Network for Studies on Pensions, Aging and Retirement

Netspar THESES

Emile van Elen

Term Structure Forecasting

Does a Good Fit Imply Reasonable Simulation Results?

BSc Thesis 2010

Term structure forecasting

Does a good fit imply reasonable simulation results?

by
Emile A.L.J. van Elen (74.30.15)

A thesis submitted in partial fulfillment of the requirements for the degree of Bachelor in Econometrics and Operations Research.

Tilburg School of Economics and Management
Tilburg University

Supervised by:
T.G.E. van der Heijden MSc

Date: July 24, 2010

Abstract

In this paper I question whether term structure models of interest rates used for pricing derivative instruments are suited to use in a simulation based context like an asset liability management (ALM) study. I consider the Vasicek, Cox-Ingersoll-Ross (CIR) and Nelson-Siegel model. First, I discuss the models and their merits and drawbacks. Second, I estimate the parameters of all three models. Estimates are based on monthly Canadian zero-coupon yields from 1986 up until and including 2009. I estimate the Vasicek model twice: once using time series of the two-year yields and once using cross-sectional data (i.e., the different maturity times). I calibrate the CIR model using time series of the two-year yields and the Nelson-Siegel model using cross-sectional data. Third, I simulate the models. I simulate the Vasicek and CIR short rate processes by their discrete representations, whereas I simulate the factors of the Nelson-Siegel with a first-order vector autoregressive (VAR(1)) model. I judge the term structure models' empirical fit as well as dynamics in the ALM study and find that the Nelson-Siegel model outperforms the Vasicek and CIR models.

Contents

1 Introduction

Scientists, banks, pension funds, trading houses and governmental institutions are all very interested in the future development of the term structure of interest rates. Trading houses, for instance, need estimations of interest rates in order to price derivative instruments. For pension funds, future interest rates are of high interest to estimate the value of their assets and liabilities. The main topic of this paper is the question whether term structure models used for pricing derivative instruments are suited to use in a simulation based context like an asset liability management (ALM) study over a long investment horizon e.g., forty years.

2008 and 2009 showed unprecedented drops in the funding rate (i.e., the assets over the liabilities) of the few largest and generally considered stable, Dutch pension funds. For instance, the funding rate of 'Pensioenfonds Metaal & Techniek' (PMT), the largest Dutch pension fund for employees in metalworking and other technical branches, shocked the nation by reporting a funding rate of a poor 80 percent at the end of February, 2009 (PMT, 2009).

A decrease in the funding rate can either be caused by a decrease of the assets, an increase of the liabilities or both. In the current credit crunch, it is both. On the one hand, assets declined up to 40 percent in some cases due to plunging stock markets. On the other hand, the most important factor: central banks lowered interest rates to nearly zero percent. Since all future payments of a pension fund need to be discounted over multiple decades, one can imagine the impact a change of the interest rate will have. That is, a drop of the interest rate of one percent causes the funding rate of an average pension fund to drop by 15 to 20 percent. Clearly, having a good estimate of interest rates is of high importance to get a decent insight into the pension fund's liabilities.

Pension funds often chose to work with the constant actuarial interest rate of four percent. Until the beginning of the second millennium, that is. Poor financial circumstances resulted in decreasing funding rates of the Dutch pension funds. Therefore, the watchdog of the Dutch pension sector (Pensioen- en Verzekeringskamer, incorporated into De Nederlandsche Bank as of 2004) decided to tighten their policy. The new rules, known as FTK (Financieel ToetsingsKader) were adopted in 2007 by the Second Chamber into the Dutch pension's law, 'Pensioenwet'. The most radical change was the obligatory way of discounting the liabilities using a term structure model instead of the constant actuarial rate.

Several term structure models have been developed in the course of time. A model that forms the basis of many other term structure models is the Vasicek model (1977). The innovating feature of the Vasicek model is that mean reversion of the interest rate (see section 2.3) is captured by the model. A famous extension to the Vasicek model is the Cox-Ingersoll-Ross or CIR model (1985), which aims to cope with some of the drawbacks of the Vasicek model. Other famous extensions are the multiple-factor (most notably two-factor) Vasicek, multiple-factor CIR, Hull-White (1990) and Black-Derman-Toy models (1990).

From a mathematical point of view totally different models are the three-factor Nelson-Siegel model (1987) and its variations e.g., the four-factor Svensson model (1994). Although the standard Nelson-Siegel model does not satisfy the no-arbitrage property, as shown by Bjork and Christensen (1999), it is widely used due to its good fit of the observed term structure. For instance, De Pooter (2007) states that nine out of thirteen central banks that report their curve estimation methods to the Bank of International Settlements use either the Nelson-Siegel model or a variation. Moreover, The European Central Bank (ECB) models the term structure with an extended Nelson-Siegel model (Coroneo, Nyholm and Vidova-Koleva, 2008).

Another class of term structure models is the Heath-Jarrow-Morton (HJM) framework (1987). The main difference with the other term structure models like e.g., the Vasicek model, is that HJM models capture the full dynamics of the entire yield curve whereas other models only capture the dynamics of a single point on the curve. Models in the HJM framework are non-markovian (as opposed to the other models discussed so far), making the HJM models in general computationally intractable.

The KISS principle of forecasting - Keep It Sophisticatedly Simple - (Zellner, 1992) makes its mark on this paper. That is, this paper discusses the Vasicek, CIR and Nelson-Siegel model, whereas only little attention is paid to extented versions of these models. To find out which of these three term structure models are suited in a simulation based context, one requires estimations of the models' parameters. I estimate the models using monthly Canadian zero-coupon data from 1986 up until and including 2009. Estimations are made through either ordinary least squares (OLS) regression, via maximum likelihood, or both. I estimate the Vasicek model twice: once using time series of the two year interest rates and once using cross-sectional data (i.e., the different maturity times). The CIR model is calibrated using time series of the two year interest rates and the Nelson-Siegel model using cross-sectional data.

I simulate the term structure with each of the three models. One motivation to simulate interest rates may be to examine the out-of-sample performance of the term structure models. An interesting read on this topic for the Nelson-Siegel model is Diebold and Li (2006). They indicate that the model produces term structure forecasts at both short and long horizons "... with encouraging results. In particular, our forecasts appear much more accurate at long horizons than various standard benchmark forecasts." Duffie (2002) forecasts term structures with the standard class of affine models. He concludes this class to forecast the term structure poorly and indicates that random walks are better predictors. Not treated very often in the literature but highly interesting, is the question whether simulation results match the larger trends and statistics (i.e., stylized facts) of the actual interest rate data. This paper aims to contribute answering this question. Therefore, I construct term structures with both the Vasicek and CIR model by simulating the short rate. I simulate the factors of the Nelson-Siegel model to deduce the full term structure using a first-order vector autoregressive (VAR(1)) model.

The remainder of this paper is structured as follows. In Section 2 I shortly elaborate on elements that are of key importance to some term structure models. Section 3 discusses the three models and their merits and drawbacks. Section 4 describes the Canadian interest rate data and estimates the parameters of the models. I present the forecasting of the term structure models in Section 5. Section 6 concludes the paper.

2 Concepts & Notations

2.1 Introduction

This section discusses properties (like the Wiener process and mean reversion) that occur in some term structure models. For the sake of readability, I use the same notations for parameters in the different models throughout the entire paper unless stated otherwise.

2.2 Wiener Process

The Wiener process (also referred to as a Brownian motion) is a continuous-time stochastic process with small, independent increments. In finance the Wiener process is used to describe changes in prices of options or to describe changes in interest rates.

A continuous-time process $\{W_t\}$ is said to be a Wiener process if it satisfies the following three properties (Schumacher, 2009):

1. $W_0 = 0$.
2. If $t_1 < t_2 \leq t_3 < t_4$, then the increments $W_{t_2} - W_{t_1}$ and $W_{t_4} - W_{t_3}$ are independent.
3. For any given t_1 and t_2 with $t_2 > t_1$, the distribution of the increment $W_{t_2} - W_{t_1}$ is the normal distribution with mean 0 and variance $t_2 - t_1$.

2.3 Mean Reversion

A process is said to be mean reverting if the process tends to fall (rise) after hitting a maximum (minimum). Interest rates are known to be subject to mean reversion over a longer horizon. Therefore, some models in this paper use an Ornstein-Uhlenbeck process for the dynamic evolution of interest rates. The Ornstein-Uhlenbeck process is given by the following stochastic differential equation (SDE):

$$dx_t = \theta(\mu - x_t)dt + \sigma dW_t. \tag{1}$$

Here, $\theta, \mu, \sigma > 0$, and W_t is a Wiener process. Note that in (1), μ denotes the mean to which the process will revert and that θ indicates the speed at which this revertion takes place. Furthermore, the amount of randomness is indicated by parameter σ. Also note that whenever x_t will be high (i.e., larger than μ), the process is likely (likely rather than surely because of the randomness involved) to move downwards. The opposite also holds true.

2.4 Risk Neutral World

In the world as we know it, investors are assumed to be risk-averse. This implies that investors bearing risk want to get compensated for it. In the risk neutral world, as the name suggests, investors are indifferent to risk. Therefore, investors bearing risk in the risk neutral world do not need a compensation now.

Under the assumption of no-arbitrage, the current value of all financial assets in the risk neutral world equals the expected value of the future payoffs discounted at the risk-free rate. Let λ denote the market price of risk. In order to capture λ into a model, the real world Wiener process $\{\tilde{W}_t\}$ is defined as:

$$\tilde{W}_t = W_t - \lambda t. \tag{2}$$

Here, W_t denotes the Wiener process in the risk neutral world.

3 Term Structure Models

3.1 Introduction

Term structures of interest rates describe the relation between interest rates and bonds with different maturity times. This section elaborates on three models that are aimed at describing the term structure. Two of them - the Vasicek and Cox-Ingersoll-Ross model - are called short rate models, as they model the dynamics of the instantaneous short rate r_t directly. The third model, being the Nelson-Siegel model, merely tries to fit the observed term structure. The three models describe the dynamic evolution of the yield curve.

The short rate is the annualized interest rate for an infinitesimally short period of time. In practice, however, the three-month rate is considered a better approximation of the short rate because, for example, overnight loans are affected by factors that term structure models do not aim to cover (Schumacher, 2009). In this paper the short rate r_t is defined as

$$r_t = R(t, 0) = \lim_{T \to 0} R(t, T),$$

where t denotes a moment in time, T the time to maturity and $R(t, T)$ the corresponding interest rate.

Let $P(t, T)$ denote the value of a zero coupon bond at time t that pays one at maturity time T and $R(t, T)$ the corresponding interest rate. In continuous time we then find:

$$P(t, T) = \exp\Big(- R(t, T)(T - t)\Big). \tag{3}$$

Rewriting (3) yields a way to describe the interest rate as a function of the value of a bond:

$$R(t, T) = -\frac{1}{T - t} \log\Big(P(t, T)\Big). \tag{4}$$

Since $P(t, T)$ is a simple discount factor, it is clear that $R(t, T)$ may not be negative to make sure the discount factor as described in (3) lies between zero and one. Let us discuss the model that builds on this introduction, the Vasicek model.

3.2 Vasicek Model

Vasicek (1977) develops a one-factor model which depends on only one uncertainty: the short rate. Vasicek defines the short rate process as:

$$dr_t = \theta\big(\mu - r_t\big)dt + \sigma dW_t. \tag{5}$$

As with the mean reverting process mentioned earlier, θ, μ, σ are strictly positive and W_t is a Wiener process.

If one wants to take the market price of risk into account, the stochastic differential equation in (5) should be adjusted for the risk neutral Wiener process as defined in (2). One then finds:

$$\begin{aligned} dr_t &= \theta\Big(\mu - r_t\Big)dt + \sigma d(W_t - \lambda t) \\ &= \theta\Big(\mu - r_t\Big)dt + \sigma dW_t - \sigma\lambda dt \\ &= \theta\Big(\Big[\mu - \frac{\lambda\sigma}{\theta}\Big] - r_t\Big)dt + \sigma dW_t. \end{aligned} \tag{6}$$

It can be shown via the Black-Scholes partial differential equation associated to the Vasicek model that the Vasicek bond pricing model, indeed, only has three parameters, namely: θ, σ and $\theta\mu - \lambda\sigma$. Schumacher (2009) gives an example of different parameters that yield the same bond prices: $\big(\theta, \mu, \sigma, \lambda\big)$ and $\big(\theta, \mu - \frac{\lambda\sigma}{\theta}, \sigma, 0\big)$. To simplify calculations, one may therefore take $\lambda = 0$. In this case μ would be a "risk-adjusted" representation of the long-term average of the short rate.

Vasicek shows that the yield curve (i.e., term structure) is a function of some constant and the short rate. Schumacher (2009) gives a representation of this, where the price of a zero-coupon bond is a function of the short rate:

$$P(t,T) = \exp\bigg(-\Big[\big(\mu - \frac{\sigma^2}{2\theta^2}\big)\big(T-t\big) + \big(r_t - \mu + \frac{\sigma^2}{\theta^2}\big)\frac{1-e^{-\theta(T-t)}}{\theta} - \frac{\sigma^2}{2\theta^2}\frac{1-e^{-2\theta(T-t)}}{2\theta}\Big]\bigg).$$

Substituting this expression for $P(t,T)$ in (4) enables one to properly denote the interest rate as a function of t:

$$R(t,T) = \Big(\mu - \frac{\sigma^2}{2\theta^2}\Big) + \Big(r_t - \mu + \frac{\sigma^2}{\theta^2}\Big)\frac{1-e^{-\theta(T-t)}}{\theta(T-t)} - \frac{\sigma^2}{2\theta^2}\frac{1-e^{-2\theta(T-t)}}{2\theta(T-t)}. \tag{7}$$

A major advantage of Vasicek's model compared to, for instance, a constant interest rate is that the model is able to reconstruct different shapes of the term structure, which occur in reality sometimes. For a large part this is due to the mean reversion which is captured by the drift in Vasicek's model. Different shapes (up- or downwarding curve or a humped shaped curve) are computed by taking different parameter values in (7). Figure 1 plots these three shapes. Table 1 denotes the corresponding paramater values.

Table 1: Parameter values for the Vasicek Model.

Parameter	Downward sloping	Upward sloping	Hump shaped
r_t	0.09	0.02	0.06
μ	0.08	0.08	0.08
σ	0.05	0.05	0.05
θ	0.2	0.2	0.2
λ	0	0	0

A drawback of the Vasicek model is that the model can produce negative interest rates. If real interest rates (i.e., interest rates corrected for inflation) are to be modeled, this does not necessarily have to be a big problem as real interest rates can be negative in reality. Nominal rates, on the contrary, will

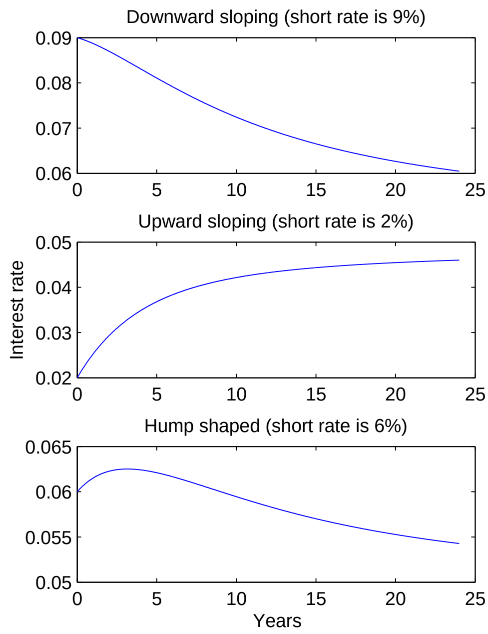

Figure 1: Different shapes of the term structure using the Vasicek model.

never be negative in practice. To show that negative interest rates might occur, a discrete version of the model is needed. Therefore, the short rate, on a time grid $0 = t_0, t_1, t_2, \ldots$ with time step $\Delta t = t_i - t_{i-1}$ is now defined as

$$r_{\Delta t} = \theta\Big(\mu - r_t\Big)\Delta t + \sigma\sqrt{\Delta t}Z. \tag{8}$$

Here, Z is standard normally distributed. In contrast with models used to price stocks, term structure models are additive:

$$r_{t+\Delta t} = r_t + r_{\Delta t} = r_t + \theta(\mu - r_t)\Delta t + \sigma\sqrt{\Delta t}Z. \tag{9}$$

We can now find an expression for the probability that $r_{t+\Delta t}$ is non-positive. That is,

$$\begin{aligned}
\mathbb{P}\Big(r_{t+\Delta t} \leq 0\Big) &= \mathbb{P}\Big(r_t + \theta(\mu - r_t)\Delta t + \sigma\sqrt{\Delta t}Z \leq 0\Big) \\
&= \mathbb{P}\left(Z \leq \frac{-r_t - \theta(\mu - r_t)\Delta t}{\sigma\sqrt{\Delta t}}\right) \\
&= \Phi\left(\frac{-r_t - \theta(\mu - r_t)\Delta t}{\sigma\sqrt{\Delta t}}\right).
\end{aligned} \tag{10}$$

Taking for example the parameters as in table 2 at $t = 0$ with $\Delta t = 0.01$, I find:

Table 2: Examples values

Parameter	Values
r_t	0.02
μ	0.03
σ	0.15
θ	0.2
λ	0

$$\begin{aligned}\mathbb{P}\Big(r_{0.01} \leq 0\Big) &= \Phi\Big(\frac{-0.02 - 0.2(0.03 - 0.02)0.01}{0.15 * \sqrt{0.01}}\Big) \\ &\approx \Phi\Big(-1.335\Big) \\ &\approx 0.091.\end{aligned}$$

Figure 2 shows negative short rates produced by the Vasicek model. Here, a simulation with 250 time steps of the short rate is performed. On various moments in time, the short rate becomes negative. It is noted that negative interest rates only form a problem in the short run. When looking at the conditional expectation of the Ornstein-Uhlenbeck process (11) it indeed turns out that, whenever $T \to \infty$, the exponential in (11) and hence the complete second term becomes zero, which implies the long-term expectation to be equal to μ:

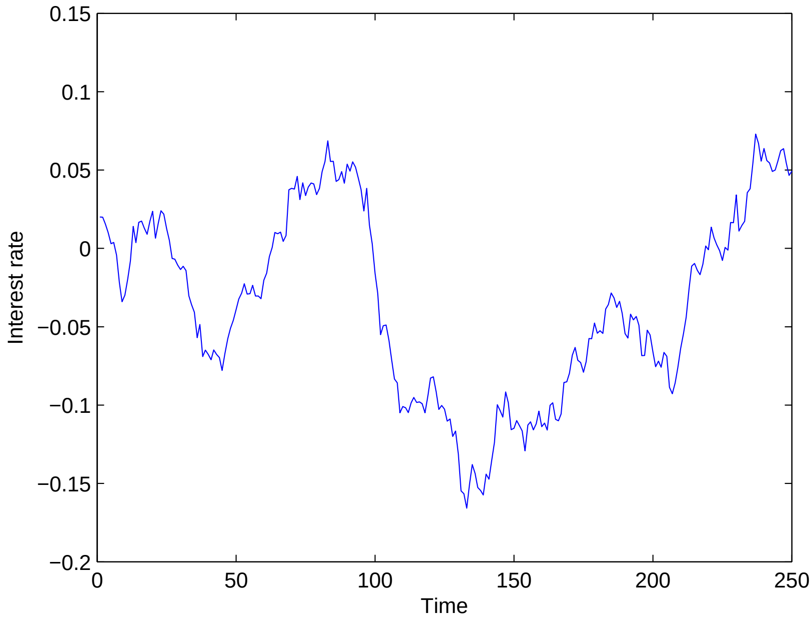

Figure 2: Simulation of the short rate using the Vasicek model.

$$\mathbb{E}_t\big(r_T\big) = \mu + \big(r_t - \mu\big) \cdot e^{-\theta(T-t)}. \tag{11}$$

3.3 Cox-Ingersoll-Ross (CIR) Model

Cox, Ingersoll and Ross (1985) propose an extension to the Vasicek model to prevent the short rate from becoming negative:

$$dr_t = \theta(\mu - r_t)dt + \sigma\sqrt{r_t}dW_t. \tag{12}$$

Again, paramaters θ, μ and σ are positive and W_t follows a Brownian motion. The first term of (12), the drift term, represents the mean reversion and is similar to the drift term in the Vasicek model. The difference between (12) and (5), is the square root in the second (or volatilty) term, which prevents the short rate from becoming negative. Indeed, when r_t approaches zero, the volatility term $\sigma\sqrt{r_t}$ approaches zero. In this case the short rate will only be affected by the drift term, resulting the short rate to revert to the mean again. In fact, Cox *et al.* (1985) show that whenever $2\theta\mu > \sigma^2$, the interest rate is strictly larger than zero. Furthermore, there is empirical evidence that whenever interest rates are high, the volatility is likely to be high as well, which justifies the volatility term in the CIR model.

Taking the market price of risk λ into account one finds, in analogy to the derivation of (6),

$$dr_t = \theta\Big(\Big[\mu - \frac{\lambda\sigma}{\theta}\Big] - r_t\Big)dt + \sigma\sqrt{r_t}dW_t,$$

where W_t represents a Wiener process in the risk neutral world.

It can be shown that the pricing function of a zero-coupon paying bond in the CIR model is given by

$$P(t,T) = A(t,T)e^{-B(t,T)r_t} \tag{13}$$

where

$$A(t,T) = \left[\frac{2\eta \cdot \exp\left\{\frac{1}{2}(\theta+\eta)(T-t)\right\}}{2\eta + (\theta+\eta)\cdot\big(\exp\left\{(T-t)\eta\right\} - 1\big)}\right]^{2\theta\mu/\sigma^2}, \tag{14}$$

$$B(t,T) = \frac{2\big(\exp\left\{(T-t)\eta\right\} - 1\big)}{2\eta + \big(\theta+\eta\big)\cdot\big(\exp\left\{(T-t)\eta\right\} - 1\big)} \tag{15}$$

and

$$\eta = \sqrt{\theta^2 + 2\sigma^2}. \tag{16}$$

Rewriting the expression for $P(t,T)$ in (13) and substituting it in (4), yields a function to compute the term structure in the CIR model:

$$R(t,T) = \frac{1}{T-t}\Big[\big(B(t,T)\cdot r_t\big) - \log\big(A(t,T)\big)\Big], \tag{17}$$

with $A(t,T)$ as in (14), $B(t,T)$ as in (15) and η as in (16).

On a time grid $0 = t_0, t_1, t_2, \ldots$ with time step $\Delta t = t_i - t_{i-1}$, the discretized version of the CIR model is defined as

$$r_{t+\Delta t} = r_t + \theta(\mu - r_t)\Delta t + \sigma\sqrt{\Delta t}\sqrt{r_t}Z \tag{18}$$

with $Z \sim N(0,1)$. The three different shapes of the term structure that could be computed by the Vasicek model, can also be computed by the CIR model by changing the parameters of (17).

3.4 Nelson-Siegel Model

Nelson and Siegel (1987) introduce a model that from a mathematical point of view deviates significantly from the classical term structure models (e.g., the Vasicek and CIR models). They propose the following three-factor term structure model:

$$R_t(\tau) = \beta_{1,t} + \beta_{2,t}\frac{1-e^{-\kappa_t\tau}}{\kappa_t\tau} - \beta_{3,t}e^{-\kappa_t\tau}. \tag{19}$$

Here, $\beta_{1,t}, \beta_{2,t}, \beta_{3,t}$ and κ_t are time-varying parameters. $R_t(\tau)$ denotes the zero coupon rate at observed time t with maturity time τ.

Diebold and Li (2006) reformulate the original Nelson-Siegel equation (19) in order to be able to give easy economic interpretations to the parameters $\beta_{1,t}, \beta_{2,t}$ and $\beta_{3,t}$. The Diebold-Li representation is given by

$$R_t(\tau) = \beta_{1,t} + \beta_{2,t}\frac{1-e^{-\kappa_t\tau}}{\kappa_t\tau} + \beta_{3,t}\left(\frac{1-e^{-\kappa_t\tau}}{\kappa_t\tau} - e^{-\kappa_t\tau}\right). \tag{20}$$

Diebold, Ji and Li (2006) state that in (20), $\beta_{1,t}, \beta_{2,t}$ and $\beta_{3,t}$ can be interpreted as a level factor, a slope factor and a curvature factor, respectively. To see this it might be handy to interpret the loadings. First of all, note that the loading of $\beta_{1,t}$ equals one which is a constant independent of the maturity, τ. Hence, the term structure at different maturities is affected by $\beta_{1,t}$ equally, which justifies the interpretation of $\beta_{1,t}$ as a level factor. Secondly, the loading of $\beta_{2,t}$ is given by $(1-e^{-\kappa_t\tau})/\kappa\tau$. Clearly, whenever $\tau \to 0$, the loading approaches one, and if $\tau \to \infty$, the loading converges to zero. Therefore, the yield curve is primarily affected by $\beta_{2,t}$ in the shorter run, so a change in $\beta_{2,t}$ implies a change in the slope of the term structure. Finally, the loading that comes with $\beta_{3,t}$ equals $\big(1-e^{-\kappa_t\tau}\big)/\kappa_t\tau - e^{\kappa_t\tau}$. Plotting this loading as a function of time, will make the graph start in zero, gradually increase and converge to zero again. Diebold *et al.* (2006) thus call $\beta_{3,t}$ the curvature factor as it affects the curvature of the term structure and has the greatest impact on medium-term yields.

The fourth parameter, κ_t indicates the rate of the exponential decay. That is, a small κ_t causes a slow decay of the term structure and it can better fit the yield curve at long maturities. Conversely, a large κ_t causes a fast decay and fits the yield curve better at short maturities. Moreover, κ_t determines the maximum location of the loading of $\beta_{3,t}$.

Figure 3 shows the loadings of $\beta_{1,t}, \beta_{2,t}$ and $\beta_{3,t}$ (κ_t taken a constant) as a function of time.[1] The limiting behavior of (20) can now readily be computed:

$$\lim_{\tau\to 0} R_t(\tau) = r_t = \beta_{1,t} + \beta_{2,t}. \tag{21}$$

$$\lim_{\tau\to\infty} R_t(\tau) = l_t = \beta_{1,t}. \tag{22}$$

Here r_t denotes the short rate as before and l_t denotes the long term rate.

The Nelson-Siegel model can generate several shapes of the yield curve including upwarding sloping, downward sloping and (inverse) hump shaped with no more than one maximum and one minimum. The different shapes are found by altering the parameters in (20). Extensions of the Nelson-Siegel model, most

[1] Diebold and Li (2006) fix κ at 0.0609. Section 4.6 elaborates on this fixed value.

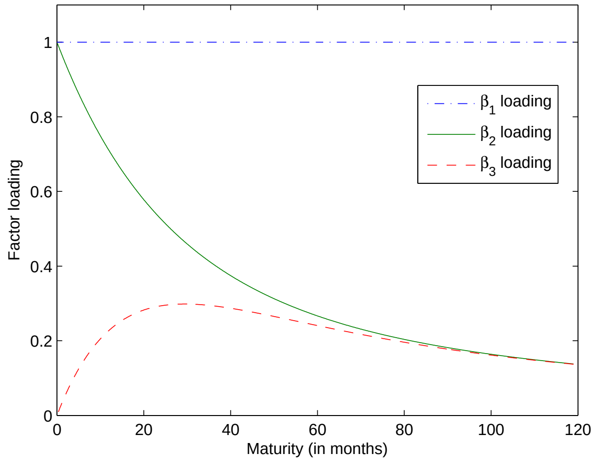

Figure 3: Factor loadings of the Nelson-Siegel model ($\kappa = 0.0609$).

notably the four-factor Svensson model (1994), are able to reconstruct term structures with multiple minima and/or maxima, which sometimes occur in reality. These extended models, as discussed in e.g., De Pooter (2007) are beyond the scope of this thesis, however.

At first sight, an important drawback of the Nelson-Siegel model seems that it lacks a solid mathematical basis. That is, the model is inconsistent with the no-arbitrage property as shown by Bjork and Christensen (1999). This means that the consistency between the dynamic evolution of interest rates and the actual shape of the term structure at a certain moment in time is not ensured. However, Coroneo *et al.* (2008) find that "the no-arbitrage parameters are not statistically different from those obtained from the Nelson-Siegel model, at a 95 percent confidence level." They hence conclude that the Nelson-Siegel model is in line with the no-arbitrage property.[2] Furthermore, they show that "...the Nelson-Siegel model performs as well as its no-arbitrage counterpart in an out-of-sample forecasting experiment" (Coroneo *et al.*, 2008).

Christensen, Diebold and Rudebusch (2007) introduce a slightly altered Nelson-Siegel model by adding a correction term, yielding a model that is consistent with the no-arbitrage property. The correction term only takes effect over an horizon of more than ten years and has little impact. Moreover, Christensen, Diebold and Rudebusch (2009) introduce a (generalized) Nelson-Siegel model that is closely related to the original model and that is in line with the the no-arbitrage restrictions.

Now that the three models considered are defined properly, the next step is to calibrate the models such that interest rates can be simulated. The upcoming section discusses the used data, estimation procedures and their results.

[2]It is noted that Coroneo *et al.* (2008) apply the model to the U.S. market.

4 Parameter Calibration and Estimation

4.1 Introduction

In order to compare the fit of term structure models in a proper way, it is imperative that one applies the same dataset to all models. Unfortunately, other papers do not compare the combination of models (Vasicek, CIR and Nelson-Siegel) that is used in this paper. Henceforth, I cannot use parameter estimations from other papers to compare the fit of the above models.

Since three dimensions - yield, time to maturity and time - should be taken into account, different estimation methods can be used. One could choose to do a cross-sectional estimation. That is, an estimation of the parameters while considering the different maturity times at a fixed moment in time. Another method is a time series estimation. In that case, maturity time is fixed and the parameters are estimated while considering the evolution of the interest rate over the different time periods in the dataset.[3]

For the Vasicek and CIR model, both methods have their merits and disadvantages. De Munnik and Schotman (1994) state that estimation of the mean reversion parameter will be difficult with time series data, because many observations spanning a large number of years are required. Clearly, in the cross-sectional procedure, only one day needs to be observed, as long as sufficient bonds are traded (i.e., bonds with different maturities). De Munnik and Schotman (1994) also indicate that a cross-sectional estimation might yield illogical estimates e.g., an estimate of the long-term average of above 20 percent.

The differences between time series and cross-sectional estimates should be small if the one-factor models of the term structure were true. Term structure models solely based on the short rate might not be very realistic in all scenarios, but still might do a good job in fitting the term structure.

For the Vasicek model, I discuss both cross-sectional and time series estimation and make a short comparison. I estimate the CIR model using time series. Since the parameters of the Nelson-Siegel model are time-dependent I estimate these via the cross-sectional data for each observed month in the dataset.

4.2 Data

The data I use are monthly yield curves for zero-coupon bonds, generated using pricing data of Canadian bonds and treasury bills. Note that zero-coupon bonds for maturities with a horizon longer than one year are not directly observable, meaning they need to be computed. Bolder, Johnson and Metzler (2004) provide a description of the methodoly used to derive the Canadian zero-coupon yields.

The data I consider run from January 1986 up until and including December 2009.[4] In total, there are 288 months in the dataset.

Until the end of December 1990, the Canadian government issued bonds with a maturity time not longer than 25 years. As of January 1991 they introduced bonds with a maturity time of 30 years. Henceforth, determining yields with maturity times of 25 years and longer for the period 1986-1991 is difficult if not

[3] Of course, other methods to estimate the parameters exist. See for instance Bolder (2001), in which the Kalman filter is introduced and applied to estimate the parameters of the Vasicek model.

[4] The data can be downloaded from the website of the Bank of Canada: http://www.bankofcanada.ca/en/rates/yield_curve.html.

impossible. If one tries to compute those yields anyway by e.g., extrapolation, cross-sectional estimations for that period will be unreliable. Therefore, I only focus on yields with maturity times not longer than 25 years.

Times to maturity have an interval of 0.25 years i.e., three months, starting at 0.25 years. Hence, I consider 100 different maturity times. Although the Canadian government (nor any other government) does not issue bonds with maturities like, for instance 15.25 years, an expression of the yield with such maturities is given by interpolation.

Figure 4 shows the monthly yield curves for zero-coupon bonds. Clearly visibile is a general decreasing trend of the interest rate for all maturity times. The central banking's policy to lower the short rates to nearly zero percent in an attempt to boost the economy after the credit crunch struck can be seen as well. Furthermore, an overall decrease of the yields occurred after central banks responded to the burst of the dot-com bubble. Table 9 presents corresponding descriptive statistics.[5] From this table one can deduce that, on average, the term structure is upward sloping. Moreover, the short rates are somewhat more variable than the long-maturity interest rates, whereas the long rates seem more persistent (i.e., higher autocorrelations) than the short rates.

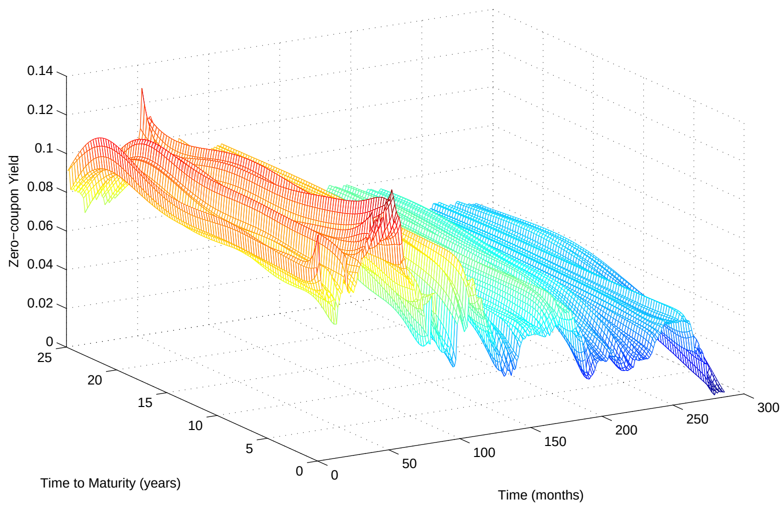

Figure 4: 3D-plot of the Canadian zero-coupon yields.

4.3 Calibration of the Vasicek Model

To estimate the parameters, a discrete time version of the model is needed. I now do not consider the discrete version as in (8), but introduce another representation of the discrete model, as stated by Brigo, Dalessandro, Neugebauer and Triki (2007):

$$r_{t_i} = c + br_{t_{i-1}} + \delta Z, \tag{23}$$

[5]See Appendix A for tables containing descriptive statistics.

where

$$c = \mu\big(1 - e^{-\theta\Delta t}\big),$$

$$b = e^{-\theta\Delta t},$$

and

$$\delta = \sigma\sqrt{\big(1 - e^{-2\theta\Delta t}\big)/2\theta}.$$

Here Z follows a standard normal distribution, the parameters θ, μ and σ denote the parameters of the continuous Vasicek model (5) and $\Delta t = t_i - t_{i-1}$.

The calibration process of the parameters is an ordinary least squares (OLS) regression, which provides maximum likelihood estimators for the parameters c, b and δ. Brigo *et al.* (2007) indicate that the following expressions for θ, μ and σ hold,

$$\theta = \frac{-\log(b)}{\Delta t}, \tag{24}$$

$$\mu = \frac{c}{1-b}, \tag{25}$$

$$\sigma = \delta/\sqrt{\big(b^2 - 1\big)\Delta t/2\log(b)}. \tag{26}$$

Next, estimators are needed such that we can actually calculate the values for θ, μ and σ. Therefore, I consider the estimators derived directly via maximum likelihood (Brigo and Mercurio (2006))[6],

$$\hat{b} = \frac{n\sum_{i=1}^{n} r_i r_{i-1} - \sum_{i=1}^{n} r_i \sum_{i=1}^{n} r_{i-1}}{n\sum_{i-1}^{n} r_{i-1}^2 - \big(\sum_{i=1}^{n} r_{i-1}\big)^2}, \tag{27}$$

$$\hat{\mu} = \frac{\sum_{i=1}^{n}\big[r_i - \hat{b}r_{i-1}\big]}{n\big(1-\hat{b}\big)}, \tag{28}$$

$$\hat{\delta^2} = \frac{1}{n}\sum_{i=1}^{n}\Big[r_i - \hat{b}r_{i-1} - \hat{\mu}\big(1-\hat{b}\big)\Big]^2. \tag{29}$$

With these estimators and equations (24) and (26), estimators for θ and σ are readily found.

One could argue which interest rate data to take to calibrate the model. Obviously, taking different yields imply different parameter estimates. I choose to calibrate the model on the two-year maturity yields, for two reasons. On the one hand the Vasicek model is a short rate model, so the time to maturity should not be too large. On the other hand, taking a short maturity time, say six months, might yield strange (e.g., negative) estimates because of the extreme low interest rates and high volatilities in the final years of the data due to the credit crunch. Since the data are on a monthly interval, the time step equals 1/12. Table 3 depicts the estimates. Implementation of the OLS regression is done in MATLAB.[7] Given the average three months rate (the interest rate with the lowest observed maturity time) to equal 0.055 and $\hat{\mu}$ (long term average) as in Table 3, it is in the line of expectation that the average fitted yield curve will be downward sloping. Figure 5 plots the average observed Canadian yields and the average fitted yield curve. Indeed, the Vasicek model plots a downward sloping average yield curve, whereas the observed Canadian yields are upward

Table 3: Estimates of the Vasicek model using time series.

Parameter	Estimation
$\hat{\theta}$	0.1167
$\hat{\mu}$	0.0285
$\hat{\sigma}$	0.0137

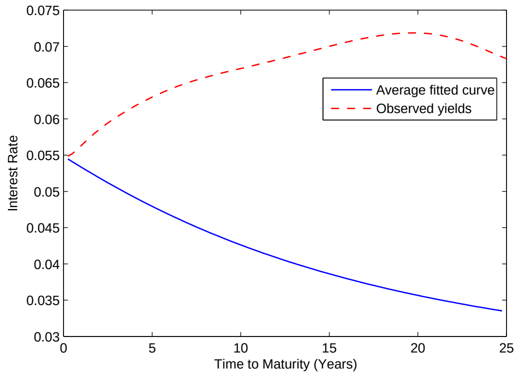

Figure 5: Fitted average yield curve with the Vasicek model using the time series estimates.

sloping on average. In the perfect case, the two graphs would match exactly. In this case, however, there is a large discrepancy between the two lines, largely due to the low interest rates as a result of the credit crunch. Calibrating the model on the yields spanning the period of 1986-2006 for instance, may produce a different average yield curve. Albeit a subset of the data may produce a better (i.e., upward sloping) average fitted curve, I decide not to omit any data.

4.4 Cross-sectional Estimation of the Vasicek Model

To calibrate the parameters via cross-sectional data, I again make use of OLS.[8] Note that I now consider the different yields to maturity on a fixed moment in time. To take all data into account, I calculate the estimates for each month (288 in total) and pool it. Table 4 shows a summary of the estimates.

Table 4: Summary of monthly pooled cross-sectional estimation.

	$\hat{r}_t$	$\hat{\theta}$	$\hat{\mu}$	$\hat{\sigma}$
Mean	0.0529	5.1554	0.2456	0.0301
Std. dev.	0.0307	12.692	0.2951	0.0463
Minimum	0 .000	0.0032	0.0000	0.0000
Maximum	0.1304	102.82	1.0000	0.2814
December 2009	0.0100	22.733	0.0003	0.0239

[6]To simplify notations, I write r_i instead of $r(t_i)$ in equations (27), (28) and (29).
[7]MATLAB version 7.7.0.471 (R2008b).
[8]See Holborow (2008).

Comparing the estimates in Table 3 with the (mean) estimates in Table 4, it turns out that none of the estimates are much alike, which can partly be explained by the fact I estimate the two-year yield when using time series and consider all yields (where the long rates are less volatile as we found in Table 9) in the cross-sectional estimation.

Indeed, as De Munnik and Schotman (1994) point out, some (mean) estimates (i.e., μ) seem to be nonsensical when using the cross-sectional estimation method.

Figure 6 plots fitted yield curves for three months, showing that the Vasicek model seems capable of adapting a term structure's observed shape in some cases.

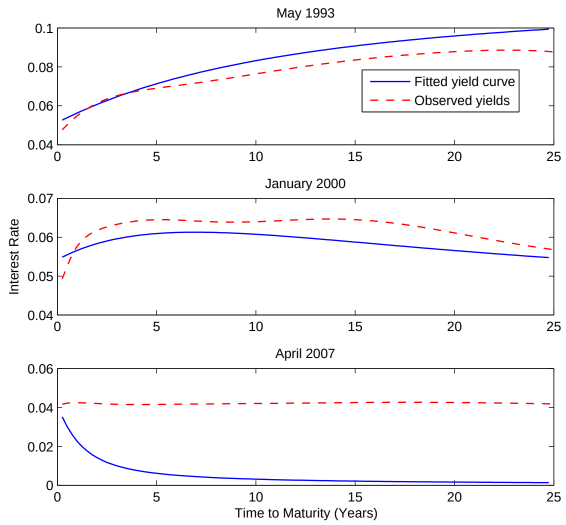

Figure 6: Fitted yield curve for specific months.

4.5 Calibration of the Cox-Ingersoll-Ross Model

As stated in the introduction of this section, I estimate the parameters of the CIR model (i.e., θ, μ and σ, see equation (12)) using time series. For the sake of consistency I use the same interest rates - the two-year rates - that I used in the time series estimation of the Vasicek model. I now consider the parameter vector $\psi \equiv (\theta, \mu, \sigma)$, which I estimate using maximum likelihood. Therefore, the density of the CIR process is required. Cox *et al.* (1985) describe the density of r_{t_i} at t_i given $r_{t_{i-1}}$ at t_{i-1} as

$$f_{CIR}(r_{t_i} \mid r_{t_{i-1}}; \psi, \Delta t) = ce^{-u-v}\left(\frac{v}{u}\right)^{\frac{q}{2}} I_q(2\sqrt{uv}), \tag{30}$$

where

$$c = \frac{2\theta}{\sigma^2\left(1 - e^{-\theta\Delta t}\right)},$$

$$u = cr_{t_{i-1}}e^{-\theta\Delta t},$$

$$v = cr_{t_i},$$

$$q = \frac{2\theta\mu}{\sigma^2} - 1,$$

and $Iq\big(2\sqrt{uv}\big)$ is a modified Bessel function of the first kind of order q. Note that for brevity I write u instead of $u_{t_{i-1}}$ and v instead of v_{t_i}.

Let N be the number of observations (e.g., the number of months the interest rate is observed; 288 in case of the Canadian yields that are considered in this thesis). Then, the likelihood function is given by:

$$L(\psi) = \prod_{i=0}^{N} f_{CIR}(r_{t_i} \mid r_{t_{i-1}}; \psi, \Delta t). \tag{31}$$

Since maximizing the log-likelihood function is often easier than maximizing the likelihood function itself we take natural logarithms on both sides in (31), resulting in:

$$\begin{aligned}
\ln\big(L(\psi)\big) &= \sum_{i=0}^{N} \ln\Big(f_{CIR}(r_{t_i} \mid r_{t_{i-1}}; \psi, \Delta t)\Big) \\
&= \sum_{i=0}^{N} \ln\Big(ce^{-u-v}\Big(\frac{v}{u}\Big)^{\frac{q}{2}} I_q(2\sqrt{uv})\Big) \\
&= \sum_{i=0}^{N} \ln(c) + \sum_{i=0}^{N}\Big(-u-v+\frac{q}{2}ln\big(\frac{v}{u}\big) + ln\big(I_q\big(2\sqrt{uv}\big)\big)\Big) \\
&= N \cdot \ln(c) + \sum_{i=0}^{N}\Big(-u-v+\frac{q}{2}ln\big(\frac{v}{u}\big) + \ln\big(I_q\big(2\sqrt{uv}\big)\big)\Big). \qquad (32)
\end{aligned}$$

Note that, since the logarithmic function is a monotonically increasing function, maximization of the likelihood function also maximizes the log-likelihood function. That is, the location of the maximum location does not change. The maximum likelihood estimates $\hat{\psi}$ of the parameter vector ψ can now be found by maximizing (32) over its parameter space. That is,

$$\hat{\psi} \equiv (\hat{\theta}, \hat{\mu}, \hat{\sigma}) = \arg\max_{\psi} \ln\big(L(\psi)\big). \tag{33}$$

For $\hat{\psi}$ to actually reach a global maximum location, it is imperative that the initial guesses of the parameter vector are sensical. Brigo *et al.* (2007) suggest the OLS estimate of the Vasicek regression as an initial estimate for θ (equation (24)),

$$\theta_0 = \frac{-\log(\hat{b})}{\Delta t}, \tag{34}$$

with $\hat{b}$ as defined in (27).

Since the long-term mean of the CIR process equals μ, the long-term variance equals $\mu\sigma^2/(2\theta)$ and because I have an initial estimate for θ, expressions for the initial guesses of μ and σ are now readily found to be

$$\mu_0 = \mathbb{E}(r_t) \tag{35}$$

and

$$\sigma_0 = \sqrt{2\theta_0 Var(r_t)/\mu_o}. \tag{36}$$

My MATLAB implementation of the maximum likelihood estimation of the CIR process makes use of the built-in MATLAB function `fminsearch` (I minimize the negative log-likelihood function).[9]

Kladivko (2007) indicates that direct implementation of the Bessel function $I_q(2\sqrt{uv})$ into MATLAB causes the program to crash. A failure occurs because the Bessel function $I_q(2\sqrt{uv})$ diverges to plus infinity on a high pace; something the MATLAB function `fminsearch` cannot handle. To cope with this problem, Kladivko (2007) suggests to use a scaled version of the Bessel function. This scaled Bessel function, which I denote by $I_q^{scaled}(2\sqrt{uv})$, is defined as $I_q(2\sqrt{uv})\exp(-2\sqrt{uv})$. The exponential here does the trick. To take the scaled Bessel function into account, the log-likelihood function in (32) needs to be adjusted to

$$\ln\big(L(\psi)\big) = N \cdot \ln(c) + \sum_{i=0}^{N}\Big(-u - v + \frac{q}{2}ln\big(\frac{v}{u}\big) + \ln\big(I_q^{scaled}\big(2\sqrt{uv}\big)\big) + 2\sqrt{uv}\Big).$$

Note that in the above equation, the term $2\sqrt{uv}$ appears because $\exp(-2\sqrt{uv})$ in the scaled Bessel function should be canceled out to keep the log-likelihood function the same.

Table 5 depicts results of the initial and MLE estimations. Figure 7 plots the

Table 5: Results of the MLE estimation of the CIR model.

	$\hat{\theta}$	$\hat{\mu}$	$\hat{\sigma}$
Initial	0.11668	0.058516	0.053132
MLE	0.10012	0.023526	0.056675

fitted average (downward sloping) yield curve and the observed average (upward sloping) yield curve.

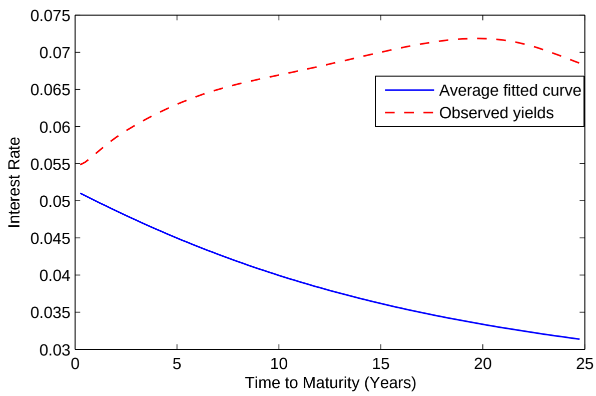

Figure 7: Fitted average yield curve with the cir model.

[9]The methodology of `fminsearch` is based on the Nelder-Mead method.

4.6 Estimation of the Nelson-Siegel Model

The Diebold and Li (2006) reformulation of the Nelson-Siegel model (equation 20) forms the basis for my estimation procedure. For the estimation of the parameters of this model I consider the functional form,

$$R_t(\tau) = H * \beta_t + \epsilon_t, \tag{37}$$

where $R_t(\tau)$ denotes the vector of interest rates at time t for n different times to maturity collected in the vector τ and $\epsilon_t \sim N(0, R)$. From (20) I derive H as

$$H = \begin{bmatrix} 1 & \frac{1-e^{-\kappa_t \tau_1}}{\kappa_t \tau_1} & \frac{1-e^{-\kappa_t \tau_1}}{\kappa_t \tau_1} - e^{-\kappa_t \tau_1} \\ 1 & \frac{1-e^{-\kappa_t \tau_2}}{\kappa_t \tau_2} & \frac{1-e^{-\kappa_t \tau_2}}{\kappa_t \tau_2} - e^{-\kappa_t \tau_2} \\ \vdots & \vdots & \vdots \\ 1 & \frac{1-e^{-\kappa_t \tau_n}}{\kappa_t \tau_n} & \frac{1-e^{-\kappa_t \tau_n}}{\kappa_t \tau_n} - e^{-\kappa_t \tau_n} \end{bmatrix}$$

and $\beta_t = \left(\beta_{1,t}, \beta_{2,t}, \beta_{3,t}\right)^T$.

From (19) it can be seen that the parameter κ_t might vary over time. Diebold and Li (2006) argue that this parameter might as well be taken as a constant with little degradation of fit. Fixing κ_t yields two major advantages. First of all, it greatly simplifies the estimation procedure of the other parameters because (37) reduces to a linear regression. Second, κ_t has no clear, intuitive economic interpretation. They furthermore argue to fix κ at 0.0609 with maturities measured in months. As section 3.4 explains, κ_t determines the maximum location of the loading of $\beta_{3,t}$ (i.e., the factor loading of the medium-term or curvature factor). Two or three years are typical medium-term maturities. Diebold and Li (2006) simply take the average, 30 months, and find that the the value for κ that maximizes the loading of the curvature factor at 30 months equals 0.0609.

Table 6 shows summary statistics of the OLS estimates.[10]

Table 6: Summary statistics of the Nelson-Siegel estimations ($\kappa = 0.0609$).

	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\epsilon}$
Mean	7.321	-1.681	-2.288	0.000
Std. dev.	2.100	1.981	2.280	0.000
Minimum	4.185	-5.376	-9.192	0.000
Maximum	11.140	3.611	4.327	0.000
$\hat{\rho}(1)$	0.986	0.886	0.763	0.085
$\hat{\rho}(12)$	0.966	0.445	0.028	-0.091
$\hat{\rho}(24)$	0.866	0.177	-0.103	0.051

To empirically test whether the factors $\beta_{1,t}, \beta_{2,t}$ and $\beta_{3,t}$ are legitimately called a level factor, slope factor and curvature factor, respectively, I do the following. From the observed yield data I construct a level, slope anc curvature and compare them with $\beta_{1,t}, \beta_{2,t}$ and $\beta_{3,t}$. That is, I define the 25-year yield as the level of the yield curve. The slope I define as the difference between the 25-year and three-month yields (a straight line through the smallest and

[10]The presented summary statistics for ϵ are averaged over the different maturity times. $\hat{\rho}(i)$ denotes the sample autocorrelation with a time-lag of i months.

largest yields). Finally, I compute the curvature as two times the two-year yield minus the sum of the 3-month and 25-year yields. Figure 8 shows a time series of the three factors found by OLS and the level, slope and curvature as I just defined. The estimated factors and the defined factors seem to follow the same pattern. To be more precise, correlations, ρ between the estimated factors and the level, slope, and curvature are $\rho(\hat{\beta}_{1,t}, l_t) = 0.943, \rho(\hat{\beta}_{2,t}, s_t) = -0.929$ and $\rho(\hat{\beta}_{3,t}, c_t) = 0.784$, where (l_t, s_t, c_t) are the level, slope and curvature, respectively. Hence, $\beta_{1,t}, \beta_{2,t}$ and $\beta_{3,t}$ may rightly be called level, slope and curvature factors, respectively. Correlations between the estimated factors are low: $\rho(\hat{\beta}_{1,t}, \hat{\beta}_{2,t}) = 0.168$, $\rho(\hat{\beta}_{1,t}, \hat{\beta}_{3,t}) = 0.033$ and $\rho(\hat{\beta}_{2,t}, \hat{\beta}_{3,t}) = 0.388$.

Diebold and Li (2006) perform a similar exercise based on zero-coupon yields generated using end-of-month price quotes for U.S. Treasuries, from January 1985 through December 2000. They find similar correlations: $\rho(\hat{\beta}_{1,t}, l_t) = 0.97, \rho(\hat{\beta}_{2,t}, s_t) = -0.99$ and $\rho(\hat{\beta}_{3,t}, c_t) = 0.99$. Exact values for the correlations between the factors are not provided, though they find these correlations to be low.

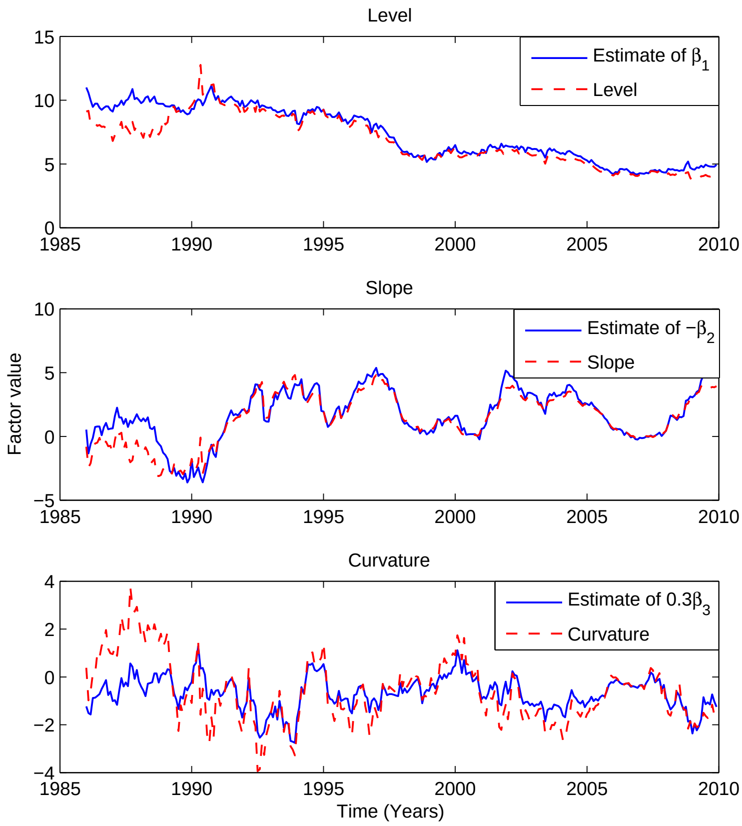

Figure 8: Time series of the Nelson-Siegel factors ($\kappa = 0.0609$).
Note: Rescaling of estimates based on Diebold and Li (2006).

Using the estimates from Table 6, Figure 9 shows the average fitted curve for the Nelson-Siegel model. It seems the curve fits pretty well. If I look more in-depth i.e., take a look at specific months, it turns out the Nelson-Siegel model fits multiple shapes of the term structure without too much difficulties (Figure 10).

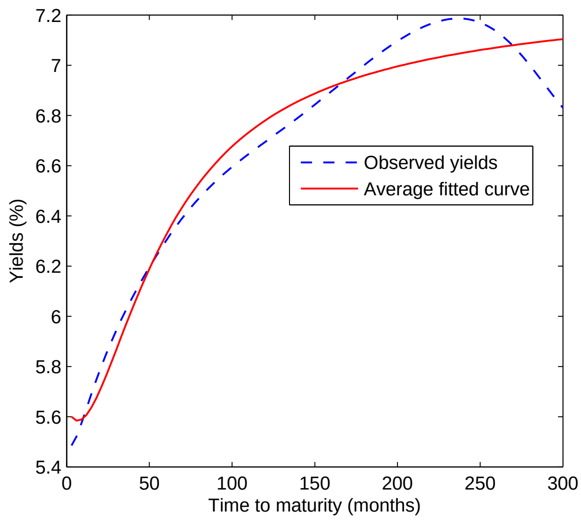

Figure 9: Fitted average yield curve ($\kappa = 0.0609$).

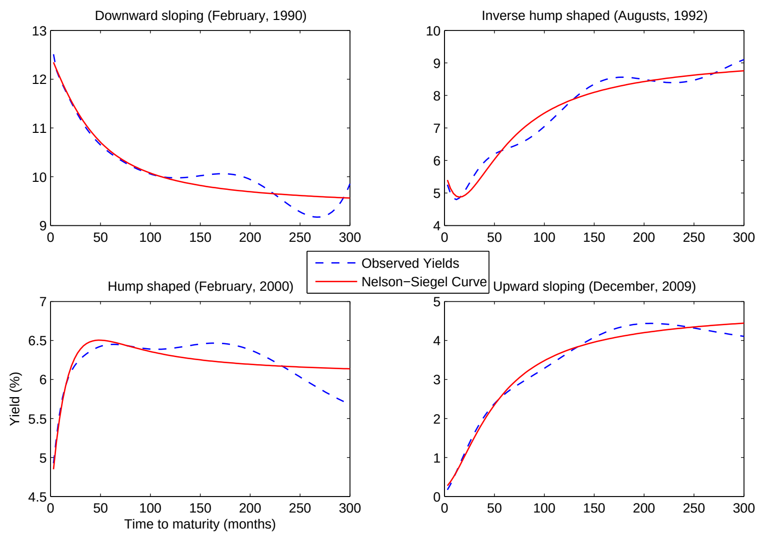

Figure 10: Fitted yield curve for specific months ($\kappa = 0.0609$).

From the table containing the descriptive statistics of the Canadian yields (Table 9) it becomes apparent that long rates are more persistent than short rates. That is, autocorrelations of long rates are higher than those of short rates. As (22) showed, $\beta_{1,t}$ solely determines the long run limiting behavior of the Nelson-Siegel model. One might therefore expect $\beta_{1,t}$ to be the most persistent factor. Descriptive statistics in Table 6 indeed indicate $\beta_{1,t}$ to be the most persistent factor. Furthermore, Figure 11 depicts autocorrelations for the three factors for different lags, confirming the observation. The figure also shows autocorrelations of the error terms, justifying the use of OLS.

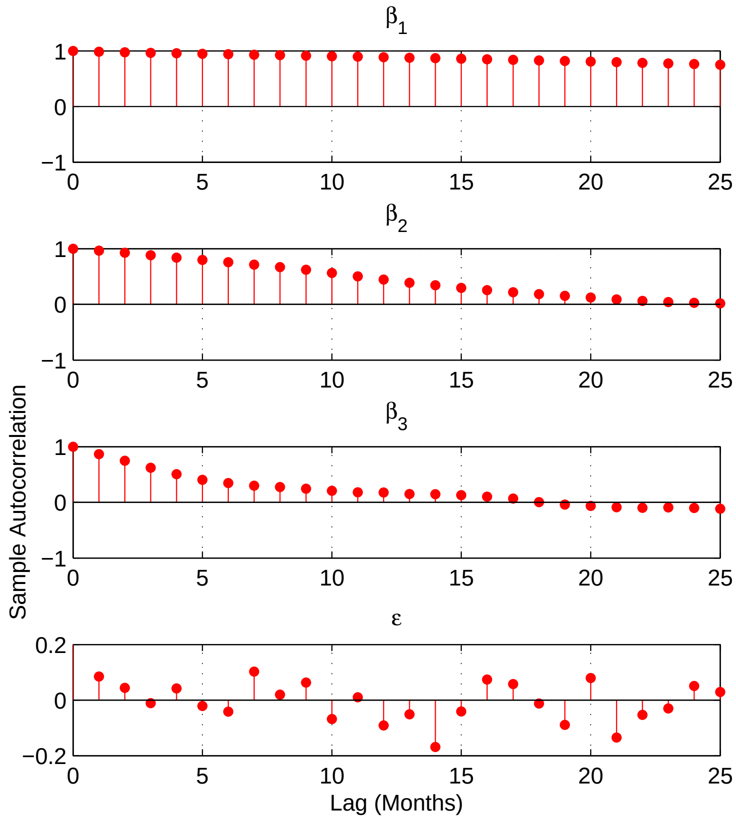

Figure 11: Autocorrelations of the three fitted factors and the error term of the Nelson-Siegel model with different lags measured in months.

All in all, looking at the residuals of the regression (Figure 12), it turns out that the fit is pretty good in most cases. Some months, however, especially those with multiple maxima and/or minima, are not fitted very well. Multiple maxima and/or minima occur in the term structure of months in the late '80s, which becomes apparent by the large residuals in these months.

Since estimates for all three models are now found, forecasting the term structure may begin. The upcoming section describes this process per model, starting with the Vasicek model and followed by the CIR and Nelson-Siegel models.

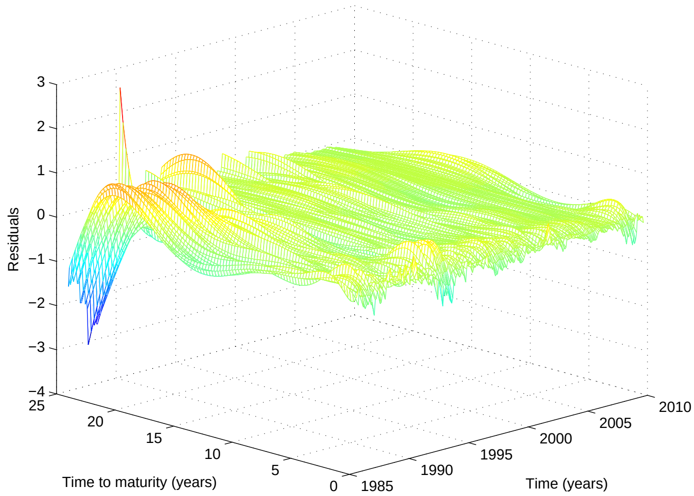

Figure 12: Residuals of the Nelson-Siegel OLS regression ($\kappa = 0.0609$).

5 Term Structure Forecasting

5.1 Introduction

This section aims to contribute answering the question whether simulated interest rates match the larger trends and statistics (i.e., stylized facts) of the observed interest rate data. Stylized facts, inspired by the Canadian yield data and Diebold and Li (2006), I will look into are:

1. The average yield curve is increasing and concave.
2. The yield curve assumes a variety of shapes through time e.g., upward and downward sloping, hump shaped, and inverted hump shaped.
3. Short rates are more volatile than long rates.
4. Long rates are more persistent than short rates.
5. Short rates are more skewed than long rates.
6. Short rates have higher kurtoses than long rates.

Since pension funds' liabilities often have long maturities, I use a forecasting period of forty years with a time step of one month. That is, I simulate 480 months, starting with the last month in the observed data, December 2009 and ending at December 2049.

In line with the Canadian yield data, I compute interest rates for maturity times varying from 0.25 years (three months) until 25 years with an interval of 0.25 years.

5.2 Forecasting with the Vasicek Model

The simulation of the short rate is based on the discretized model as described in (9). I run the simulation once using the time series estimates from Table 3 and once using the cross-sectional estimates from Table 4.

Starting with the time series estimates, the starting point of the short rate simulation process is the two-year yield at December 2009, being 0.013.[11] Of 10,000 simulations it turns out that approximately 17 percent of the simulated short rates are negative. Section 3.2 argues that observed nominal yields are never negative. Hence, if the simulation computes a negative yield, I replace it by the value of zero.

I find an average (over both the number of simulations and time) short rate of approximately 0.028, from which the full term structure can be deduced.

Figure 13 presents a plot of the average yield curve, which is downward sloping and hence not in line with the first stylized fact. Note that the no-arbitrage restrictions imposed on the Vasicek model have a large influence on the term structure's shape. Figure 5 in Section 4.3 shows a downward sloping average fitted yield curve. Due to the no-arbitrage fulfillments, the average fitted yield curve and the simulated average fitted yield curve show in some sense a similar path i.e., downward sloping.

Looking at specific dates, it turns out that inverse humped and upward sloping shapes occur as well. Note that by definition neither the Vasicek model nor the CIR model is able to replicate inverse hump shaped term structures.

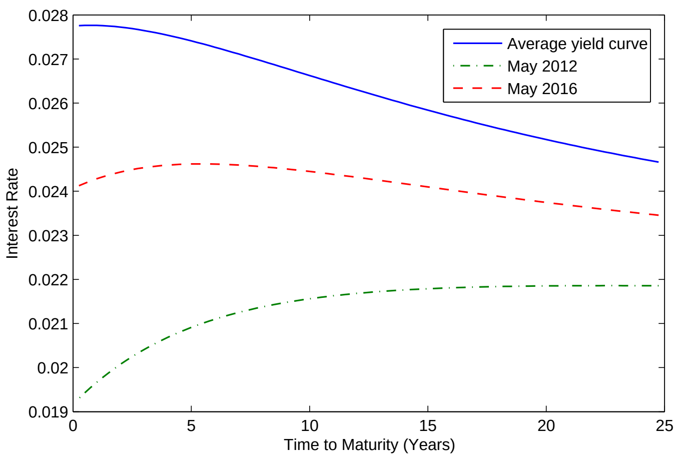

Figure 13: Simulated yield curves with the Vasicek model using the time series estimates.

By using (7) and the simulated short rate for each scenario, expressions arise for the yields at different maturity times. That is, I compute 10,000 matrices of size 480×100, containing yields for all maturity times and for all months.

Table 10 shows descriptive statistics of the simulated yields. Note that all yields with their respective maturity times are a function of the short rate. Hence, the variance of a yield with an arbitrary maturity time strongly depends

[11] Recall I calibrated the Vasicek model on the two-year yields.

on the variance of the simulated short rate. Therefore, values for the variances of the from (7) deduced interest rates are not very sensible. It does turn out, however, that the simulated short rates are more volatile than the simulated long rates, which is in line with the real interest rate data. Unfortunately, simulated volatilities are underestimated in the sense they are lower than the observed standard deviations.

For the same reason values for the autocorrelations, skewness coefficients and kurtoses are not very sensible either. Because a yield with an arbitrary maturity time is a function of the simulated short rate, kurtoses, skewness coefficients and autocorrelations will be much alike if not equal for all maturities. Table 10 indeed shows the moments and autocorrelations to be equal for all maturity times.

I rerun the simulation, but now using the cross-sectional estimates. To that end, I use the (nonsensical) OLS estimates of the last month in the observed dataset (December 2009), depicted in the last row of Table 4. For the sake of consistency, I take the same yield, 0.013, as the starting point of the simulation. After 10,000 simulations it turns out that a frightening 49 percent of the simulated yields are negative. I find an average (over both the number of simulations and time) short rate of approximately 0.006. Figure 14 shows the average yield curve. It may be clear that the shape of the term structure can at least be called odd, which is largely due to the unrealistic nature of $\hat{\theta}$. The simulation produces no shape other than downward sloping. Again, I find short rates to be more volatile than long rates and find constant values for the skewness coefficients, kurtoses and autocorrelations.

One may compare Table 11, containing descriptive statistics of the simulation based on the cross-sectional estimates, with Table 10.[12] Statistical properties (e.g., averages and autocorrelations) of the two simulations are not much alike. Stylized facts produced using the time series estimates seem somewhat more in line with the observed yield data than those produced with the cross-sectional estimates of December 2009.

Summarizing, the term structure simulation using the Vasicek model does not seem to be able to replicate many of the term structure's stylized facts. Moreover, it produces many negative yields, justifying ones doubt of the credibility of the simulation. One might conclude that the Vasicek model does not seem very suitable in a simulation based context. Since the CIR model is a simple extension to the Vasicek model, it is in the line of expectation that the CIR model will not be very suitable either.

5.3 Forecasting with the Cox-Ingersoll-Ross model

I simulate the short rate by making use of the discrete version of the short rate CIR process (Equation 18). Via (17) the full term structure can be deduced. I use the time series estimations for θ, μ and σ from Table 5 for the simulation. By construction, the continuous CIR model does not produce negative short rates. Its discrete counterpart, however, does. Therefore, I again replace negative yields by zeros.[13] Since the Vasicek and CIR models are calibrated on the same

[12] Note that the variation of the averages and the standard deviations in Table 11 primarily lie in the fourth decimal. See Figure 14.

[13] Approximately one half percent of the simulated yields are negative.

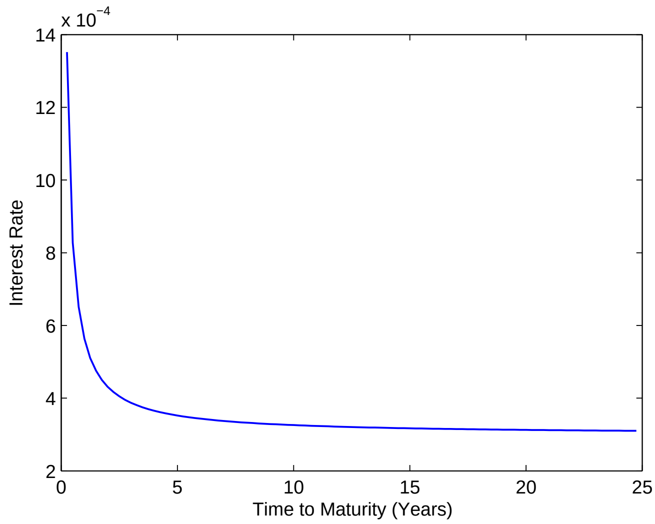

Figure 14: Simulated yield curve with the Vasicek model using the cross-sectional estimates of December 2009.

yields, I start the simulation of the short rate CIR process from the same point I started the simulation of the short rate Vasicek process: 0.013.

Of 10,000 simulations I find an average (over both the number of simulations and time) short rate of approximately 0.021. Figure 15 shows a plot of the hump shaped average yield curve, implying the simulated yield curve not to be in line with the first stylized fact. The figure also shows the term structure's other shapes produced by the simulation.

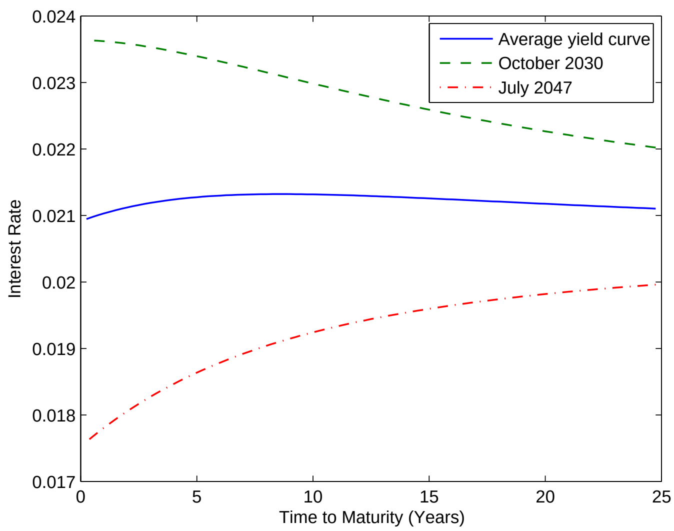

Figure 15: Simulated yield curves with the CIR model.

Summary statistics in Table 12 indicate that also the CIR model is not capable of replicating the interest rates' general trends.[14] Like the simulated Vasicek model, the CIR model generates the same skewness coefficients, the same kurtoses and the same autocorrelations for all maturity times. The short rates seem more volatile than the long rates, although the volatility is underestimated for all maturity times compared to the observed nominal yields.

One may conclude the CIR model to perform unsatisfactory in a simulation based context. Both the Vasicek and the CIR model seem not very useful in a pension fund's ALM study. As opposed to the Vasicek and CIR model, the Nelson-Siegel model does not fall not within the standard class of affine term structure models. Therefore, yields - and hence their stylized facts - simulated with the Nelson-Siegel model will likely be significantly different from the yields produced by the two other models.

5.4 Forecasting with the Nelson-Siegel Model

In contrast with the Vasicek and the CIR models, no general framework exists to simulate yields using the Nelson-Siegel model, leaving some room for ones own creativity. Since the three parameters $\beta_{1,t}, \beta_{2,t}$ and $\beta_{3,t}$ (assuming κ to be fixed) of the Nelson-Siegel model give a full description of the yield for a certain maturity, a model for those parameters should be build. In section 4.6 I found the estimated factors to be moderately correlated. Therefore, I choose to model the factors $\beta_{1,t}, \beta_{2,t}$ and $\beta_{3,t}$ as a first-order vector autoregressive (VAR(1)) model. That is,

$$\beta_{1,t} = c_1 + \pi_{11}\beta_{1,t-1} + \pi_{12}\beta_{2,t-1} + \pi_{13}\beta_{3,t-1} + \epsilon_{1,t}$$

$$\beta_{2,t} = c_2 + \pi_{21}\beta_{1,t-1} + \pi_{22}\beta_{2,t-1} + \pi_{23}\beta_{3,t-1} + \epsilon_{2,t}$$

$$\beta_{3,t} = c_3 + \pi_{31}\beta_{1,t-1} + \pi_{32}\beta_{2,t-1} + \pi_{33}\beta_{3,t-1} + \epsilon_{3,t}$$

where $\epsilon_{1,t}, \epsilon_{2,t}$ and $\epsilon_{3,t}$ are three white noise processes with covariance matrix Σ. This three-equations system may also be written as

$$\begin{pmatrix} \beta_{1,t} \\ \beta_{2,t} \\ \beta_{3,t} \end{pmatrix} = \begin{pmatrix} c_1 \\ c_1 \\ c_1 \end{pmatrix} + \begin{pmatrix} \pi_{11} & \pi_{12} & \pi_{13} \\ \pi_{21} & \pi_{22} & \pi_{23} \\ \pi_{31} & \pi_{32} & \pi_{33} \end{pmatrix} \begin{pmatrix} \beta_{1,t-1} \\ \beta_{2,t-1} \\ \beta_{3,t-1} \end{pmatrix} + \begin{pmatrix} \epsilon_{1,t} \\ \epsilon_{2,t} \\ \epsilon_{3,t} \end{pmatrix}. \tag{38}$$

I calibrate the parameters of (38) on the time series of $\beta_{1,t}, \beta_{2,t}$ and $\beta_{3,t}$ that I obtained from the OLS regression on (37) using maximum likelihood.[15] Therefore, I make use of the MATLAB function `vgxvarx`, part of MATLAB's Econometrics toolbox. I indicate that the program should stop after 1000 iterations with a convergence tolerance of $\epsilon^{\frac{3}{4}}$.[16] Table 7 shows the maximum likelihood estimates for the parameters of (38), including standard errors and t-statistics. Several t-statistics suggest the insignificance of some parameters. Although I am

[14]Figure 15 shows that the average yield curve is not flat, while the averages depicted in Table 12 may suggest otherwise. It is noted that this table considers only three decimals, while the variation in the average yield curve lies in the fourth decimal.

[15]Figure 8 shows time series of $\beta_{1,t}, \beta_{2,t}$ and $\beta_{3,t}$.

[16]MATLAB's default value for ϵ equals $2.22044604925031 \times 10^{-16}$.

Table 7: Maximum likelihood estimates, standard errors and t-statistics for the paramaters of (38).

	Value	Std. Error	t-statistic
$\hat{c}_1$	0.660	0.152	4.329
$\hat{c}_2$	-0.081	0.106	-0.767
$\hat{c}_3$	-0.413	0.261	-1.581
$\hat{\pi}_{11}$	0.924	0.019	49.350
$\hat{\pi}_{12}$	0.034	0.022	1.562
$\hat{\pi}_{13}$	0.014	0.019	0.758
$\hat{\pi}_{21}$	0.013	0.013	1.001
$\hat{\pi}_{22}$	0.948	0.015	61.940
$\hat{\pi}_{23}$	0.052	0.013	3.932
$\hat{\pi}_{31}$	0.017	0.032	0.522
$\hat{\pi}_{32}$	0.042	0.038	1.117
$\hat{\pi}_{33}$	0.848	0.032	26.303

aware of this, I choose not to omit any variables. (39) shows the corresponding (co)variance matrix $\hat{\Sigma}$.

$$\hat{\Sigma} = \begin{pmatrix} 0.445 & & \\ -0.070 & 0.218 & \\ -0.124 & -0.049 & 1.317 \end{pmatrix}. \tag{39}$$

For the actual simulation I make use of the MATLAB function `vgxsim`. The input for `vgxsim` are the parameters in Table 7 and the expressions for the (co)variances in (39). Due to computational limitations I simulate the time series of size 480 months only 5,000 times. Table 8 displays summary statistics of the three simulated factors and the corresponding error terms. This table may be compared with the actual estimated factors from Table 6.

Table 8: Summary statistics of the simulated Nelson-Siegel factors and its' error terms ($\kappa = 0.0609$). Averaged over both the number of simulations and different months.

	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\epsilon}_1$	$\hat{\epsilon}_2$	$\hat{\epsilon}_3$
Mean	7.076	-2.172	-2.507	0.000	0.000	0.000
Std. dev.	1.969	1.914	2.221	0.670	0.466	1.147
Minimum	0.482	-6.879	-8.792	-2.026	-1.410	-3.472
Maximum	12.108	2.493	3.770	2.025	1.414	3.469
$\hat{\rho}(1)$	0.907	0.947	0.841	0.012	-0.006	0.004
$\hat{\rho}(12)$	0.298	0.491	0.159	0.015	-0.013	0.009
$\hat{\rho}(24)$	0.071	0.262	0.089	-0.003	-0.026	-0.019

Averaged over both the number of simulations and the different months, the yield curve indeed is upward sloping (Figure 16). Comparing this graph with the fitted average yield curve from Section 4.6, one notices the graphs to be much alike. This may be attributed to the fact that the standard Nelson-Siegel is not statistically different from its arbitrage-free counterpart at the 95 percent confidence level.[17] Looking at specific months to identify different shapes of the term structure, it turns out that upward sloping and inverse hump shaped are the only shapes produced by the simulation. That is, if one looks at specific months averaged over the number of simulations. If one would not average over the number of simulations, the term structure assumes all four possible shapes.

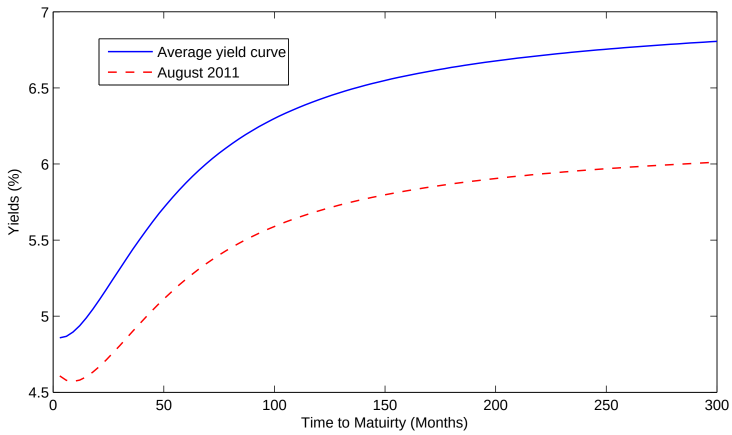

Figure 16: Simulated yield curves using the Nelson-Siegel model.

To check for the other stylized facts, I compute yields for all maturity times and for all different months 5,000 times. I thus compute 5,000 matrices - one for each scenario - with dimensions 480×100, containing the yields on every month for all maturity times. Then, I compute the statistical properties that are of interest (e.g., variances and autocorrelations) per scenario. After having done this, I average over the number of scenarios, yielding the required statistical properties.

Table 13 shows summary statistics of the simulation. Here, it can be seen that the simulated short rates indeed are more volatile than the long rates. It also seems that the simulation catches the downward sloping trend of the skewness of the yields. Moreover, kurtoses of the simulated short rates are lower than those of the simulated long rates, as can also be found in the observed nominal yields.

The numeric values of the averages and volatilites show much resemblance with the actual Canadian yield data. Numeric values for the skewness coefficients and kurtoses, however, deviate from the observed yields. Where the Canadian data shows skewness coefficients between zero and one, the simulation shows values somewhere between 0.01 and -0.2. Furthermore, the Canadian yields have kurtoses ranging from approximately -0.3 and -1, while the simulation produces kurtoses ranging from roughly 2.5 up to 3.2.

One may also deduce from Table 13 that the simulation fails when it comes to the autocorrelations of the yields. Where in the observed nominal data long rates are more persistent (higher autocorrelations) than short rates, the simulation shows the opposite. Figure 17 shows sample autocorrelations of the factors

[17] See Coroneo *et al.* (2008).

which also support this observation.[18] The figure also shows autocorrelations of the error terms, justifying the use of OLS.

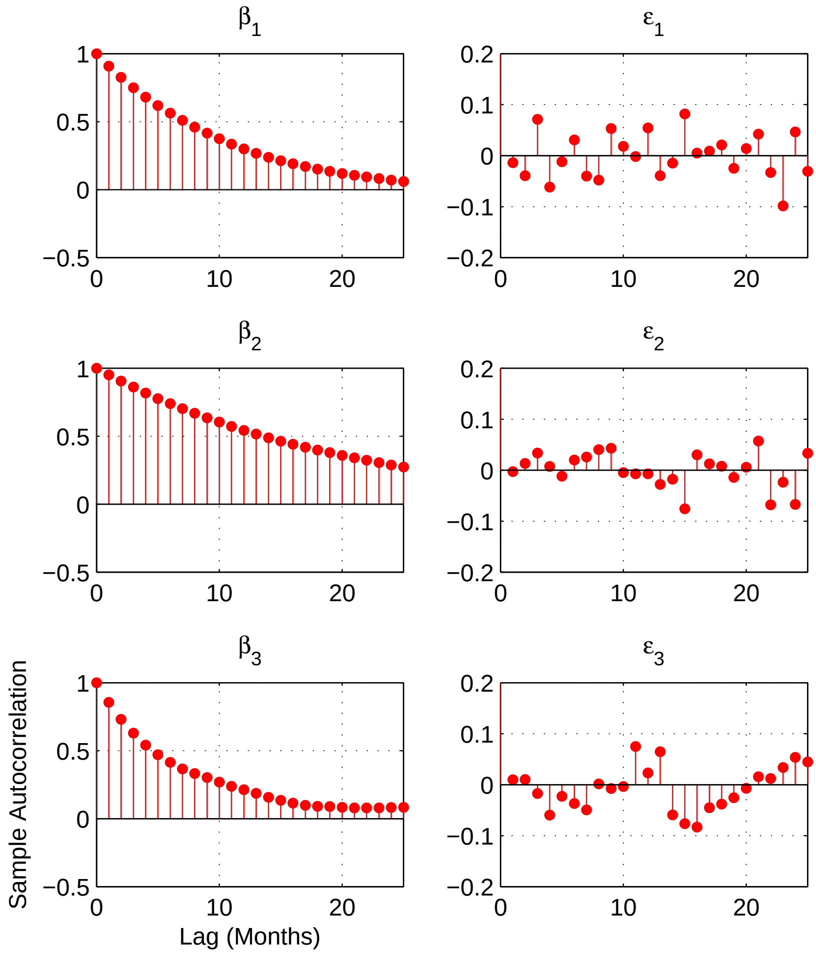

Figure 17: Sample autocorrelations of the simulated factors and error terms as a function of lag (in months).

One may be interested how the term structure reacts over time to exogenous impulses (shocks). In the Nelson-Siegel model this boils down to how the three factors $\beta_{1,t}, \beta_{2,t}$ and $\beta_{3,t}$ respond to such impulses. Therefore, I construct impulse response functions for the three factors. That is, the deterministic response of my VAR(1) model to an innovations process that has the value of one standard deviation in one component at the initial time, and zeros in all other components and times. MATLAB's function `vgxproc` simulates the evolution of

[18]Recall that for the Nelson-Siegel model it holds that $\lim_{\tau \to \infty} R_t(\tau) = \beta_{1,t}$, implying $\beta_{1,t}$ to be the most persistent factor in case long rates are more persistent than short rates. Figure 17 shows the opposite: $\beta 2, t$ is more persistent than $\beta_{1,t}$.

a time series model from a given innovations process, making it suitable to construct impulse response functions. Using `vgxproc`, I simulate the VAR(1) model two more times. Once with all innovations equal to zero and once with the first innovation term equal to the square root of $\hat{\Sigma}(i,i)$ (39) for factor $\beta_{i,t}$ and the rest zeros. Computing and plotting the relative differences between these two simulations returns the desired impulse response functions.

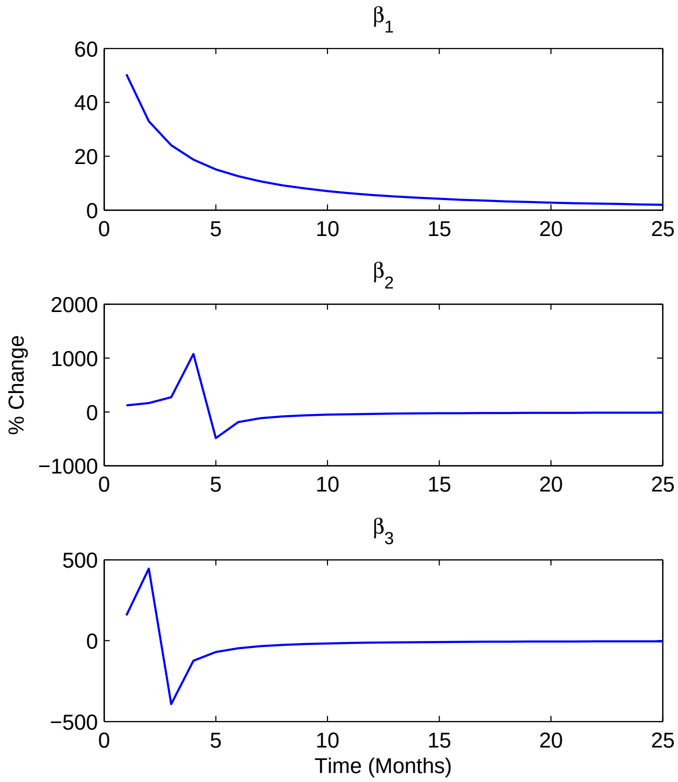

Figure 18: Impulse response function of the Nelson-Siegel factors measured in months.

Figure 18, presenting the impulse response functions, shows that $\beta_{1,t}$ (level factor) requires a longer period to recover from a shock of size one standard deviation than $\beta_{2,t}$ (slope factor) and $\beta_{3,t}$ (curvature factor). To be more precise, it seems it takes approximately six months for $\beta_{2,t}$ and $\beta_{3,t}$ to recover, whereas it takes over twenty months for $\beta_{1,t}$ to fully recover. The relative changes of $\beta_{2,t}$ and $\beta_{3,t}$ are much higher than the relative change of $\beta_{1,t}$. This can be explained by the fact that values for $\beta_{2,t}$ and $\beta_{3,t}$ are smaller in magnitude than $\beta_{1,t}$, implying a fixed shock to have more effect on $\beta_{2,t}$ and $\beta_{3,t}$ than it does on $\beta_{2,t}$.

Summarizing, it turns out that the Nelson-Siegel model performs well in a simulation based context. Therefore, it is legitimate to say that the Nelson-Siegel model outperforms the Vasicek and CIR models when it comes to term structure simulation.

6 Conclusion

6.1 Summary

In this paper I focus on whether term structure models of interest rates used for pricing derivative instruments are suited to use in a simulation based context like an asset liability management (ALM) study. I consider the Vasicek, Cox-Ingersoll-Ross (CIR) and Nelson-Siegel model. After elaborating on the three models, I estimate the models' parameters. To that end, I consider Canadian zero-coupon yields from 1986 up until and including 2009 with maturity times ranging from three months to 25 years with an interval of three months.

Using OLS I estimate the Vasicek model once using time series of the two-year interest rates and once using cross-sectional data i.e., the different times to maturity. The time series estimates seem somewhat more sensible than the (mean) estimates computed using the cross-sectional data.

I apply maximum likelihood to calibrate the CIR model on time series of the two-year yields.

I calibrate the Nelson-Siegel model by fixing the parameter that indicates the rate of the exponential decay (κ). As a result, the estimation procedure of the Nelson-Siegel model reduces to a simple OLS regression. I find the Nelson-Siegel model to fit the term structure well in many cases. Moreover, I find that, in line with Diebold and Li (2006), the three factors $\beta_{1,t}, \beta_{1,t}$ and $\beta_{3,t}$ are rightly called a level, slope and curvature factor respectively. Furthermore, the estimated factors are moderately correlated.

Comparing stylized facts of the Canadian data with stylized facts produced by the simulation of the Vasicek model based on the time series estimates, it turns out that there are a lot of mismatches. To name a few: the average term structure of the observed nominal yields is upward sloping whereas the simulation produces a downward sloping yield curve. Moreover, skewness coefficients, kurtoses and autocorrelations are equal for all maturity times according to the simulation, whereas the differences in the observed data are evident. The simulation replicates a higher volatility for the short rates than it does for the long rates, which is in line with the actual data. However, the volatility is underestimated. Stylized facts produced by the simulation based on the cross-sectional estimates of December 2009 (last observed month) show also mismatches with the observed data. The simulation based on the time series seems to do a better job in the sense that values for averages and volatilities are somewhat more realistic.

The simulation of the CIR model produces similar results to the Vasicek simulation based on the time series estimates. Skewness coefficients, kurtoses and autocorrelations are equal for all maturity times. The average curve seems almost flat at 2.1 percent, with a little hump in the medium run. Simulated short rates are more volatile than long rates but again underestimated.

I model the three factors (fixing κ) of the Nelson-Siegel model with a first-order vector autoregressive (VAR(1)) model. This simulation replicates most of the stylized facts. Although there is a mismatch between the actual values, the simulation captures the general trends of the skewness coefficients and kurtoses. Moreover, the average term structure is upward sloping and values for the averages and volatilities seem realistic. Where the level factor, $\beta_{1,t}$ is the most persisent factor in the observed data, the slope factor, $\beta_{2,t}$ turns out to be the most persistent in the simulation.

All in all, one may conclude that using time series or cross-sectional data, neither the Vasicek model nor the CIR model seems suited to use in a simulation based context. The Nelson-Siegel model, however, replicates many of the observed stylized facts, making it more suitable for e.g., an ALM study than the Vasicek or CIR model.

6.2 Recommendations for Future Research

Since the parameter values have a large influence on the term structure's shape, it would be interesting to investigate the effect of the estimation procedure on the simulation. This could be translated into calibrating the model onto a subset of the observed data such that e.g., low yields due to the credit crunch are omitted.

This could also be translated into using a totally different estimation procedure: the time series estimation procedure neglects all yields except for one array of yields with a fixed maturity time. The cross-sectional estimation procedure neglects even more interest rates, as it considers the yields at a fixed moment in time. Therefore, one might look into an estimation procedure that considers all data e.g., panel data estimation.

Although the Vasicek and CIR models are in line with the KISS principle compared to its' extensions, one might question a model that deduces the full term structure from solely the short rate. Therefore, an interesting study would be to look into the simulating power of multiple-factor models like the two-factor Vasicek model or a time-dependent model like the Hull-White model.

As the VAR(1) model is unable to identify $\beta_{1,t}$ as the most persisent factor in the Nelson-Siegel model, it might be interesting to model the factors with a univarite model like an AR(1) or ARMA(1,1) model. Moreover, investigating whether dependence between the factors should be taken into account would make an interesting study as well.

References

[1] Bjork, T. and Christensen, B.J. (1999). Interest Rate Dynamics and Consistent Forward Rate Curves. *Mathematical Finance*, 9 : pp 323 - 348.

[2] Black, F., Derman, E. and Toy, E. (1990). A one-factor model of interest rates and its application to treasury bond options. *Financial Analyst Journal*, 46 : pp 33 - 39.

[3] Bolder, D.J. (2001). Affine Term-Structure Models: Theory and Implementation. *Bank of Canada. Working Paper 2001-15.*

[4] Bolder, D.J., Johnson, G. and Metzler, A. (2004). An Empirical Analysis of the Canadian Term Structure of Zero-Coupon Interest Rates. *Bank of Canada. Working Paper 2004-48.*

[5] Brigo, D., Dalessandro, A., Neugebauer, M. and Triki, F. (2007). A Stochastic Processes Toolkit for Risk Management. Available at: `http://ssrn.com/abstract=1109160` (Accessed April 12, 2010).

[6] Brigo, D. and Mercurio, F. (2006). Interest Rate Models: Theory and Practice - With Smile, Inflation and Credit. *New York, NY: Springer Verlag.*

[7] Christensen, J.H.E., Diebold, F.X. and Rudebusch, G.D. (2007). The Affine Arbitrage-Free Class of Nelson-Siegel Term Structure Models. *FRB of San Francisco Working Paper No. 2007-20.*

[8] Christensen, J.H.E., Diebold, F.X. and Rudebusch, G.D. (2009). An arbitrage-free generalized Nelson-Siegel term structure model. *Econometrics Journal*, 12 : pp C33 - C64.

[9] Cox, J.C., Ingersoll, J.E and Ross, S.A. (1985). A Theory of the Term Structure of Interest Rates. *Econometrica*, 3 : pp 385 - 407.

[10] Coroneo, L., Nyholm, K. and Vidova-Koleva, R. (2008). How Arbitrage-Free is the Nelson-Siegel Model? *European Central Bank. Working Paper No. 874.*

[11] De Munnik, J.F.J. and Schotman, P.C. (1994). Cross-sectional versus time series estimation of term structure models: empirical results for the Dutch bond market. *Journal of Banking & Finance*, 18 : pp 997 - 1025.

[12] De Pooter, M. (2007). Examining the Nelson-Siegel Class of Term Structure Models. *Tinbergen Institute Discussion paper. No. 2007-043/4.*

[13] Diebold, F. X., Ji, L. and Li, C. (2006). A Three-Factor Yield Curve Model: Non-Affine Structure, Systematic Risk Sources, and Generalized Duration, in L. Klein (ed.). *Macroeconomics, Finance and Econometrics: Essays in Memory of Albert Ando, Cheltenham, U.K.* : pp 240 - 274.

[14] Diebold, F.X. and Li, C. (2006). Forecasting the Term Structure of Government Bond Yields. *Journal of Econometrics*, 130 : pp 337 - 364.

[15] Duffee, G. (2002). Term premia and interest rate forecasts in affine models. *Journal of Finance*, 57 : pp 405 - 443.

[16] Heath, D., Jarrow, R. and Morton, A. (1987). Bond Princing and the Term Structure of Interest Rates: A New Methodology for Contingent Claims Valuation. *Cornell University, Ithaca, United States.*

[17] Holborow, D. (2008). The Effectiveness of Structural Bond Pricing Models with Maximum Likelihood Estimation of Asset Dynamics Parameters. *University of Queensland, Brisbane, Australia.*

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

[19] Kladivko, K. (2007). Maximum Likelihood Estimation of the Cox-Ingersoll-Ross Process: the MATLAB Implementation. *University of Economics, Prague, Czech Republic.*

[20] Mankiw, N.G. and Miron, J.A. (1986). The Changing Behavior of the Term Structure of Interest Rates. *NBER Working Papers 1669, National Bureau of Economic Research, Inc.*

[21] Nelson, C. R. and Siegel, A. F. (1987). Parsimonious Modeling Of Yield Curves. *Journal of Business*, 60 : pp 473 - 489.

[22] Pensioenfonds Metaal & Techniek (2010). Dekkingsgraden. Available at: `http://www.bpmt.nl/over-pmt/algemeen/nieuws/257,275/dekkingsgraden` (Accessed May 6, 2010).

[23] Schumacher, J.M. (2009). Syllabus Financial Models. *Tilburg University, Tilburg, the Netherlands.*

[24] Svensson, L. E. O. (1994). Estimating and Interpreting Forward Interest Rates: Sweden 1992-1994, *NBER Working Paper Series. No. 4871.*

[25] Vasicek, O.A. (1977). An Equilibrium Characterisation of the Term Structure. *Journal of Financial Economics*, 5 : pp 177 - 188.

[26] Zellner, A. (1992). Statistics Science and Public Policy. *Journal of the American Statistical Association* 87, pp: 1 - 6.

Appendix A: Descriptive statistics of (simulated) yields

Note: $\hat{\rho}(i)$ denotes the sample autocorrelation with a time-lag of i months.

Table 9: Descriptive Statistics of the Canadian zero-coupon yields.

	Maturity					
	3 month	1 year	5 year	10 year	20 year	25 year
Mean	0.055	0.056	0.063	0.067	0.072	0.068
Stand.Dev.	0.031	0.029	0.024	0.022	0.022	0.020
Minimum	0.001	0.005	0.017	0.029	0.040	0.037
Maximum	0.133	0.128	0.117	0.111	0.116	0.128
Skewness	0.705	0.500	0.284	0.204	0.162	0.242
Kurtosis	-0.235	-0.579	-1.065	-1.339	-1.422	-1.003
$\hat{\rho}(1)$	0.983	0.982	0.981	0.983	0.984	0.979
$\hat{\rho}(12)$	0.750	0.767	0.819	0.851	0.872	0.862
$\hat{\rho}(24)$	0.541	0.572	0.676	0.730	0.747	0.725

Table 10: Descriptive Statistics of the simulated yields using the Vasicek model based on the time series estimates.

		$\hat{\theta} =$ 0.1167	$\hat{\mu} =$ 0.0285	$\hat{\sigma} =$ 0.0137		
	Maturity					
	3 month	1 year	5 year	10 year	20 year	25 year
Mean	0.028	0.027	0.027	0.026	0.025	0.025
Stand.Dev.	0.018	0.017	0.014	0.011	0.007	0.006
Skewness	0.458	0.458	0.458	0.458	0.458	0.458
Kurtosis	2.637	2.637	2.637	2.637	2.637	2.637
$\hat{\rho}(1)$	0.975	0.975	0.975	0.975	0.975	0.975
$\hat{\rho}(12)$	0.738	0.738	0.738	0.738	0.738	0.738
$\hat{\rho}(24)$	0.540	0.540	0.540	0.540	0.540	0.540

Table 11: Descriptive Statistics of the simulated yields using the Vasicek model based on the cross-sectional estimates of December 2009.

		$\hat{\theta} =$ 22.733	$\hat{\mu} =$ 0.0003	$\hat{\sigma} =$ 0.0239		
			Maturity			
	3 month	1 year	5 year	10 year	20 year	25 year
Mean	0.001	0.001	0.000	0.000	0.000	0.000
Stand.Dev.	0.002	0.001	0.000	0.000	0.000	0.000
Skewness	1.555	1.555	1.555	1.555	1.555	1.555
Kurtosis	4.992	4.992	4.992	4.992	4.992	4.992
$\hat{\rho}(1)$	-0.469	-0.469	-0.469	-0.469	-0.469	-0.469
$\hat{\rho}(12)$	0.192	0.192	0.192	0.192	0.192	0.192
$\hat{\rho}(24)$	0.042	0.042	0.042	0.042	0.042	0.042

Table 12: Descriptive Statistics of the simulated yields using the CIR model.

		$\hat{\theta} =$ 0.1001	$\hat{\mu} =$ 0.0235	$\hat{\sigma} =$ 0.0567		
			Maturity			
	3 month	1 year	5 year	10 year	20 year	25 year
Mean	0.021	0.021	0.021	0.021	0.021	0.021
Stand.Dev.	0.012	0.012	0.010	0.008	0.005	0.004
Skewness	0.715	0.715	0.715	0.715	0.715	0.715
Kurtosis	3.064	3.064	3.064	3.064	3.064	3.064
$\hat{\rho}(1)$	0.975	0.975	0.975	0.975	0.975	0.975
$\hat{\rho}(12)$	0.736	0.736	0.736	0.736	0.736	0.736
$\hat{\rho}(24)$	0.532	0.532	0.532	0.532	0.532	0.532

Table 13: Descriptive Statistics of simulated yields using the Nelson-Siegel model ($\kappa = 0.0609$).

			Maturity			
	3 month	1 year	5 year	10 year	20 year	25 year
Mean	0.049	0.050	0.059	0.064	0.068	0.068
Stand.Dev.	0.029	0.028	0.023	0.021	0.020	0.020
Skewness	0.015	0.010	-0.058	-0.125	-0.173	-0.183
Kurtosis	2.662	2.669	2.815	2.984	3.118	3.147
$\hat{\rho}(1)$	0.963	0.961	0.944	0.936	0.930	0.929
$\hat{\rho}(12)$	0.635	0.624	0.531	0.478	0.4442	0.434
$\hat{\rho}(24)$	0.385	0.385	0.300	0.248	0.213	0.206