual measurement at that time or "correct". The predicting step equations are:

$$\begin{aligned} y_t^- &= C + Dy_{t-1}, \\ v_t^- &= Dv_{t-1}D^T + Q, \end{aligned}$$

where the negative power y_t^- indicates the prior knowledge and the correcting step equations are:

$$\begin{aligned} K_t &= v_t^- B^T (B v_t^- B^T + R)^{-1}, \\ y_t &= y_t^- + K_t(x_t - By_t^-), \\ v_t &= v_t^-(I - K_t B), \end{aligned}$$

where R is the measurement error covariance and x_t is the actual measurement. Eventually, the parameters are estimated using by maximizing the following log-likelihood function:

$$L = -\frac{N \cdot m}{2}\ln(2\pi) - \frac{1}{2}(\ln(\mathrm{Det}(Bv_t^- B^T + R)) + (x_t - By_t^-)^T(Bv_t^- B^T + R)^{-1}(x_t - By_t^-)).$$

The two-factor CIR parameters are estimated using Kalman filter from the total of 87 loops, where we calibrated using the yield curve and obtained parameters as shown in Table 4.

Parameters	Estimates	
κ	0.960009182 (0.080480847)	0.049776547 (0.002295931)
λ	-0.102155613 (0.005670062)	-0.050607585 (0.00313773)
σ	0.213243831 (0.033279424)	0.099290394 (0.010696689)
θ	0.097197115 (0.008878529)	0.00099276 (0.00004773)
$\hat{\sigma_1}$	0.006172291 (0.782156219)	
$\hat{\sigma_2}$	0.152029436 (0.922952245)	
$\hat{\sigma_3}$	-0.033713879 (0.843342428)	
$\hat{\sigma_4}$	0.006736857 (0.746635426)	
$\hat{\sigma_5}$	0.132825503 (1.003501827)	
$\hat{\sigma_6}$	-0.121600789 (0.866313306)	

Table 4: Parameter estimations for the Heston model

From Table 4, the values in the parenthesis are standard deviation of each parameters. The number in front for κ, λ, σ and θ are the first factor and the later is the second factor. By using this parameters, we plotted the yield curve of two-factor CIR along with the actual data.

Figure 10 illustrates that the two-factor CIR model follows the actual data closely in the same trend but not perfectly coincide with the actual value. We can briefly conclude that two-factor CIR model using Kalman filter to estimate the parameters produced well results. Seeing as the modelled values do not deviate much from the actual one. We can see the evaluation of this results later in the model evaluation part.

4.2.2 Calibration of Complex Models

Because the CIR model with stochastic volatility has no analytical solution for the yield it is infeasible to find the market price of parameter risk as is done with a Kalman filter in the Heston model. Instead, the model is recalibrated with risk-neutral parameters. In

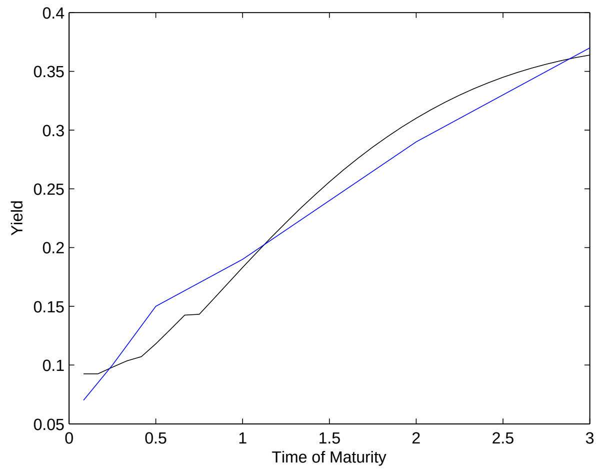


Figure 10: The plot of yield curve of two-factor CIR with the actual data

this way the market price of risk will be reflected in the parameter estimates themselves, instead of being an independent parameter. This is done by minimizing the L^2-norm of the difference between the model yield and the current yield curve. Here, the yield for the model is calculated using Monte Carlo simulation.

The current short rate is obtained from the yield curve, whereas the current volatility is measured as the unexplained variance of the last few measurements of the short rate. N=10^4 trajectories of the interest rate are then simulated over the relevant time span, $\tau = T - t$, and the yield is calculated. The models we have chosen to investigate cannot fit all the complex shapes of the yield curve. If we would not fit the models to the relevant time span they would perform less ideally. This is because they would either not take into account all the available information from the market or they would provide an average fit to a too long time span. For example, if a derivative with expiration within four years is to be priced, the model is calibrated to the first four years of the yield curve. The market expectations about the time after the maturity of the derivative is irrelevant for the pricing of said derivative. As the calibrated parameters are only valid for the actual day of calibration and because there are numerous sets of parameters they are listed in the appendix.

To get a good view of the models' potentials, they are also calibrated to swaption prices. Because there is a possibility that the yield curve does not contain all relevant market information, the models might not come as close to the swaption prices as their flexibility allows. Therefore we also investigate how well the models can fit swaption prices by calibrating the models directly to the swaption prices. This will paint a better picture of the actual flexibility of the models.

4.2.3 Calibration of Heston Model

Heston model is calibrated using the Caplet price. The data is obtained on May 14th, 2012 for the options time to maturity range of 1 month, 3 month, 6 month, 1 year, 2 years and 3 years with the underlying time to maturity of 1 year, 2 years, 3 years, 4 years, 5 years, 7 years, 10 years and 15 years. The Heston model is convenient as it has the closed-form solution hence we can obtain the Caplet price directly from the model, then use it the calibrate the model to be more precise. The parameters obtained are shown in Table 5 .

	κ	θ	σ	ρ	$v0$
1 Year	0.525435261	0.999930024	0.000015872	-0.000000102	0.000000003
2 Years	1.550679426	0.999124088	0.025168180	-0.004255795	0.000398513
3 Years	1.178803788	0.990197632	0.000554063	-0.020030888	0
4 Years	5.839651686	0.597719620	2.512773095	-0.339826433	0.012415896
5 Years	11.261378543	0.999997735	0.001474514	-0.000093182	0.995735899
7 Years	10.674520823	0.851867993	0.085842532	-0.012758259	0.000021683
10 Years	0.999349794	0.414769050	0.245768553	-0.000000394	0.000049043
15 Years	4.843913448	0.682357856	2.041824346	-0.999975226	0.860885106

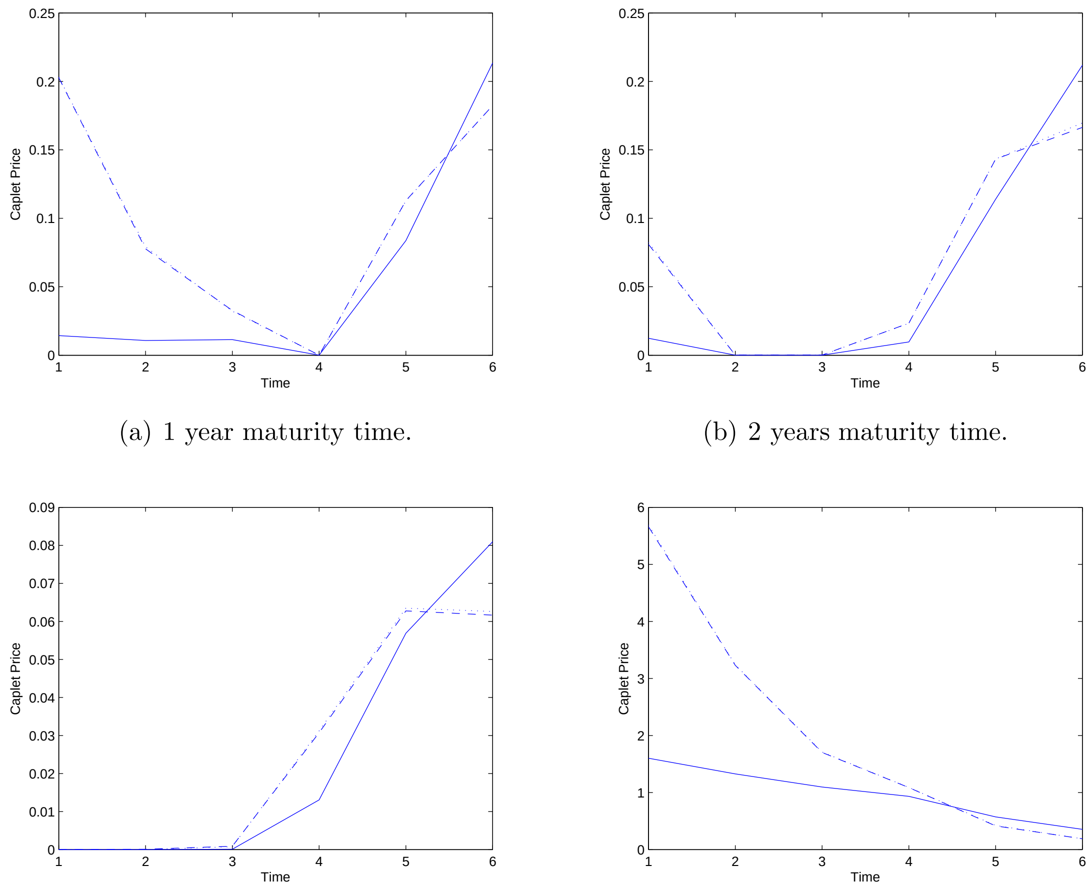
Table 5: Parameter estimates obtained by the Heston method.

By applying these parameters back to the model to get the Caplet price, we have the results as shown in Figure 11 and 12. The dotted line is the calibrated Heston Caplet price, the dashed line is the calibrated FFT Heston Caplet price and the solid line is the actual price of the Caplet.

From Figure 11 and 12, it is clear that the model cannot resemble the actual price perfectly but still, they are going in the same trend, mostly. This can be implied that Heston model is better when calibrated with the call option prices as many other papers have done so. The average time taken for all to calibrate the model and the time difference between using the normal Heston model and Fourier transform method are shown in Table 6.

	Calibration	Heston	FFT Heston
Time(sec)	57.39973038	0.646965875	0.34137575

Table 6: Time taken for the Heston method.



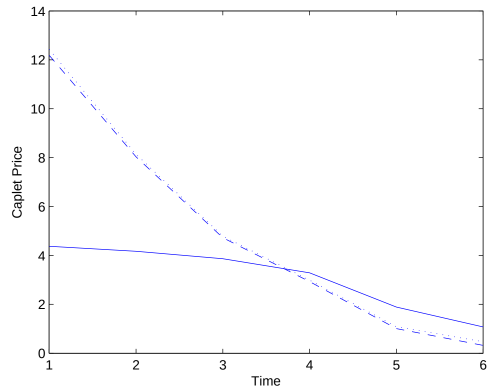
(a) 1 year maturity time.

(b) 2 years maturity time.

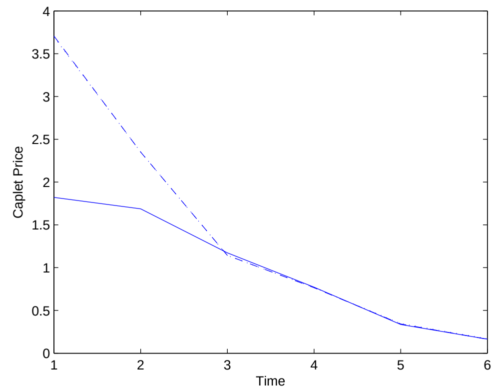
(c) 3 years maturity time.

(d) 4 years maturity time.

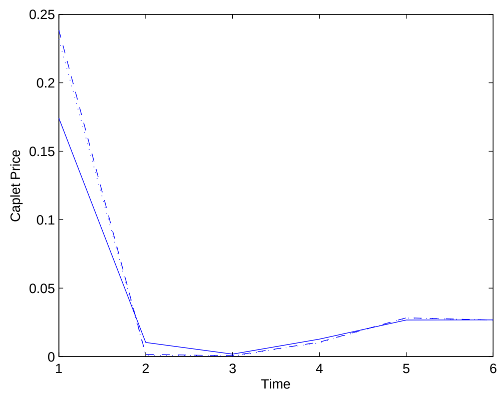
Figure 11: The calibrated Heston Caplet price for 1-4 years maturity time.



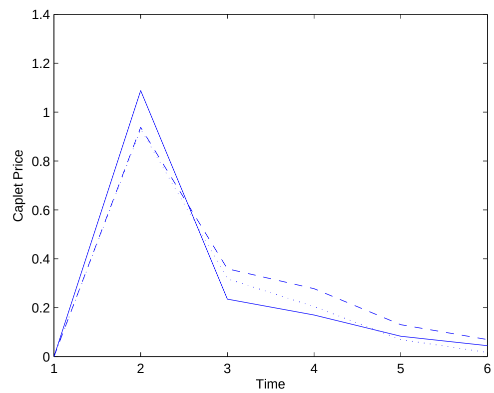
(a) 5 years maturity time.



(b) 7 years maturity time.



(c) 10 years maturity time.



(d) 15 years maturity time.

Figure 12: The calibrated Heston Caplet price for 5, 7, 10 and 15 years maturity time.

5 Evaluating the Model

In the previous sections we have calibrated the models in two ways for two different purposes. In this section we will investigate how the models perform with respect to the two purposes, i.e. describing the interest rate dynamics and pricing interest rate assets.

5.1 Evaluating the Short-Rate Model Dynamics

In evaluating the short term model dynamics, the transition function of the SDE's play a big role. If the transition function of the model closely matches the empirical transition function then the model performs well. In general, the closer the match the better the model. We will investigate the transition function using a simple rank test. We will also look at the long term distribution of the variance and investigate the distance from the empirical variance distribution.

5.1.1 A Rank Test

To test how the different models perform in describing the interest rate dynamics we choose a general rank test that works for all models. For each time interval, [t-1,t], we simulate M trajectories of the interest rate starting in state (r_{t-1}, v_{t-1}). The true value, r_t is then ranked among the endpoints, x_t of the simulated interest rate. Denote R_t the rank of r_t at time t, then

$$P(R_t = q) = p_{t,q} = \frac{1}{M+1},\ q = 1,\ \dots, M+1;\ \ t = 2+l,\ \dots, N$$

where l is the number of lags required for the measurement of the variance. To simplify we assume that the probability is independent of time, meaning that $p_{t,q} = p_q$. This leads to the equality

$$\hat{p}_q = \frac{\Omega_{N-1}(q)}{N-1}\ q = 1,\ \dots, M+1$$

where

$$\Omega_{N-1}(q) = \sum_{t=2}^{N} I\{R_t = q\};\ \ q = 1,\ \dots, M+1$$

Under the hypothesis that the interest rate is generated from our model, we have

$$E\left[\hat{p}_q\right] = E\left[\frac{\Omega_{N-1}(q)}{N-1}\right] = \frac{1}{M+1}.$$

We use the test statistic

$$X^2 = \sum_{q=1}^{M+1} \frac{(\Omega_{N-1}(q) - \frac{N-1}{M+1})^2}{\frac{N-1}{M+1}},$$

which is under H_0 asymptotically distributed as $\chi^2(M)$.

To obtain an appropriate starting point (r_{t-1}, v_{t-1}) we need to measure the variance at that time. It is impossible to observe the variance, so the best we can do is to approximate it. This is done by using a GARCH process as before and assuming that $\hat{\sigma}^2_{t-1} = \sqrt{r(t)}\sqrt{v(t)}$. The fact that we cannot observe the variance directly and have to use a proxy is going to affect our goodness of fit test. In an attempt to reduce this influence we choose to evaluate the estimated SDE's over time intervals of differing size. Our assumption here is that the longer the the time interval, the lower the affect of the initial value. This is because of the autocorrelation of the variance which decays rather quickly. Because the autocorrelation decays quickly the importance of the starting position will also decrease rapidly with time.

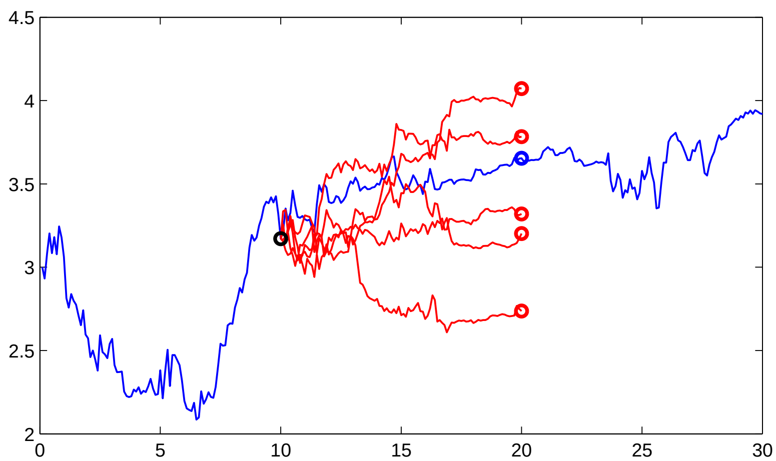


Figure 13: An illustration of the rank test for one data point.

Five time intervals were arbitrarily chosen: 1, 5, 10, 15 and 20 days. According to Bak, Nielsen and Madsen the χ^2 approximation fails when the expected rank frequencies are low, and there is a consensus that this expected frequency should be no lower than 5 [2]. This means that the number of simulated trajectories and thus the degrees of freedom, M, are less than or equal to $\frac{N-1}{5} - 1$, where N is the number of time intervals. The results of the simulations are summarized in the table below.

As can be seen in table 7, both the CIR model and the 3/2 model give substantially lower χ^2-values than the GARCH and CKLS models. The P-values in table 8 that correspond to the χ^2-tests tell us that the CIR model is the only model that is reasonably close to the true marginal distribution. The striking thing is that the CIR model outperforms the CKLS model. This should not be the case because the CKLS model is more flexible and should end up with the same parameter estimates as the CIR model, if these were the most feasible.

TIME INTERVAL	1	5	10	15	20
MODEL/$\chi^2(N)$	$\chi^2(953)$	$\chi^2(188)$	$\chi^2(93)$	$\chi^2(61)$	$\chi^2(45)$
CIR	2800.07	229.65	88.39	66.20	53.08
GARCH	7910.16	959.04	378.29	225.13	143.92
3/2	2960.64	253.87	146.06	101.91	58.10
CKLS	5831.87	887.18	365.18	200.50	135.29

Table 7: χ^2-values for the different models in the rank test.

MODEL / TIME INTERVAL	1	5	10	15	20
CIR	0	0.0206	0.6185	0.3022	0.1908
GARCH	0	0	0	0	0
3/2	0	0.001	0.0004	0.008	0.091
CKLS	0	0	0	0	0

Table 8: Corresponding P-values to the χ^2-tests.

5.1.2 Stationary Distribution Tests

In our interest rate model, the increments divided by the square root of the current interest rate are stationary. This means that as the number of normalized increments increases, they will converge in distribution. A criterion for our SDE model to be correct is that the increments converge to the same distribution as the one observed from the data. We measure this by comparing the Kolmogorov-Smirnov distance.

For this test to work, we need to have so many data points that they are deemed independent from the starting value of the process. To see how long time this might take, we look at the autocorrelation function of the variance process. We choose to discard a burn in period of the first N values that are deemed affected by the starting point of the process.

It can be seen in Figure 14 that the autocorrelation function of the variance goes toward zero at around 150 lags in the CIR model. Values that are separated by more than 150 days will thus be uncorrelated and the state of a simulated Markov Chain that is longer than this will be independent of its starting value. The other models did not express autocorrelation to the same degree as the CIR model, why the CIR model is the limiting factor in choosing a burn in period. To be on the safe side we choose to discard the first 500 values of our simulations.

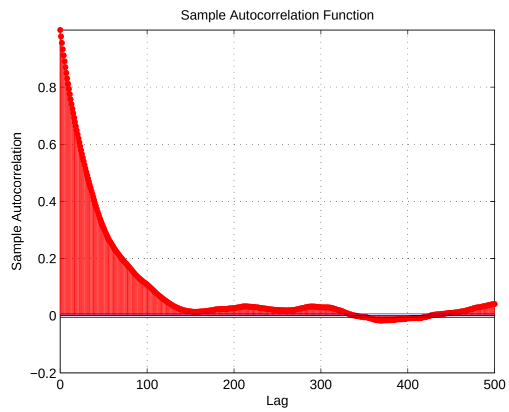


Figure 14: Autocorrelation of the CIR Variance Process

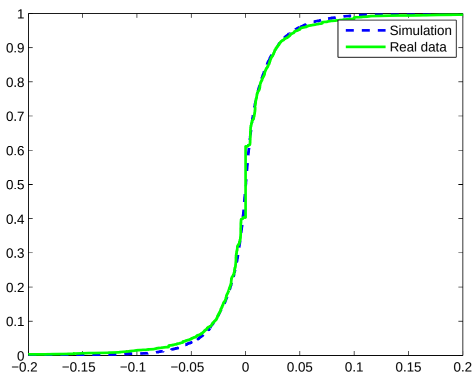
Model	CIR	GARCH	3/2	CKLS	Two factor CIR
KS Statistic	0.113101	0.123462	0.109700	0.118534	0.131148
L^2 Norm	$9.3171 \cdot 10^{-5}$	$8.1715 \cdot 10^{-4}$	$3.7261 \cdot 10^{-4}$	$7.4806 \cdot 10^{-4}$	$2.6048 \cdot 10^{-4}$

Table 9: KS statistics and the distance between model distributions and the empirical distribution as measured by the L^2-Norm.

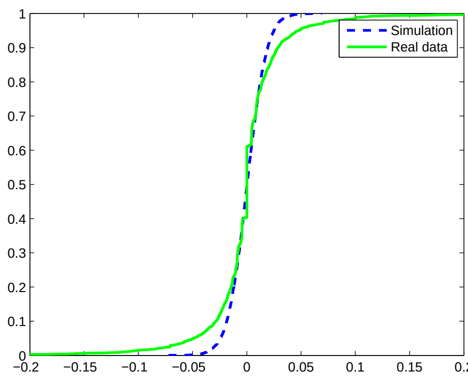
As can be seen in Table 9, the CIR model is the one that performs the best with the lowest L^2 norm. The KS statistic is on the other hand not the best and in addition is very high. A plausible explanation is that this is caused by the data, which has many increments of size zero. This because bonds and thus the interest rate are limited in how small increments can be. The allowed tic size (in effect the smallest size of increments in bond prices) is 1/32 percent. This result is that interest rate exhibits increments of zero size. As can be seen in the empirical distribution functions, the empirical distribution takes a jump in the middle. The largest distance in distribution, which is exactly what is measures by the KS statistic, will thus be at this point. The KS statistic is thus not optimal as a goodness of fit measure, the L^2 norm will most likely give a more fair description of goodness of fit.

In Figure 15 one can see that all but the CIR models have a rather poor fit to the empirical dustribution. For the two factor CIR model this is probably due to limitations in the model itself, as it does not support stochastic volatility. For the GARCH, 3/2 and CKLS models we suspect that parameter estimation has not worked as well as it could have. Especially the CKLS model should in a worst case scenario perform just as well as the CIR model. This is because the lever parameter γ is not fixed but can take any value. This makes it a more flexible model than the others and lets it take any parameter set that the CIR model does. Therefore it should be at least as good and in most cases better Then the CIR model.

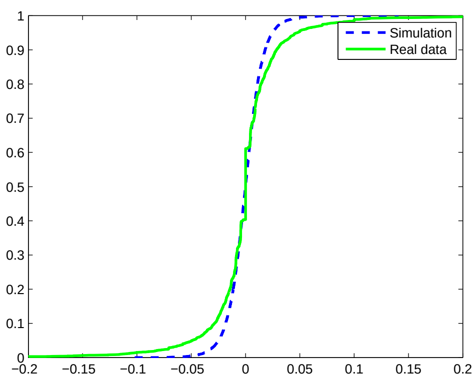
When simulating the GARCH, 3/2 and CKLS, there is a risk of obtaining extreme spikes in volatility if the time step is not short enough. If this happened during parameter estimation in the EMM method it could cause big damage to the estimates, because the EMM method is not very robust. For illustrative purposes, estimated variance processes from the different models as well as from the data are shown in Figure 16. As can be seen in the pictures, the CIR variance process in the only one that resembles the variance process of the data. This leads only increases our suspicion that parameter estimation has not worked as it was supposed to for the GARCH, 3/2 and CKLS models.



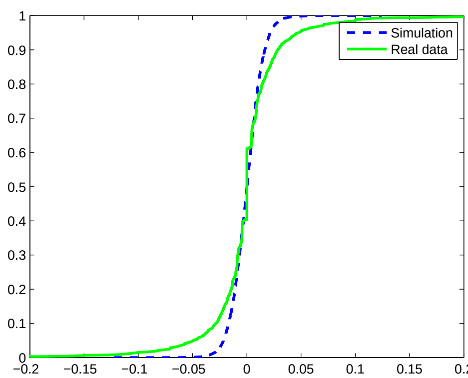
(a) CIR Distribution



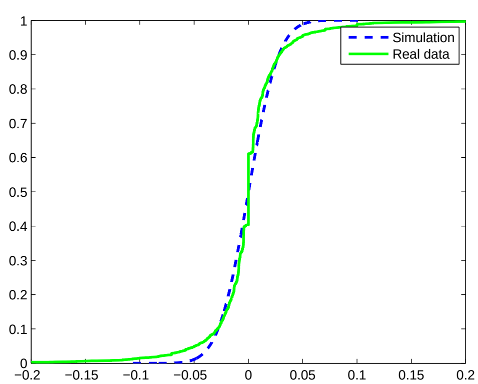
(b) GARCH Distribution



(c) 3/2 Distribution

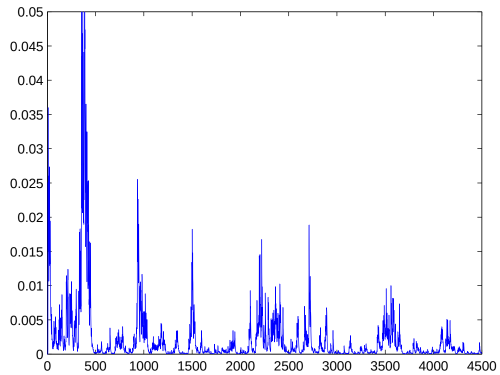


(d) CKLS Distribution

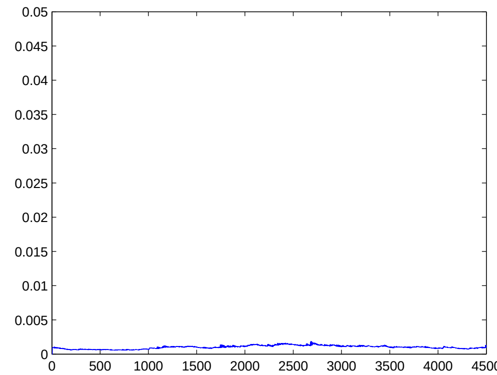


(e) Two Factor CIR Distribution

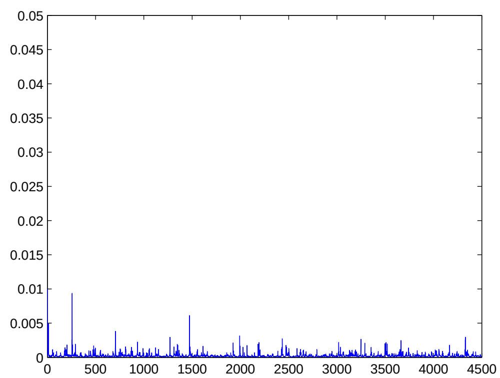
Figure 15: Cumulative stationary distributions of the volatility models compared to the empirical volatility.



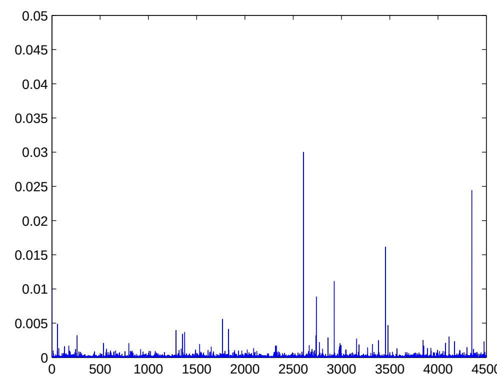
(a) CIR Variance Process



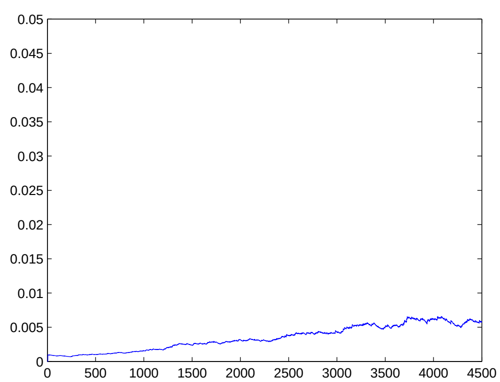
(b) GARCH Variance Process



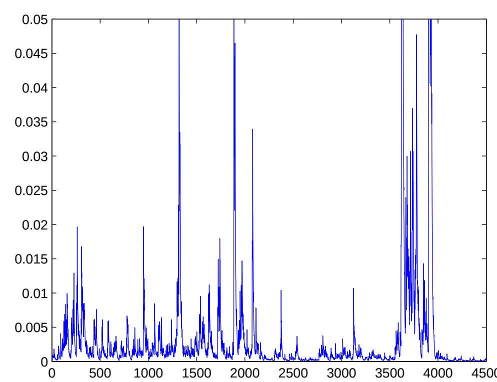
(c) 3/2 Variance Process



(d) CKLS Variance Process



(e) Two Factor CIR Variance Process



(f) Empirical Variance Process

Figure 16: Estimation of Simulated and Empirical Variance Processes

5.1.3 A Likelihood Test

The last test that is performed on the interest rate dynamics is a multidimensional likelihood test. Using the calibrated models, N=100 trajectories of each process are simulated. The variance process of each trajectory is then estimated using a GARCH process, as is done in previous sections. Then, for each time step in the trajectory the transition density function from time t to $t+1$ is estimated by simulating from the model SDE and fitting a non-parametric pdf using the MATLAB function "fitdist". Each increment in the trajectory is then given a likelihood value using the estimated transition density. For each of the 100 trajectories, the likelihood values are logarithmized and summed to give one likelihood value for each trajectory. The same process is then performed using the simulated model transition density but on the real interest rate data, to obtain a likelihood value based on the fitted model.

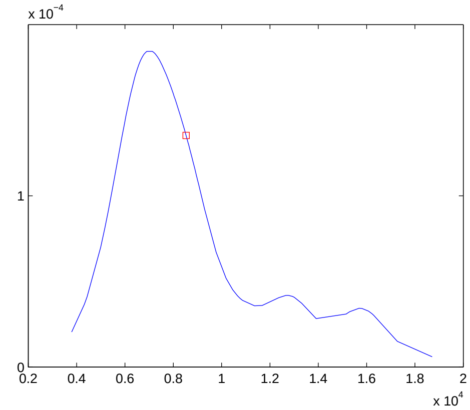
When all trajectories, including the real interest rate process are scored with a likelihood value, they are compared against each other. If the empirical likelihood is somewhere around the mean value of the simulated log likelihoods this is considered excellent. If it on the other hand is much lower than all simulated log likelihoods, this points towards the model being wrong. The empirical density functions of the likelihood values are visualized below, with the likelihood value of the real interest rate process marked in red. In these functions, the 100 likelihood values obtained from the simulation are smoothed with the MATLAB function "fitdist" to give a better illustration of the actual density.

As in the previous tests, Figure 17 shows that it is the CIR model that seems to perform best. The GARCH model Likelihood value is extremely much lower that the likelihood values from the simulations, indicating that this model is wrong.

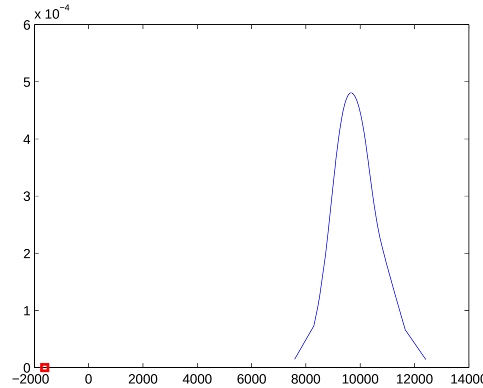
5.2 Evaluating Market behaviour

In this section we look at how well the different models perform in the pricing of interest rate derivatives. We have chosen to look at the pricing of swaptions because of the market's liquidity and the possibility of obtaining good market data. The performance of the models is evaluated using the mean average percentage deviation (MAPD) of model prices from market prices. The reason for using this measure is because prices will increase with the tenor (which is the time span of the underlying swap). This will cause models that perform well on shorter tenors to appear perform worse than they actually do, i.e. a bias in performance evaluation.

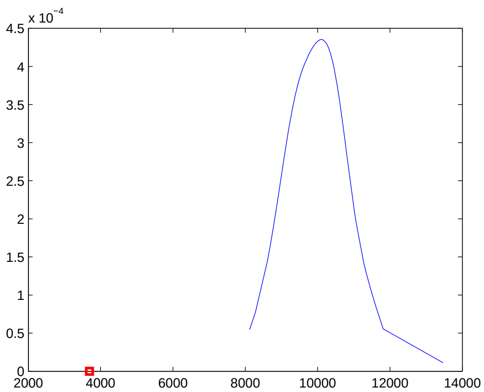
The options on the swaps have different times to maturity, ranging from 1 month to 30 years. We have chosen to limit the evaluation to contracts that do not range further into the future than around 20 years. This means that that the time to maturity of the option plus the time span of the underlying swap have to be less than or equal to 20 years. For



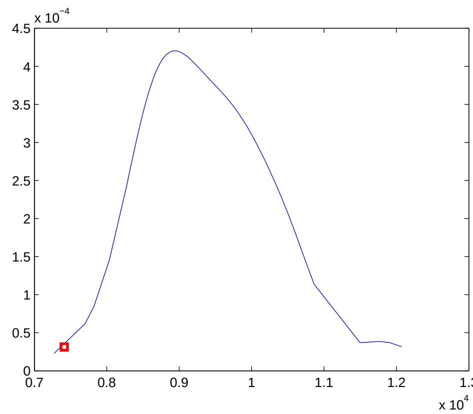
(a) CIR Likelihood Density



(b) GARCH Likelihood Density



(c) 3/2 Likelihood Density



(d) CKLS Likelihood Density

Figure 17: Densities from likelihood evaluations of simulated models. The likelihood of the real interest rate process according to the given model is marked as a square.

example a swaption with time to maturity of 5 years with underlying swap ranging 15 years equals a total of 20 years. The models were tested on swaptions with time to maturity of 1 month, 1 year and 5 years with underlying swaps of different maturities. The results are shown in Table 10 and are as mentioned before in the form of MAPD.

It can be seen in table 10 that the models perform quite poorly, especially the two factor CIR which deviations of up to 1200 percent, which is absolutely awful. The other models perform better, but are nowhere near an acceptable fit. Here again, it is the two-factor CIR stochastic volatility model that performs the best, with the best value showing 11.22 percent deviation from the market price on average.

We suspect that not all market information is available in the yield curve and that it is

T_0 / Model	CIR	GARCH	3/2	CKLS	CIR2	CIR2(Kalman)
1m	35.29	34.43	41.60	37.37	206.01	30.63
3m	27.31	46.04	53.74	39.55	293.52	11.22
6m	38.93	62.61	49.00	45.46	374.45	31.34
1y	37.89	68.55	72.03	44.46	416.00	20.54
2y	62.95	48.95	69.78	68.98	829.26	14.83
3y	63.96	33.87	71.37	67.45	1194.05	14.76

Table 10: Mean average percentage deviation of the model price from market swaption prices, calibrated to the yield curve.

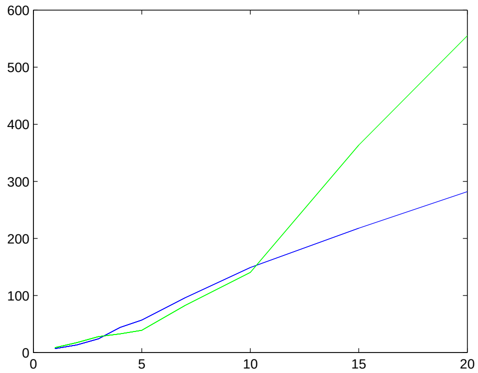
this that causes the awful prices above. When calibrating to the yield curve, it seems that it is mostly the average level of the interest rate that is take into account, not so much the volatility. This is supported by the good results of the two factor CIR model calibrated with the Kalman filter. In this method, the model is calibrated to not only the current yield curve, but the historical evolution of the yield curve over time. The Kalman filter will thus capture volatility of not only the short rate, but the entire yield curve which is why it performs so well in pricing.

To get a better evaluation of the models than the one above, we choose to calibrate the model directly to swaption prices to see how flexible the model is. As we saw before, the model was not flexible enough to fit the yield curve, it could only produce a simple curve. In Table 11, the MAPD values for this procedure are illustrated. Here swaption prices have been calibrated to all swaptions within a 5 year as well as a 15 year horizon.

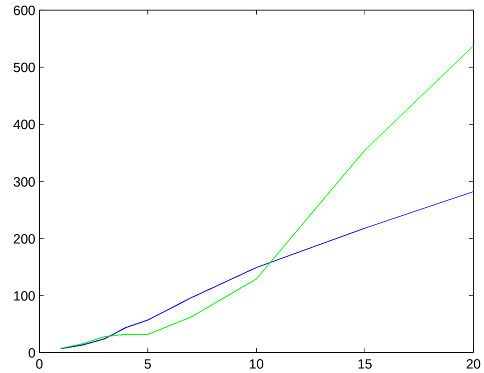
Range / Model	CIR	GARCH	3/2	CKLS	CIR2	CIR2(Kalman)	Heston
5 years	9.36	12.78	13.74	18.09	11.94	12.84	11.71
15 years	14.09	16.74	15.46	16.55	8.97	9.37	13.23

Table 11: Mean average percentage deviation of the model price from market swaption prices, calibrated to swaption prices.

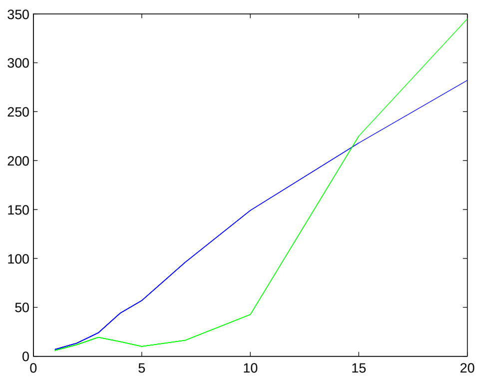
In the above Table 11 as well as in Figure 19 we can see that model and market prices coincide much better. Using this method of calibration the two factor CIR model is by far the most accurate model, with an average deviation of 8.97 percent over a 15 year period. These results point toward the conclusion that the models should be calibrated against swaption prices, or even to swaption prices and the yield curve simultaneously.



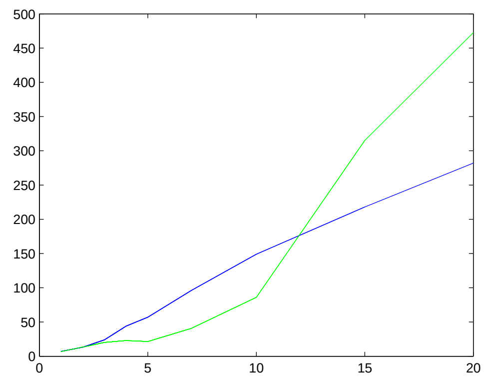
(a) CIR Swaption Prices



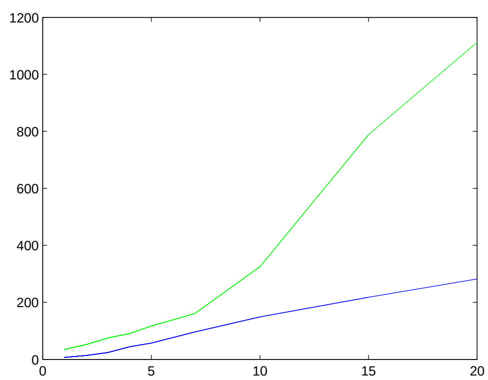
(b) GARCH Swaption Prices



(c) 3/2 Swaption Prices

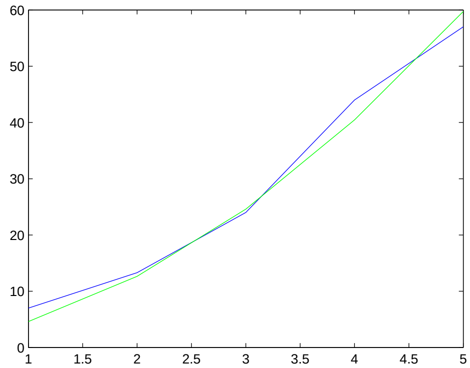


(d) CKLS Swaption Prices

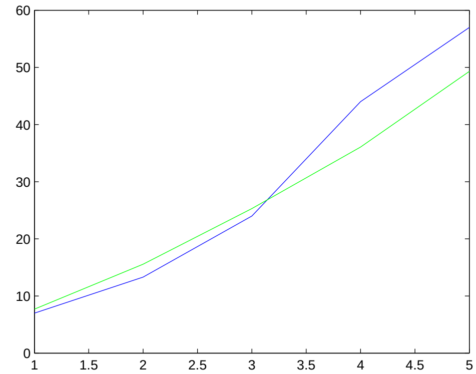


(e) Two Factor CIR Swaption Prices

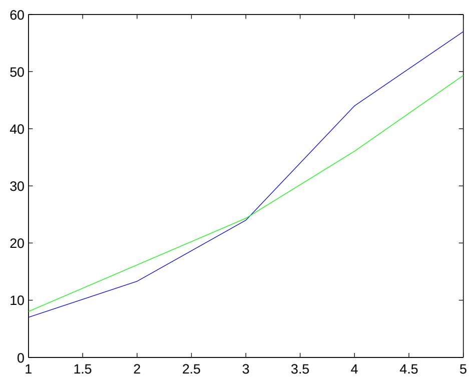
Figure 18: Predicted swaption prices form the respective models against real swaption prices.



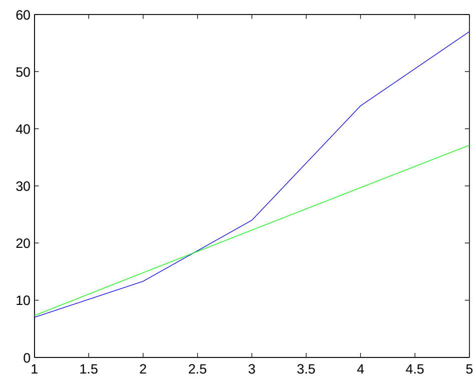
(a) CIR Swaption Prices



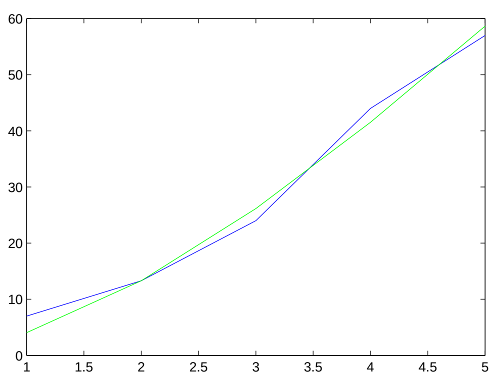
(b) GARCH Swaption Prices



(c) 3/2 Swaption Prices



(d) CKLS Swaption Prices



(e) Two Factor CIR Swaption Prices

Figure 19: Model swaption prices against market prices.

The GARCH, 3/2 and CKLS models perform worse than the other models here again. As was mentioned before, we suspect that is has to do with the difficulty in simulating the models, as well as the optimization procedure used. When simulating these processes some parameter combinations tend to cause the variance process to produce an extreme spike. These parameter sets need a shorter time step when simulated in order to not produce a spike. It is possible that spikes have occured during the optimization procedure that caused faulty parameter estimates here as well. The optimization algorithm might also have caused problems, this is because a gradient-free algorithm had to be used. This gradient-free algorithm can only hadle nice, convex problems and it is entirely possible that the optimization algorith converged to a local minimum.

5.3 Goodness-Of-Fit Test using Coefficient of Determination (R^2)

As the unknown parameters in Heston model are estimated using least squared method, the goodness-of-fit is better evaluate using similar principle. So, in this model, we choose R^2 method to see the goodness-of-fit of the model. R^2 is defined by

$$R^2 = 1 - \frac{\text{SSE}}{\text{SST}},$$

where

$$\text{SSE} = \sum (y - \hat{y})^2 \quad \text{and} \quad \text{SST} = \sum (y - \bar{y})^2,$$

where R^2 value describe how the model price describe the actual market price in the percentage form. Here we use the results obtained from the Heston model comparing with the actual Caplet price. The evaluation of the model are as shown in Table 12

T_0	R^2	R^2 (FFT)
6m	90.21	91.17
1y	98.76	98.32
2y	76.15	71.57
3y	53.16	28.03

Table 12: Coefficient of determination (R^2) percentage that the Caplet price can be explained by Heston model.

From Table 12, it can be seen that at 6 months and 1 year options time to maturity, the Caplet price can be explained by Heston model exceptionally well at 90.21% (91.17%) and 98.76% (98.32%) respectively and FFT value is in the parenthesis. The ability to explain decreases down at longer options time to maturity. Comparing the method of FFT and the quadratic one, quadratic seems to perform better by looking at the results overall.

However, if we look at the table 6, the quality of quadratic estimation came with the price of time taken. Hence, it purely depends on the user to decide which method suits best for ones' purpose.

We also test this goodness-of-fit method with other models and the results are shown in Table 13.

Range / Model	CIR	GARCH	3/2	CKLS	two factor CIR
5 years	98.85	90.69	91.57	92.67	98.95
15 years	98.59	95.74	96.97	97.43	98.80

Table 13: Coefficient of determination (R^2) for the respective models, calibrated to swaption prices.

6 Conclusion

In the tests of the short range dynamics it seems that the CIR model is the only model that cannot be rejected. It seems to work very well in all tests and when looking at a CIR variance process trajectory it closely resembles the empirical variance process. As the two factors CIR model for the interest rate has (nearly) constant volatility we can draw the conclusion that models of the interest rate dynamics are greatly improved by adding a stochastic volatility process.

As we suspect that the parameter estimates have been compromized for the GARCH, 3/2 and CKLS models we cannot really draw any conclusions about the effectiveness of these models. The only thing we can say is that the CIR model is considerably easier to simulate. In order to get good dynamics without any spikes, the CIR model requires only one time step per day. The other stochastic volatility models require at least 100 time steps per day in order to not produce spikes. For these small time steps the models are considerably slower than the CIR model, on the border of being infeasible if used on a regular PC. Because of the suspicious parameter estimates we recommend the GARCH, 3/2 and CKLS models be evaluated again to see if they perform better with new parameter estimations.

We could see that the CIR stochastic volatility model performed just as well as the two factor CIR model with constant volatility when it came to pricing. The GARCH, 3/2 and CKLS stochastic volatility models performed a bit worse, but were not far behind. This makes the choice of pricing model very easy, namely the two factor CIR model. The two factor CIR model has an analytical solution for the yield as well as to some derivatives, because of this it only takes a few minutes to calibrate. The CIR stochastic volatility model on the other hand takes many hours to calculate.

A thing that was not investigated was the numerical solution of the pricing PDE for the stochastic volatility models. This might have shortened the calibration time in that solving the PDE is potentially faster than simulating the really complex models.

While Heston model which is a stochastic volatility model assumes that the volatility is not constant but, on the other hand, randomized, has a closed-form solution. Using a simple statistical data test R^2 of model explanation, Heston seems to be working fine except when it encounters zero value. However, we could not say much if the precision is actually improved because of the stochastic process. In our opinion, we think Heston model is useful when all the information are available to incorporate into the model. As mentioned before, the heston model does not model interest rates, but instead it models bonds and bond options directly. In spite of this, the model performs very well in comparison.

In conclusion, it seems that stochastic volatility is not so significant to interest rate dynamics when used for pricing. The short rate dynamics are on the other hand greatly improved and the accuracy in pricing is nearly the same as with a constant volatility. We, therefore, recommend modelling the stochastic volatility when the goal is accurate short rate dynamics, but to leave it constant when using interest rate models for pricing.

We are not certain that short rate models are the best way to go when trying to price interest rate derivatives. The two factor CIR calibrated with the Kalman filter method tries to capture the dynamics of the entire yield curve. As this model performed very well, we feel that modelling the dynamics of the yield curve might be a better method than using a short rate model. This points in the direction of LIBOR market models that model the evolution of forward rates.

7 Appendix

7.1 Asset Pricing and Related Theorems

The Feynman-Kac Theorem. *Consider the stochastic differential equation*

$$dX(t) = \beta(t, X(t))dt + \gamma(t, X(t))dW(t)$$

Let h(y) be a Borel-measurable function. Fix T>0, and let $t \in$ [0,T] be given. Define the function

$$g(t, x) = E^{t,x}\Big[h(X(T))\Big]$$

(We assume that $E^{t,x}\big|h(X(T))\big| < \infty$ for all t and x.) Then g(t,x) satisfies the partial differential equation

$$g_t^{'}(t, x) + \beta(t, x)g_x^{'}(t, x) + \frac{1}{2}\gamma^2(t, x)g_{xx}^{''}(t, x) = 0$$

and the terminal condition

$$g(T, x) = h(x) \ \forall \ x$$

The No Arbitrage Pricing Formula. *Assume there exists an equivalent martingale measure for Q and let Y be an attainable contingent claim. Then for each time t, 0$\leq t \leq T$, there exists a unique price $\Pi(t)$ associated with Y, i.e.,*

$$\Pi(t) = E\Big[D(t, T)Y|\mathcal{F}_t\Big]$$

7.2 Heston Model Proof

$$C(S(t), V(t), t) = S(t)P_1 - Ke^{-r\tau}P_2$$

and by letting $x = \ln S$, the derivatives are

$$\begin{aligned}
\frac{\partial C}{\partial t} &= e^x\frac{\partial P_1}{\partial t} - Kre^{-r\tau}P_2 - Ke^{-r\tau}\frac{\partial P_2}{\partial t}, \\
\frac{\partial C}{\partial x} &= e^xP_1 + e^x\frac{\partial P_1}{\partial x} - Ke^{-r\tau}\frac{\partial P_2}{\partial x}, \\
\frac{\partial^2 C}{\partial x^2} &= e^xP_1 + 2e^x\frac{\partial P_1}{\partial x} + e^x\frac{\partial^2 P_1}{\partial x^2} - Ke^{-r\tau}\frac{\partial^2 P_2}{\partial x^2}, \\
\frac{\partial C}{\partial V} &= S\frac{\partial P_1}{\partial V} - Ke^{-r\tau}\frac{\partial P_2}{\partial V}, \\
\frac{\partial^2 C}{\partial V^2} &= S\frac{\partial^2 P_1}{\partial V^2} - Ke^{-r\tau}\frac{\partial^2 P_2}{\partial V^2}, \\
\frac{\partial^2 C}{\partial x\partial V} &= e^x\frac{\partial P_1}{\partial V} + e^x\frac{\partial^2 P_1}{\partial x\partial V} - Ke^{-r\tau}\frac{\partial^2 P_2}{\partial x\partial V}
\end{aligned}$$

By substituting these derivatives into PDE equation and set $K = 0$ and $S = 1$, we obtain,

$$\frac{1}{2}V\frac{\partial^2 P_1}{\partial x^2} + \rho\sigma V\frac{\partial^2 P_1}{\partial x \partial V} + \frac{1}{2}\sigma^2 V\frac{\partial^2 P_1}{\partial V^2} + (\frac{1}{2}V + r)\frac{\partial P_1}{\partial x} + [\kappa(\theta - V) - \lambda V + \rho\sigma V]\frac{\partial P_1}{\partial V} + \frac{\partial P_1}{\partial t} = 0$$

and by setting $S = 0$, $K = -1$ are $r = 0$, we get,

$$\frac{1}{2}V\frac{\partial^2 P_2}{\partial x^2} + \rho\sigma V\frac{\partial^2 P_2}{\partial x \partial V} + \frac{1}{2}\sigma^2 V\frac{\partial^2 P_2}{\partial V^2} + (-\frac{1}{2}V + r)\frac{\partial P_2}{\partial x} + [\kappa(\theta - V) - \lambda V]\frac{\partial P_2}{\partial V} + \frac{\partial P_2}{\partial t} = 0$$

which can be expressed by

$$\frac{1}{2}V\frac{\partial^2 P_j}{\partial x^2} + \rho\sigma V\frac{\partial^2 P_j}{\partial x \partial V} + \frac{1}{2}\sigma^2 V\frac{\partial^2 P_j}{\partial V^2} + (u_j V + r)\frac{\partial P_j}{\partial x} + (a - b_j V\frac{\partial P_j}{\partial V} + \frac{\partial P_j}{\partial t} = 0$$

for $j = 1, 2$, where

$$u_1 = 1/2, \quad u_2 = -1/2, \quad a = \kappa\theta, \quad b_1 = \kappa + \lambda - \rho\sigma, \quad b_2 = \kappa + \lambda.$$

7.3 Risk Neutral Parameter Estimates

Below are the risk neutral parameter estimates from section 4.1.4. The estimates were obtained by minimizing the distance from the model yield to the yield curve.

Term / Param.	b	a	d	c	η	γ
1	0.00104415	0.00042391	0.00002770	0.02927969	0.01048674	1/2
2	0.00088370	0.00044635	0.00002850	0.03147365	0.01107118	1/2
3	0.00088370	0.00046867	0.00002850	0.03147365	0.01107118	1/2
4	0.00094369	0.00050966	0.00002652	0.03340273	0.01027645	1/2
5	0.00108839	0.00048020	0.00002570	0.03624724	0.00963793	1/2
7	0.00135051	0.00037725	0.00002599	0.03178235	0.00965817	1/2
10	0.00160118	0.00038089	0.00002424	0.03049744	0.00926495	1/2
15	0.00175196	0.00033543	0.00002377	0.03443598	0.00893111	1/2
20	0.00174837	0.00033713	0.00002449	0.03495618	0.00891611	1/2

Table 14: Risk neutral parameter estimates for the CIR model

Term / Param.	b	a	d	c	η	γ
1	0.00001205	0.00000462	0.00011839	1.10540005	0.19369786	1
2	0.00000976	0.00000510	0.00012303	1.10690552	0.20359387	1
3	0.00000944	0.00000534	0.00012740	1.09829954	0.20183879	1
4	0.00000979	0.00000505	0.00012900	1.11427707	0.20475900	1
5	0.00001063	0.00000496	0.00013447	1.18076907	0.18626191	1
7	0.00001362	0.00000426	0.00013382	1.00918410	0.17252970	1
10	0.00001612	0.00000353	0.00010616	1.04670545	0.19721594	1
15	0.00001581	0.00000244	0.00011035	1.24193492	0.19528274	1
20	0.00001575	0.00000243	0.00011153	1.26541525	0.19751018	1

Table 15: Risk neutral parameter estimates for the GARCH model

Term / Param.	b	a	d	c	η	γ
2	0.00095362	0.00040501	0.00001269	-0.05398962	45.20501400	3/2
3	0.00093454	0.00041635	0.00001304	-0.05172206	46.47075440	3/2
4	0.00094143	0.00040764	0.00001327	-0.05252701	46.08809818	3/2
5	0.00103056	0.00037730	0.00001363	-0.05188089	47.92409507	3/2
7	0.00129995	0.00032119	0.00001296	-0.05076395	49.28744057	3/2
10	0.00151917	0.00026672	0.00001265	-0.05055119	52.97489909	3/2
15	0.00160287	0.00025318	0.00001246	-0.04974459	56.37307076	3/2
20	0.00158003	0.00025623	0.00001261	-0.05034401	56.71412784	3/2

Table 16: Risk neutral parameter estimates for the 3/2 model

Term / Param.	b	a	d	c	η	γ
1	0.00122305	0.00057165	0.00597972	44.41835977	27.93390488	1.12951698
2	0.00099081	0.00060063	0.00648536	47.10722789	28.97191671	1.12848968
3	0.00093724	0.00064651	0.00653192	47.54272668	29.80602119	1.12892813
4	0.00101601	0.00061768	0.00616414	51.26663158	30.00697509	1.12876705
5	0.00110051	0.00052169	0.00636981	57.09936238	29.24757752	1.12860603
7	0.00136956	0.00044727	0.00655155	49.44138959	27.72786965	1.12963578
10	0.00155533	0.00033762	0.00613510	52.48279727	29.39393076	1.13010149
15	0.00168626	0.00030741	0.00624735	52.72508959	29.77721213	1.13035793
20	0.00169009	0.00031434	0.00638622	52.89269606	29.78293456	1.12919458

Table 17: Risk neutral parameter estimates for the CKLS model

Term / Param.	b_1	a_1	σ_1	b_1	a_1	σ_1
1	0.00001068	0.00004184	0.00729997	0.00114187	0.00001746	0.02023978
2	0.00001058	0.00004486	0.00762195	0.00086688	0.00001914	0.02101880
3	0.00001042	0.00004652	0.00760068	0.00082793	0.00001923	0.02227276
4	0.00001056	0.00004756	0.00763829	0.00082953	0.00001957	0.02269377
5	0.00001048	0.00004970	0.00755190	0.00091066	0.00001967	0.02167231
7	0.00001096	0.00004388	0.00695753	0.00112392	0.00001899	0.02108767
10	0.00001086	0.00004780	0.00667100	0.00129536	0.00001811	0.02160815
15	0.00001090	0.00004795	0.00685929	0.00130379	0.00001817	0.02167784
20	0.00001125	0.00004644	0.00714662	0.00121355	0.00001872	0.02255263

Table 18: Risk neutral parameter estimates for the two factor CIR model

Bibliography

[1] T. G. Andersen and J. Lund. Estimating continous-time stochastic volatility models of the short-term interest rate. *Journal of Econometrics*, 77:343–377, 1997.

[2] J. Bak, H. A. Nielsen, and H. Madsen. Goodness of fit of stochastic differential equations, 1999.

[3] D. Brigo and F. Mercurio. *Interest Rate Models - Theory and Practice*. Springer Finance, 2nd edition, August 2006.

[4] P. Carr and D. B. Madan. Option valuation using the fast fourier transform. *Journal of Computational Finance*, 2(4), 1999.

[5] R. R. Chen and L. Scott. Multi-factor cox-ingersoll-ross models of the term structure. Master's thesis, School of Business, Rutgers University, September 2002.

[6] J. C. Cox, J. E. Ingersoll, and S. A. Ross. A theory of the term structure of interest rates. *Econometrica*, 53(2):385–407, March 1985.

[7] A.R. Gallant and G. Tauchen. Which moments to match. *Econometric Theory*, 12(4):657–681, October 1996.

[8] S. L. Heston. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *The Review of Financial Studies*, 6(3):327–343, 1993.

[9] G. Hong. Forward smile and derivative pricing. *Equity Quantitative Strategists Working Paper, UBS*, 2004.

[10] J. Hull and A. White. Pricing interest-rate-derivative securities. *The Review of Financial Studies*, 3(4):573–592, 1990.

[11] N. Moodley. The heston model: A practical approach. Master's thesis, University of the Witwatersrand, 2005.

[12] K. Ben Nowman. The volatility of japanese interest rates, evidence for certificate of deposit and gensaki rates. *International Review of Financial Analysis*, 11:29–38, 2002.

[13] E. M. Stein and J. C. Stein. Stock price distributions with stochastic volatility: An analytic approach. *Review of Financial Studies*, 4(4), 1991.

[14] O. Vasicek. An equilibrium charachterization of the term structure. *Journal of Financial Economics*, 5:177–188, 1977.