

# **A 3-FACTOR MODEL FOR THE YIELD-CURVE DYNAMICS - THE CASE OF STOCHASTIC SPOT-RATE, MARKET PRICE OF RISK AND VOLATILITY**

**Claus Madsen**

**First version 11 November 1998**

**This Version 18 November 1998**

**Abstract:** This paper proposes a new 3-factor model for the dynamic in the yield-curve which belongs to the Affine class of term structure models. Using a yield-factor approach combined with a maximum likelihood estimation technique we conclude the following using Danish bond data over the period 2 January 1990 - 30 June 1998:

- We find all the parameters to be significant and when comparing the model generated interest rate series against the actual (observed) interest rate series we find for all maturity dates that we cannot reject the hypothesis that both samples have been randomly selected from the same distribution.
- We also find that there is a high degree of correlation between the predicted interest rates and the observed interest rates - given the state-variables

We also analyse the implied state-variables - that is the state-variables that arise when converting the model into a yield-factor model - and find that there are three dominant factors that govern the dynamic in the term structure of interest rates which can be recognized as, a Short-factor (sometimes called a Steepness-factor), a Level-factor and a Curvature-factor. This is in line with what is generally reported in the literature, see among others Litterman and Scheinkmann (1991).

As is generally the case we also find evidence that the Level-factor is close to being non-stationary.

**Keywords:** Multi-factor models, stochastic volatility, numerical solution of ODEs, spot-rate models, Affine yield-curve models, maximum likelihood

**e-mail:** [cam@fineanalytics.com](mailto:cam@fineanalytics.com)

## A 3-Factor Model for the Yield-Curve Dynamics - The case of stochastic spot-rate, market price of risk and volatility<sup>1</sup>

### 1. Introduction

In the literature a great deal of attention has been drawn to the specification and estimation of multi-factor term structure models, see among others Balduzzi, Das and Foresi (1997), Duan and Simonato (1995), Dai and Singleton (1997) and Chen and Scott (1995). Furthermore most of the focus has been on either multi-factor CIR-models or multi-factor Gaussian representations.

Due to the fact that there is evidence that points toward a high degree of volatility persistence in the interest rate process, researches have recently been inspired to include stochastic volatility in multi-factor yield-curve models, see especially Vertzal (1997) and Andersen and Lund (1997).

In this paper we suggest a new 3-factor model that is a member of the linear Affine class of term structure models. We specify the model to belong to the Affine class because it makes the solution of the PDE (the system of ODEs) relatively straightforward even when no closed form solution is available. Our work can therefore be seen to be related to the work of Chen (1995) - who also specifies a 3-factor model with stochastic volatility belonging to the Affine class.

As our model belongs to the Affine class of yield-curve models - it is a spot-rate model - more precisely the spot-rate is a linear combination of the state-variables. Our 3-state-variables have been selected as:

- The spot-rate
- The market price of risk
- The volatility

Our motive in choosing this specification is discussed in section 2. As the model belongs to the Affine class we present in section 3 some well known results (and make some generalizations) about Affine yield-curve models, both for the Gaussian case and the Exponential case.

As we will utilize the yield-factor approach in our estimation procedure we show in section 4 how to imply the state-variables from an Affine yield-curve model given a subjective

---

<sup>1</sup> I thank Christian Dahl for usefull comments, especially in connection with the estimation of the 3-factor model.

## A 3-Factor Model for the Yield-Curve Dynamics

---

specification of yield-factors.

In section 5 our model is specified and derived and furthermore its implication is discussed.

Section 6 is dedicated to the estimation of the model using information across the whole yield-curve. For that purpose we use Danish bond data over the period 2 January 1990 - 30 June 1998.

Because we are using the yield-factor approach when estimating the model we can derive the exact maximum likelihood function. Compared to the Kalman filter algorithm the yield-factor approach is just as efficient seen from a computational perspective and the estimator is more efficient, as will be explained in more detail in section 6. The reason for this is that for a general specification of Affine term structure models the QML-estimator is not consistent - it will actually only be consistent if the transition equation is treated as if it is Gaussian.

When estimating the model we find all the parameters to be significant and when comparing the model-generated interest rate series against the actual (observed) interest rate series we find for all maturity dates that we cannot reject the hypothesis that both samples have been randomly selected from the same distribution.

We also analyse the implied state-variables - that is the state-variables that arise when converting the model into a yield-factor model - and find that there are three dominant factors that govern the dynamic in the term structure of interest rates which can be recognized as, a Short-factor (sometimes called a Steepness-factor), a Level-factor and a Curvature-factor. This is in line with what is generally reported in the literature, see among others Litterman and Scheinkmann (1991).

In section 8 we investigate the prediction of yield-curve movements given knowledge about the state-variables (the yield-factors). We find a high degree of correlation between the actual yield-curve movements and the model-generated yield-curve movements. However, the model is found to be inadequate forecasting the level of interest rates across the whole yield-curve - which is due to the fact that the Level-factor turned out to be close to non-stationary.

In the paper we also find evidence that there is possibly a non-stationary component in interest data. Firstly, this kind of result is generally what is reported in the literature for multi-factor models, see among others Zheng (1993) and Chen and Scott (1995). Secondly, as mentioned by Madsen (1998) if we utilize a PCA method and there is a non-stationary component in the interest rate series we will have one (1) dominant factor which will turn out as a Level-factor.

### 2. Modelling the dynamic in the yield-curve - a survey

In general we have that the process for the yield-curve can be captured by specifying the process for the spot-rate - this is even true in a multi-factor environment.

## - The case of stochastic spot-rate, market price of risk and volatility

---

This central role played by the spot-rate in all yield-curve modelling is nowhere more evident than in the pricing of fixed income securities. The reason for this is that we under general conditions find that the price at time  $t$  of a zero-coupon bond that matures at time  $T$  can be expressed using the Feynmac-Kac representation:

$$P(t;T) = \tilde{E}_t \left[ e^{-\int_t^T g(X_s) ds} \right] \quad (1)$$

where the expectation is taken under the risk-neutral probability measure<sup>2</sup>. This equation clearly illustrates that modelling the dynamic in the yield-curve can be established by specifying a process for the spot-rate.

In the literature on analysing the dynamic in the yield-curve two different approaches have been utilized<sup>3</sup>, even though they conceptionally are equal as both methods rely on the connection between the state-variables and the spot-rate, ie:  $r_t = g(X_t)$ . There is however one important point where the line of research diverges in the two approaches:

- Methods that “just” focus on the dynamic of the spot-rate - that is in the estimation of the parameters in the system of SDEs only information about the spot-rate is utilized. This I will denote the pure spot-rate model approach
- Methods that use information about yields across the yield-curve (or actual bond-prices). This I will denote the full spot-rate model approach

Let me briefly discuss the line of research that has been adopted in the two approaches in order. The order will however not be strict as cross relationships, when appropriate, will be addressed.

The main reference in the area of research for pure spot-rate models is the paper of Chan, Karolyi, Longstaff and Sanders (1992). They consider the following general spot-rate model:

$$dr_t = [a_1 + a_2 r_t] dt + b_2 r_t^\gamma dW \quad (2)$$

where this specification of the spot-rate process nests most of the parametric specifications from the literature. For example we get the Vasicek model for  $\gamma = 0$  and the CIR model for  $\gamma = 0.5$ .

In order to estimate the unrestricted model they perform a discrete time-approximation in line

<sup>2</sup> Where  $g(X_t)$  is a function that relates the state-variable/s  $X_t$  to the spot-rate.

<sup>3</sup> We disregard here approaches which only rely on panel-data and for that reason violate the time-homogenous assumption in the SDEs. These approaches are generally referred to as implied-volatility-approaches, see among others Madsen (1994), Brown and Dybvig (1986), Brown and Schaefer (1994) and Barone, Cuoco and Zautzik (1991).

## A 3-Factor Model for the Yield-Curve Dynamics

---

with Dietrich-Campbell and Schwartz (1986) and Babbell (1988) as follows:

$$\begin{aligned} r_{t+1} - r_t &= a_1 + a_2 r_t + \varepsilon_{t+1} \\ &\text{for} \\ E[\varepsilon_{t+1}] &= 0 \\ &\text{and} \\ E[\varepsilon_{t+1}^2] &= b_2^2 r_t^{2\gamma} \end{aligned} \tag{3}$$

and estimate the parameters using the GMM<sup>4</sup> approach<sup>5</sup>. The general conclusion in their analysis is that models that assume  $\gamma < 1$  give the worst description of data<sup>6</sup>. In this setting it rules out the Vasicek model and the CIR model. This result indicates the presence of a form of conditional heteroskedasticity, as it implies that the volatility depends on the level of the spot-rate - the so-called level-effect. From the empirical work in the literature<sup>7</sup> - for spot-rate models of the kind in equation 3 - it is appropriate to assume that  $\gamma$  is approximately equal to 1 (one), that is a log-normal specification seems reasonable.

Even though the one-factor model in formula 2 is fairly general it does not in a satisfactory manner<sup>8</sup> describe the kind of dynamic which can be observed in the market for the whole yield-curve.

This has inspired research in two directions:

- Non-parametric methods
- Introducing more state-variables - that is a move to multi-factor models

I will shortly discuss these two approaches in order.

---

<sup>4</sup> The GMM-technique is fairly popular. The reason for this is firstly that in GMM it is not required that the variables are normally distributed, it is "only" required that the variables are stationary and the relevant expectations operators exist. Secondly, the GMM-estimator is consistent (with respect to the discrete time approximation) even though the residuals contain heteroskedasticity and serial correlation, see Greene (1993, section 13).

<sup>5</sup> However as mentioned by Honore (1998) then by estimation the parameters using a discretization as in formula 3 the estimates of  $a_2$  and  $b_2$  will be biased. The reason for this is that in the discretization in formula 3 it is assumed that the transition density is normal - which is only the case for  $\gamma = 0$ . Actually we have that in general it is only when the time-step approaches 0 (zero) that the estimates are not biased. Nowman (1997) proposes a different discretization scheme for the Gaussian case which ramifies that problem - it is however worth mentioning that Nowmans discretization is known as the exact discretization, see Gouriou, Monfort and Renault (1993 section 7.3).

<sup>6</sup> This result is also obtained by Honore (1998).

<sup>7</sup> See among others Marsh and Rosenfeld (1983), Fisher and Zechner (1984) and Uhrig and Walter (1994).

<sup>8</sup> For elaboration see later in this section.

## - The case of stochastic spot-rate, market price of risk and volatility

---

A new methodology has recently been applied for one-factor diffusion models - namely the non-parametric approach, see Ait-Sahalia (1995) and Jiang (1997). The overall results in these analyses points toward a highly non-linear specification for the drift and diffusion term.

Ait-Sahalia develops an estimation technique based on the matching of the unconditional kernel density (see Scott (1992)) and applies this method to the class of one-factor models nested by the specification in formula 2 - and finds that none of the usual parametric models (including the general specification in formula 2) can describe the dynamic in the short-rate adequately.

Instead he proposes a general non-linear one-factor model of the following form<sup>9</sup>:

$$dr_t = \left[ \frac{a_0}{r_t} + a_1 + a_2 r_t + a_3 r_t^2 \right] dt + \sqrt{b_0 + b_1 r_t + b_2 r_t^2} dW \quad (4)$$

However this formulation is not supported by Honore (1998)<sup>10</sup> in his analysis of the Ait-Sahalia non-linear one-factor model - more precisely he finds that it is sufficient with a square-root diffusion term and that  $a_0 = a_2 = 0$ . The explanation for this finding is probably that Honore include rates along the whole yield-curve in the estimation whereas Ait-Sahalia “only” focus on the evolution in the short-rate.

This however does not indicate that the specification by Ait-Sahalia does not give a better description of the data-generating process for the short-rate - instead it indicates that the presence of non-linearities in interest-rate data is not as pronounced for longer term maturities as for short-term maturities.

On the other hand from the literature of GARCH<sup>11</sup> models there is evidence that points toward a high degree of volatility persistence in the interest rate process<sup>12</sup>. This has inspired among others Brenner, Harjes and Kroner (1993), Vetzal (1997) and Brailsford and Maheswaran (1997).

The results from these analyses indicate firstly, that the importance of interest rate volatility

---

<sup>9</sup> This is the only specification which is not rejected by his test.

<sup>10</sup> Honore's method is not what I have termed a pure spot-rate method as it also uses panel-data. I have however incorporated it here as it seems more fitting.

<sup>11</sup> GARCH is short for (generalized) autoregressive conditionally heteroskedastic. The GARCH was initially introduced by Bollerslev (1986) as an extension of the ARCH-model from Engle (1982). For a good introduction see Giannopoulos (1995).

<sup>12</sup> Adding a process for the volatility was however originally suggested by Brennan and Schwartz (1982).

## A 3-Factor Model for the Yield-Curve Dynamics

---

on levels has been exaggerated<sup>13</sup>, and secondly that the appropriate GARCH specification for the short-rate seems to belong to the class of asymmetric GARCH-models<sup>14</sup>.

In spirit of the findings from the GARCH literature Andersen and Lund (1997) suggest the following two-factor stochastic volatility model:

$$\begin{aligned} dr_t &= \kappa(\theta - r_t)dt + \sigma r^\gamma dW_1 \\ d\ln\sigma^2 &= \kappa_3(\theta_3 - \ln\sigma^2)dt + \sigma_3 dW_3 \end{aligned} \quad (5)$$

Their results indicate that it is more appropriate to assume that  $\gamma = 0.5$  - which is in line with the results from Brenner, Harjes and Kroner<sup>15</sup>.

However, conflicting results have been reported in the literature, for example Hørdahl (1997) estimates the Andersen and Lund model and get results for  $\gamma$  close to 1<sup>16</sup>. Vetzal (1997) for a similar model gets results that indicate that  $\gamma$  is close to 1<sup>17</sup>.

Let us now consider the three factor models<sup>18</sup> of Chen (1995) and Andersen and Lund (1997) - which can be understood as a central tendency model with stochastic volatility in the process for the spot-rate - then we find that the spot-rate is governed by the following system of SDEs:

$$\begin{aligned} dr_t &= \kappa(\theta - r_t)dt + \sigma r^\gamma dW_1 \\ d\theta &= \kappa_2(\theta_2 - \theta)dt + \sigma_2\sqrt{\theta}dW_2 \\ d\ln\sigma^2 &= \kappa_3(\theta_3 - \ln\sigma^2)dt + \sigma_3 dW_3 \end{aligned} \quad (6)$$

The specification in equation 6 is identical to Andersen and Lund's model if we assume that

<sup>13</sup> That is, the findings in for example Chan, Karolyi, Longstaff and Sanders that  $\gamma \geq 1$  gives the best description is not correct if the spot-rate process is adjusted for GARCH effects.

<sup>14</sup> That include for example EGARCH, AGARCH and Treshold-GARCH-models, see Giannopoulos (1995).

<sup>15</sup> This is (as mentioned above) in line with the result from Honore (1998) using however a completely different approach.

<sup>16</sup> Hørdahl however uses a different estimation method than the one utilized by Andersen and Lund. Andersen and Lund use the EMM method from Gallant and Tauchen (1996) where as Hørdahl uses the method from Ruiz (1994).

<sup>17</sup> Vetzal mentions (footnote 22) that the reason for the different conclusions between his results and the results from Andersen and Lund about  $\gamma$  could be that Andersen and Lund use data for the 3-month T-bill whereas he uses the 30-days T-bill. However Hørdahls result is obtained using the 3-month T-bill - which indicates that the explanation from Vetzal probably does not fully explain the difference.

<sup>18</sup> Which (as of now) is the most advanced pure spot-rate models that has been proposed in the literature.

## - The case of stochastic spot-rate, market price of risk and volatility

there are no correlations among the Wiener-processes. Chen's model is established if we fix  $\gamma = 0.5$ , assume that the process for the variance (not the log-variance) follows a square-root process, and if all the three Wiener-processes are correlated.

The motivation for expanding with a process for the unconditional mean in the spot-rate is that data shows that there are extended periods of strong drift - and furthermore the spot-rate series often seems to drift downwards, even when the rate is below the mean<sup>19</sup>. Incorporating a process for the unconditional mean in the spot-rate process is also supported in the work of Balduzzi, Das and Foresi (1997) .

As in Andersen and Lund's two-factor volatility model Andersen and Lund get results that indicate that the process for the short-rate seems to be driven by a CIR-type process.

From the results in the literature on pure spot-rate models we will then draw the following overall conclusion:

- If we restrict ourselves to one-factor models, in order to capture the complex behaviour of the short-rate process we need highly non-linear specification for both the drift and diffusion term
- Expanding to multi-factor models generally indicates that a reasonable assumption for the spot-rate process is a CIR-type specification. This result is however not unique - so I might be more appropriate to assume that  $\frac{1}{2} \leq \gamma \leq 1$  .

In general however we are interested in the dynamic in the whole yield-curve and not only the spot-rate. What I mean by that is that getting a good model for the short-rate validated using the pure spot-rate model approach does not necessarily indicate that we have succeeded in specifying a good model for the dynamic across the whole yield-curve.

Andersen and Lund shows that the factor-loadings that can be derived from their model in a sense mimic the results obtained using PCA, see for example Litterman and Scheinkman (1991). That is they can also identify the presence of three dominant factors - which can be interpreted as level, steepness and curvature. This however also seems to be the observation for other three-factor models, for example Chen and Scott (1995) and Zheng (1993) for very different model specifications<sup>20</sup> - so this observation might not directly (only) be a function of how the different SDEs in the three-factor model is specified but instead (probably more) related to the number of factors.

---

<sup>19</sup> This is especially true for the U.S short-term interest rate. Most of the analysis in the literature has also been concentrated on U.S data - probably because of the long time-series (daily and weekly data from 1954-) and which furthermore is available for free from several websites.

<sup>20</sup> Chen and Scott use a three-factor CIR-model, whereas Zheng focuses on a three-factor version of the SAINTS model from Constantinides (1992).

## A 3-Factor Model for the Yield-Curve Dynamics

---

Furthermore with the result from Honore (1998) in mind it seems as if the dynamic for longer-rates does not exhibit as many non-linearities as is the case for the short-rate. This also supports the issue addressed above - namely, that a good model for the short-rate does not by nature indicate a good model for the dynamic across the yield-curve.

Models that focus on modelling the yield-curve using what we have denoted the full spot-rate method are all of the Affine class<sup>21</sup>. This line of research has mainly been inspired by the fact that the correlation between changes in rates usually falls when the time-to-maturity distance rises - which of course makes one-factor models<sup>22</sup> inadequate for describing the co-movements along the whole yield-curve. One other thing that make linear one-factor models appropriate is the fact that twist in the yield-curve is not allowed in these models. The reason for this is that the risk-premium in these models can only be of either negative or positive signs<sup>23</sup>.

This has led to the suggestion of a number of two-factor models. The second factor chosen is variously the long rate (Brennan and Schwartz (1979,1980,1982)), the spread between the long-and the short-rate (Schwartz and Schaefer (1984)), inflation (Buraschi (1994)), the unconditional mean of the short-rate<sup>24</sup> (Balduzzi, Das and Foresi (1997) and the double-decay model of Beaglehole and Tenney (1991)), the volatility of the short-rate (Longstaff and Schwartz (1990) and Vasicek and Fong (1991)), and duration (Schwartz and Schaefer (1987)).

Analysis of multi-factor models that use information across the yield-curve include for example Chen and Scott (1993,1995), Singh (1995), Duan and Simonato (1995)<sup>25</sup>, Dai and Singleton (1997), Geyer and Pichler (1995) and Duffie and Singleton (1996).

Even though the authors uses different data<sup>26</sup> it is of interest to see what kind of results are

---

<sup>21</sup> For further information see section 3. It should however be mentioned here that the SAINTS model does not fall under this category. Another thing worth emphasizing is that by a minor change of the Andersen and Lund specification this model can also be put into the Affine family.

<sup>22</sup> At least one-factor models of the Affine linear class.

<sup>23</sup> Twist is on the other hand (in general) allowed for non-linear spot-rate models, for example the Ait-Sahalia model. However, it is worth emphasizing that the first model to address the evidence of non-linearities in the process for the spot-rate was Longstaff (1989) - later corrected by Beaglehole and Tenney (1992). By corrected I mean that that the solution given in Longstaff does not solve the problem he intends to solve, for further information see Beaglehole and Tenney (1992).

<sup>24</sup> This is sometimes referred to as the central tendency.

<sup>25</sup> However, as Duan and Simonato (1995) only consider maturities of less than 1-year it is questionable that additional information is obtained compared with only using the short-rate.

<sup>26</sup> Most of the analysis (as mentioned before) is however performed on U.S data - where the results obtained might be more or less relevant for our purpose as we only consider Danish data. But the information in the

## - The case of stochastic spot-rate, market price of risk and volatility

---

generally found when estimating multi-factor models of the Affine class. The main result in these investigation is:

- There is evidence of non-stationarity. This is because one of the mean-reversion parameters is not “significantly” different from 0 (zero)<sup>27</sup>. Furthermore the estimates for the mean-reversion parameter - when it is close to 0 (zero) - is in most cases lower than the standard-error

As pointed out in Madsen (1998) if there is a non-stationary component in bond yield-data then the result obtained from PCA will give rise to one factor which an eigenvalue significantly higher than the others and furthermore this factor will appear as a level-factor. In this case we even have that the second factor will look close to a slope-factor. If this is the case then it might explain the existence of three factors that can be interpreted as level, slope and curvature in all analysis no matter what data is used and which country data they are taken from. On U.S. data see Litterman and Scheinkman (1991) and Garbade (1986), on English data see Caverhill and Strickland (1992) and on Danish data see Madsen (1995) and Dahl (1989).

There is however a fundamental problem in estimating the effect of each component in the drift - which could be another explanation for the result obtained in the literature that one of the processes appear to be non-stationary. For the purpose of illustration let us consider the risk-adjusted drift for the CIR-model which is given by:  $[\kappa\theta - \kappa r_t - \lambda r_t]$  - from this it is

obvious that we cannot separate the impact of  $\kappa$  and  $\lambda$  on the spot-rate. In general we find that the drift is overidentified - because we have 3-parameters but can only observe 2-parameters - namely  $\kappa\theta$  and  $\kappa + \lambda$ . If we now assume that the drift in the spot-rate can be expressed as:

$[\kappa\theta - \kappa r_t - \Gamma]$  <sup>28</sup> - then we have the same problem - 3 parameters to estimate but we can only

observe 2, namely  $\kappa$  and  $\kappa\theta - \Gamma$ . Using this observation in the estimation of the parameters in multi-factor yield-curve models has (as far as I know) only been considered by Singh (1995) - for the 3-factor CIR-model<sup>29</sup>.

From fixed-income theory we always work under the risk-neutral probability measure - as we here focus on the pricing of contingent claims. The true probability measure which is related to the risk-neutral probability measure through the market price of risk has in that connection

---

line of research still contains a lot of useful information - if nothing else it might indicate formulations which could be amendable for Danish data.

<sup>27</sup> When the mean-reversion parameter is close to 0 (zero) the autoregression coefficient is close to one - which is equivalent to having a unit root close to the unit circle.

<sup>28</sup> This can be recognized to be equal to the risk-neutral drift in the Vasicek model for  $\Gamma = \lambda\sigma$ .

<sup>29</sup> From his estimated parameters there is (in one of the cases) evidence of non-stationarity. His approach for estimation is however completely different than related work, as he estimates the 3-factor CIR-model using the 3-factors obtained by PCA.

## A 3-Factor Model for the Yield-Curve Dynamics

---

no interest - it is of importance in two cases:

- When we wish to forecast the price of the underlying security
- When we use historical data to determine the parameters in the process for the spot-rate (the yield-curve)<sup>30</sup>

With respect to the first point I am not aware of any work that focuses on the recovery of subjective probabilities in connection with interest rate data<sup>31</sup>.

With respect to the second point we have limited evidence of the relationship between the market price of risk and the difference between the drifts along the yield-curve. The only general result we have is that it is negative - which is logical if we wish to restrict ourselves to the positive interest rate environment<sup>32</sup>.

The question is here - is it appropriate that the market price of risk is negative? In general, we think that the risk premium is negative if we have a rising yield-curve. On the other hand, if we are in a falling yield-curve environment, we expect the risk premium to be positive. So we cannot in general expect the market price of risk to be negative and even more important we cannot expect it to be constant. In this connection it seems - if we wish to assume that one of the parameters that determines the drift in the spot-rate is stochastic - it is more appropriate to assume it is the market price of risk parameter and not the unconditional mean<sup>33</sup>.

From this short discussion about determining the drift I think the following is of importance:

- It is important that we do not try to estimate the overidentified system for the drift specification under the original probability measure
- It is important that the mean-reversion parameter can be uniquely observed. For that reason I prefer drift specifications where we can observe the following two identities  $\kappa$  and  $\kappa\theta - \Gamma$ . This kind of drift separation principle is different from what has usually been used in the literature - the principle

---

<sup>30</sup> It is however worth mentioning that if we estimate the parameters in the SDE system for the spot-rate using the pure spot-rate model approach we cannot get information about the market price of risk from the data. The reason for this is that if the data used as a proxy for the short-rate is appropriate then we by construction find that the risk-neutral drift is equal to the drift under the original probability measure - they are namely both equal to the spot-rate (short-rate) - at least approximately. The market price of risk can only be estimated if we use information across the whole yield-curve in the estimation.

<sup>31</sup> In connection with options on stocks, there is some preliminary work by Jackwerth (1996) and Jackwerth and Rubenstein (1996). For further information I refer to these two papers.

<sup>32</sup> Of course a negative market price of risk does not necessarily mean positive interest rates - but on the other hand, if we assume a square-root process then a positive market price of risk can drive interest to become negative - which of course in this case is not valid.

<sup>33</sup> This conclusion also supports that it is not possible to determine the market price of risk by "only" using data for the short-rate - we need data across the whole yield-curve.

## - The case of stochastic spot-rate, market price of risk and volatility

---

proposed here has however the nice property that it is possible to observe the mean-reversion parameter. The reason why this is important is because mean-reversion is also very important in connection with for example option pricing as the volatility is a function of the level of mean-reversion.

- The market price of risk should not be considered as being constant

The new three-factor model for the dynamic in the yield-curve I will propose taking into account the discussion in this section will be of the Affine class (though even then fairly untractable), so for that reason I will first in section 3 recall some results from the literature on Affine yield-curve models.

With respect to the exponential class of Affine yield-curve models I will however derive a fairly general analytical expression - more precisely I will derive the analytical expression for the Duffie and Kan (1996) model (the extended stacked multi-dimensional CIR-model).

### 3. State Space Models

A general multi factor model is based on the specification of a multi dimensional Markov process  $X$ . If we assume that  $X = \{X_1, X_2, \dots, X_m\}$  follows a multi dimensional diffusion process<sup>34</sup> defined as a unique strong solution of the SDE:

$$dX_t = \mu(X_t)dt + S(X_t)dW_t \quad (7)$$

where  $W$  is an  $m$ -dimensional Wiener process and the coefficients  $\mu$  and  $S$  take values in  $\mathfrak{R}^m$  and  $\mathfrak{R}^m \otimes \mathfrak{R}^m$  respectively. The variables  $X_1, X_2, \dots, X_m$  are termed state-variables.

Notice that in the definition of formula 7 we have assumed that the number of factors coincides with the number of state-variables, in fact, generally we have that the number of state-variables may be greater than the number of factors<sup>35</sup>.

In a multi-factor setting we have that the spot-rate is given by  $r_t = g(X_t)$ , for some given  $g: \mathfrak{R}^m \rightarrow \mathfrak{R}$ .

A special class and popular representation of multi-factor models is obtained by assuming that  $\mu(X_t) = a + BX_t$ . This mean that we can rewrite formula 7 as:

$$dX_t = (a + BX_t)dt + S(X_t)dW_t \quad (8)$$

---

<sup>34</sup> Generally speaking, a diffusion process is an arbitrary strong Markov process with continous sample path.

<sup>35</sup> The opposite is however not the case.

## A 3-Factor Model for the Yield-Curve Dynamics

---

In the literature it is usually assumed that the spot-rate is given as a linear combination of the state-variables, ie:

$$r_t = r_t(X_t) = \omega^T X_t \quad (9)$$

where  $\omega^T$  is a Boolean column vector.

### 3.1 Gaussian term-structure models

If we assume that the volatility in formula 8 is deterministic, we have a Gaussian representation. Multi-factor Gaussian models among others, have been treated by Langetieg (1980), Jamshidian (1990), Karoui and Lacoste (1992) and Duffie and Kan (1996).

We have now that the dynamics of  $X_t$  is governed by the following SDE:

$$dX_t = (a + BX_t)dt + SdW_t \quad (10)$$

where  $a$  is an  $m \times 1$  vector,  $B$  is an  $m \times m$  matrix. In the general case we have that the volatility matrix is found by Cholesky factorization of the variance-covariance matrix and that the  $m$ -Brownian motions are correlated<sup>36</sup>. Furthermore we assume that the spot rate is given by a linear combination of the state-variables, ie:  $r_t = \omega^T X_t$ .

A special case of this general Gaussian model is the one-factor model of Vasicek (1977), which is of the following form:

$$dr_t = \kappa(\theta - r_t)dt + \sigma dW_t \quad (11)$$

As Vasicek assumes that  $\kappa > 0$ , formula 11 is known as an elastic random walk. Assuming that  $\kappa > 0$  is necessary if the volatility is to be bounded<sup>37</sup>.

In Vasicek's model  $\kappa$  represents the degree of mean-reversion toward the unconditional mean  $\theta$ . The parameter  $\sigma$  is the volatility in the spot-rate.

The solution to formula 10 can be written as (see Karatzas and Schreve (1988)):

$$X_s = e^{-B(s-t)}X_t + \int_t^s e^{-B(s-v)}a dv + \int_t^s e^{-B(s-v)}SdW_v \quad (12)$$

---

<sup>36</sup> We have, as is general the case in the literature also assumed that the number of state-variables equal the number of Brownian motions.

<sup>37</sup> In the general case  $\kappa \leq 0$  the volatility can only be bounded in a finite time-dimension.

## - The case of stochastic spot-rate, market price of risk and volatility

---

where it is implicitly assumed that  $B$ ,  $a$  and  $S$  satisfy the suitable integrability conditions so that the process is well defined.

The reason why we have the fundamental solution matrix given as a matrix exponential function is because we have assumed time-homogeneity. The general fundamental solution matrix defined by  $\Phi$  is given by:

$$\Phi(s - t) = B\Phi(s - t) \quad \text{for } \Phi(t - t) = I \quad (13)$$

which under the assumption of time-homogeneity has the form:

$$\Phi(s - t) = e^{-B(s - t)} = I + \sum_{n=1}^{\infty} B^n \frac{(s - t)^n}{n!} \quad (14)$$

The conditional distribution of  $X_t$  is multivariate normal and its conditional mean and variance is given by:

$$\begin{aligned} E[X_s|X_t] &= e^{-B(s-t)}X_t + \int_t^s e^{-B(s-v)}a dv \\ V[X_s|X_t] &= \int_t^s e^{-B(s-v)}SS^T e^{-B^T(s-v)}dv \end{aligned} \quad (15)$$

Let us now denote the price of a zero coupon bond maturing at time  $T$  as  $P(t,T)$ .

By using a standard arbitrage argument we can show that  $P(t,T)$  satisfy the following PDE:

$$\frac{1}{2}tr\left(\frac{\partial^2 P}{\partial X_t \partial X_t^T} SS^T\right) + \frac{\partial P}{\partial X_t^T}[a + BX_t - S\lambda^T] + \frac{\partial P}{\partial t} - r(X_t)P = 0 \quad (16)$$

subject to the terminal condition  $P(T,T) = 1$ .

As can be seen from formula 16 the use of the arbitrage argument results in a change of drift. Technically speaking we have changed the probability measure from  $P$  to  $Q$  and are now working under the risk-neutral probability measure.

Let us recall, for convenience, a few basic facts concerning the notion of equivalent probability measures on a filtration of a Brownian motion. Firstly, it is well known that any probability measure  $Q$  equivalent to  $P$  on  $(\Omega, \mathcal{F})$  is given by the Radon-Nikodym derivative:

$$\frac{dP}{dQ} = \exp\left(-\int_0^t \lambda(s)dW(s) - \frac{1}{2}\int_0^t \|\lambda(s)\|^2 ds\right) \quad (17)$$

## A 3-Factor Model for the Yield-Curve Dynamics

---

for some predictable  $\mathfrak{R}^m$ -valued process  $\lambda$ . Given an adapted process  $\lambda$  we define  $Q$  as the probability measure where the Radon-Nikodym derivative with respect to  $P$  is given by the right hand side of formula 17.

Using Girsanovs theorem<sup>38</sup>, the process:

$$W_t^Q = W_t - \int_0^t \lambda(s) ds \quad (18)$$

follows an  $m$ -dimensional standard Brownian motion under  $Q$ . This means that under the probability measure  $Q$  we have that  $X_t$  is governed by the following SDE:

$$\begin{aligned} dX_t &= (b + BX_t)dt + Sd\tilde{W}_t \\ \text{for} \\ b &= a - S\lambda^T \end{aligned} \quad (19)$$

The arbitrary process  $\pi_t$  of any attainable contingent claim  $c(t,T)$  which is of the form  $c(t,T) = f(X_T, T)$  for some function  $f: \mathfrak{R} \times (0, T) \rightarrow \mathfrak{R}$  is given by the risk neutral valuation formula:

$$c(t,T) = \pi_t = \tilde{E} \left[ f(X_T, T) e^{-\int_t^T r_s(X_s) ds} \middle| F_t \right] \quad (20)$$

It now follows from the general theory of diffusion processes, more precisely from the result known as Feynman-Kac formula (see Karatzas and Schreve (1988 section 5.7.6)), that under mild technical assumptions, the valuation function  $c(t,T)$  solves the fundamental PDE from formula 17.

If we now consider a zero-coupon bond, where  $P(T,T) = 1$ , we can rewrite 20 as:

$$P(t,T) = \tilde{E} \left[ e^{-\int_t^T r_s(X_s) ds} \middle| F_t \right] = \exp \left[ -\omega^T \int_t^T E[X_s | X_t] ds + \frac{1}{2} \omega^T \int_t^T V[X_s | X_t] ds \omega \right] \quad (21)$$

The last equality in formula 21 follows from the normal distribution assumption. From this we deduce that the price of a zero coupon bond can be written as:

$$P(t,T) = e^{[C(\tau) + D(\tau)^T X_t]} \quad (22)$$

---

<sup>38</sup> We need of course, to show that an application of Girsanovs theorem is justified. Essentially this means imposing certain restrictions which guarantee that  $Q$  is indeed a probability measure equivalent to  $P$ .

## - The case of stochastic spot-rate, market price of risk and volatility

---

where  $C(\tau)$  and  $D(\tau)$  satisfy the joint system of ordinary differential equations ODEs:

$$\begin{aligned} \frac{dD(\tau)}{d\tau} &= -B^T D(\tau) - \omega \\ \frac{dC(\tau)}{d\tau} &= \frac{1}{2} \text{tr}[D(\tau)D(\tau)^T SS^T] + D(\tau)^T b \end{aligned} \quad (23)$$

with the boundary condition  $C(0) = 0$  and  $D(0) = 0$ . Another way to derive the functional form for  $C(\tau)$  and  $D(\tau)$  is to work out the integrals in equation 21.

Langetieg (1980) derives the general functional forms for  $C(\tau)$  and  $D(\tau)$  relying on the assumption that  $B$  is non-singular. His method is however more complicated than the risk neutral valuation method employed here.

In Appendix A - using a probabilistic approach - it is shown that  $C(\tau)$  and  $D(\tau)$  under the assumption of a non-singular  $B$  can be expressed as:

$$\begin{aligned} D(\tau) &= \omega^T B^{-1} [e^{-B\tau} - I] \\ C(\tau) &= \omega^T B^{-1} [I - e^{-B\tau}] B^{-1} b - \omega^T B^{-1} b(\tau) \\ &- \frac{1}{2} [\omega^T B^{-1} [B^{-1} [I - e^{-B\tau}] SS^T + SS^T [I - e^{-B\tau}] (B^{-1})^T] (B^{-1})^T \omega \\ &- \omega^T B^{-1} SS^T (B^{-1})^T \omega(\tau) - \omega^T B^{-1} \Omega(\tau) (B^{-1})^T \omega] \end{aligned} \quad (24)$$

*where*

$$\Omega(\tau) = \int_t^T e^{-B(s-t)} SS^T e^{-B(s-t)^T} ds$$

This expression is not completely worked out since the integral - represented by  $\Omega(\tau)$  - still remains to be evaluated<sup>39</sup>.

The relationship from formula 24 will for  $B = \kappa$ ,  $X_t = r$ ,  $S = \sigma$ ,  $a = \kappa\theta$  and  $\lambda = \lambda$  degenerate into the Vasicek model - more precisely we get:

---

<sup>39</sup> Even though we have written the solution in formula 24 using an eigenvalue decomposition in order to evaluate the matrix exponential (and the integral of the matrix exponential) - this might not be the most efficient way, as mentioned by Golub and Van Lan (1993) and Lund (1997). In the case of a non-singular  $B$  matrix we have however - with our implementation of the eigenvalue decomposition (the general QZ-algorithm from Moler and Stewart (1973)) - not encountered any numerical problems. With respect to the integral represented by  $\Omega(\tau)$  it is worth pointing out that the most efficient way is to work out each element in the matrix  $e^{-B(T-t)} SS^T e^{-B(T-t)^T}$  wise - an eigenvalue decomposition as mentioned by Langetieg (1980) is not recommendable for solving the integral  $\Omega(\tau)$ . More generally the method advocated by Van Loan (1978) - which employs diagonal Padé approximation with scaling and squaring - can be used.

## A 3-Factor Model for the Yield-Curve Dynamics

---

$$\begin{aligned}
 D(\tau) &= \frac{e^{-\kappa(\tau)} - 1}{\kappa} \\
 C(\tau) &= \frac{[\kappa\theta - \sigma\lambda][1 - e^{-\kappa(\tau)}]}{\kappa^2} - \frac{[\kappa\theta - \sigma\lambda](\tau)}{\kappa} - \frac{\sigma^2}{\kappa^3}[1 - e^{-\kappa(\tau)}] \\
 &\quad + \frac{\sigma^2}{2\kappa^2}(\tau) + \frac{\sigma^2}{2\kappa^2} \int_t^T e^{-2\kappa(s-t)} ds \\
 &= R(\infty) \left[ \frac{1 - e^{-\kappa(\tau)}}{\kappa} - \tau \right] - \frac{1}{4} \frac{\sigma^2}{\kappa^3} [1 - e^{-\kappa\tau}]^2
 \end{aligned} \tag{25}$$

*where*

$$R(\infty) = \theta - \frac{\sigma\lambda}{\kappa} - \frac{1}{2} \frac{\sigma^2}{\kappa^2}$$

For the multi-factor Gaussian yield-curve model given by the relations in formula 24 the asymptotic behavior for  $T \rightarrow \infty$  can be written as<sup>40</sup>:

$$R(\infty) = B^{-1}b - \frac{1}{2} B^{-1}SS^T(B^{-1})^T \tag{26}$$

In order for formula 26 to be valid the model has to be stationary. Stationarity is obtained if all B's eigenvalues are negative - or more generally - have negative real components. The stationarity condition is obviously fulfilled for the Vasiceks, as it is assumed that the spot rate follows an elastic random walk - ie  $\kappa > 0$ .

Unless all the eigenvalues have negative real components then the variance will explode in finite time - which means that the process from an econometric perspective will not be stationary. However, even if one or more of the eigenvalues have positive real components this might not be a problem from an economic point of view. From an econometric perspective non-stationarity is most important in the probability density function for the state space - whereas from a pricing perspective we are focusing on the risk-neutral measure. Unless the security in question has an infinite life - then non-stationarity in the risk neutral measure is less significant - as the conditional densities (from formula 15) is (in general) well defined for finite time-horizons.

If I relate this general multi factor model to the Heath, Jarrow and Mortons framework, it can be shown that the forward volatility structure can be expressed as<sup>41</sup>:

$$\sigma^F(t, T) = Q e^{-A(T-t)} Q^{-1} S \tag{27}$$

---

<sup>40</sup> Which can be recognized as being identical to  $R(\infty)$  - for suitable definitions of a, B and  $SS^T$  - for the Vasicek model in formula 25.

<sup>41</sup> See Madsen (1995).

## - The case of stochastic spot-rate, market price of risk and volatility

---

which means the the bond price volatility structure can be written as:

$$\sigma_p(t,T) = \int_t^T \sigma^F(t,s) ds = QA^{-1}[I - e^{-A(T-t)}]Q^{-1}S \quad (28)$$

and the yield-curve (spot rate) volatility structure as:

$$\sigma_R(t,T) = \frac{1}{D_{T-t}} \sigma_p(t,T) \quad (29)$$

where  $D_{T-t}$  is the duration at time  $t$  for a zero-coupon bond that matures at time  $T$ .

The SDE for bond prices can therefore be expressed as:

$$\frac{dP(t,T)}{P(t,T)} = [r + \sigma_p^T \lambda] dt + \sigma_p^T dW \quad (30)$$

where  $\sigma_p(t,T)$  is defined as in formula 28.

Models that are included in this general formulation are for example the Beaglehole and Tenney (1991) double-decay model, Vasicek, and the time-homogenous version of the two-factor Hull and White (1994) model.

### 3.2 The Exponential-Affine class of term-structure models

Duffie and Kan (1996) propose a general class of term-structure models that include the Gaussian models from the last section as a special case.

Under the risk neutral probability measure we have that  $X_t$  is governed by the following SDE:

$$\begin{aligned} dX_t &= (b + BX_t)dt + \Sigma S(X_t) d\tilde{W}_t \\ &\text{for} \\ b &= a - S(X_t)^2 \lambda^T \end{aligned} \quad (31)$$

where  $S(X_t)$  is an  $m \times m$  diagonal matrix of the following form:

$$S(X_t) = \begin{bmatrix} \sqrt{\alpha_1 + \beta_1 X_t(1)} & 0 & \cdot & 0 \\ 0 & \sqrt{\alpha_2 + \beta_2 X_t(2)} & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & \sqrt{\alpha_m + \beta_m X_t(m)} \end{bmatrix}. \quad (32)$$

## A 3-Factor Model for the Yield-Curve Dynamics

---

that is, the Wiener processes are independent. The dependency between the processes  $X_t$  is captured by the  $m \times m$  matrix  $\Sigma$ .

A solution to 31 given 32 must keep the processes  $\alpha_i + \beta_i X_t$ , for all  $i$ , always on the positive real line. Sufficient conditions for this are (see Duffie and Kan (1996))<sup>42</sup>:

$$\begin{aligned} \forall X_t \in \mathfrak{R}^m, \quad S(X_t)_{ii} = 0 &\Rightarrow \beta_i [b + BX_t] \geq \frac{1}{2} \beta_i \Sigma \Sigma^T \beta_i \\ \forall j, \quad (\Sigma \beta)_j \neq 0 &\Rightarrow \beta_i = \beta_j \quad \text{and} \quad \alpha_i = \alpha_j \end{aligned} \quad (33)$$

It should however be pointed out that the Condition in line 1 in formula 33 is generally not satisfied - beyond the Gaussian case  $\beta_i = 0$ , for all  $i$ .

It should here be emphasised that condition 1 is for example satisfied for the (stacked) multi dimensional extended CIR model<sup>43,44</sup>. It also follows from equation 31 that the vector of market prices of risk can be generated from a general equilibrium model (simply take the multi-dimensional version of the extended Cox, Ingersoll and Ross (CIR) (1985) model).

As the model is Affine in the state-variables we have that the bond price can be expressed as:

$$P(t, T) = e^{[C(\tau) + D(\tau)^T X_t]} \quad (34)$$

Under the general specification of the SDE in equation 31 we have that  $C(\tau)$  and  $D(\tau)$  satisfy the joint system of ordinary differential equations ODEs:

$$\begin{aligned} \frac{dD(\tau)}{d\tau} &= \frac{1}{2} \sum_{i=1}^m [\Sigma^T D(\tau)]_i^2 \beta_i - B^T D(\tau)^T - \sum_{i=1}^m \lambda_i [\Sigma^T D(\tau)]_i \beta_i - \omega \\ \frac{dC(\tau)}{d\tau} &= \frac{1}{2} \sum_{i=1}^m [\Sigma^T D(\tau)]_i^2 \alpha_i + D(\tau)^T a - \sum_{i=1}^m \lambda_i [\Sigma^T D(\tau)]_i \alpha_i \end{aligned} \quad (35)$$

with the boundary condition  $C(0) = 0$  and  $D(0) = 0$ . For general specifications it does not seem possible to obtain a closed form solution for  $C(\tau)$  and  $D(\tau)$ , however these ODEs can

<sup>42</sup> Condition 1 implies the usual Feller condition.

<sup>43</sup> In this case  $B$  is a diagonal matrix with the mean-reversion for each of the factors in the diagonal. If  $B$  is a diagonal matrix then Condition 1 implies that  $\Sigma$  must be a diagonal matrix - that is correlation between the Wiener processes are not allowed.

<sup>44</sup> Generally speaking, the intuition in Condition 1 is that a sufficiently positive drift near the boundary where the volatility for that particular process is zero will ensure that this boundary is never hit. In this connection it is worth mentioning that Condition 2 ensures that at the boundary where the  $i$ 'th process is zero, then no other volatility terms will interact with the  $i$ 'th process so as to drive it to become negative. For elaboration see Duffie and Kan (1996).

## - The case of stochastic spot-rate, market price of risk and volatility

---

easily be solved through numerical integration.

If we now consider the (stacked) multi dimensional extended CIR model, which among others has been analysed in Cox, Ingersoll and Ross (1985), Chen and Scott (1995), Duffie and Singleton (1996) and Duan and Simonato (1995), equation 31 can be simplified as<sup>45</sup>:

$$\begin{aligned} dX_i &= (b_i + \kappa_i X_i)dt + c_i \sqrt{\alpha_i + \beta_i X_i} d\tilde{W}_i(i) \quad \text{for } i \in (1, 2, \dots, m) \\ &\text{for} \\ b_i &= a_i - \lambda_i [\alpha_i + \beta_i X_i] \end{aligned} \quad (36)$$

where it usually is assumed that  $\alpha_i = 0$  and  $c_i = 1$ , for all  $i$ .

Without loss of generality we assume that  $c_i = 1$ , and as it is assumed that the spot rate is defined through a linear combination of the state-variables ie  $r_t = \omega^T X_t$ , we can express the price of a zero-coupon bond  $P(t, T)$  as<sup>46</sup>:

$$\begin{aligned} P(t, T) &= \prod_{i=1}^m e^{[C_i(t) + D_i(t)X_i]} \\ &\text{where} \\ C_i(\tau) &= \ln \left[ \frac{2\gamma_i e^{\frac{1}{2}[\gamma_i + \kappa_i + \lambda_i \beta_i]\tau}}{[\gamma_i + \kappa_i + \lambda_i \beta_i]e^{\gamma_i \tau} + [\gamma_i - \kappa_i - \lambda_i \beta_i]} \right]^{\frac{2[\alpha_i - \lambda_i \alpha_i]}{\beta_i}} \\ &+ \frac{1}{2} \alpha_i \left[ \frac{r_{1i} [\beta_i r_{1i} - 4\gamma_i]}{\beta_i} \tau - \frac{4r_{1i}}{\beta_i} \ln \left[ \frac{r_{2i} - r_{1i}}{g_i(\tau)} \right] + \left( \frac{2\gamma_i r_{1i}}{\beta_i} \right)^2 \left[ \frac{\tau}{r_{1i}^2} - \frac{1}{\gamma_i r_{1i} g_i(\tau)} - \frac{1}{\gamma_i r_{1i} [r_{1i} - r_{2i}]} + \frac{1}{\gamma_i r_{1i}^2} \ln \left[ \frac{r_{2i} - r_{1i}}{g_i(\tau)} \right] \right] \right] \\ &\text{and} \\ D_i(\tau) &= - \frac{2[e^{\gamma_i \tau} - 1]}{[\gamma_i + \kappa_i + \lambda_i \beta_i]e^{\gamma_i \tau} + [\gamma_i - \kappa_i - \lambda_i \beta_i]} \\ &\text{where} \\ \lambda_i &= \sqrt{[\kappa_i + \lambda_i \beta_i]^2 + 2\beta_i} \\ g_i(\tau) &= r_{2i} e^{\gamma_i \tau} - r_{1i} \\ r_{1i} &= \frac{-[\kappa_i + \lambda_i \beta_i] + \gamma_i}{\beta_i} \\ r_{2i} &= \frac{-[\kappa_i + \lambda_i \beta_i] - \gamma_i}{\beta_i} \end{aligned} \quad (37)$$

This is a fairly general specification for the exponential class of Affine yield-curve models. It

---

<sup>45</sup> In Dai and Singleton (1997) these models are termed AYD(N,N) models. In AYD(N,N) models we have

that  $\sum_{i=1}^m \omega = m$  - that is the stochastic system for specifying the short rate is not over-identified - that is there is no feedback between the state-variables through, for example, the drifts, ie  $B$  is a diagonal matrix. An over-identified specification is when the short rate is determined through a strict subset of the state-variables.

<sup>46</sup> See Appendix B for the derivation of formula 37. It is however worth mentioning that this is the first time in the literature (as far as I know) that a solution to the general stacked exponential Affine yield-curve specification from equation 36 has been derived.

## A 3-Factor Model for the Yield-Curve Dynamics

---

can for example be seen that for  $\alpha_i = 0$ , the expression will degenerate into the stacked multi-dimensional CIR model<sup>47</sup>.

From equation 37 we can deduce that the bond price volatility is given by:

$$\sigma_P(t, T) = \prod_{i=1}^m \left[ \frac{2[e^{\gamma_i t} - 1]}{[\gamma_i + \kappa_i + \lambda_i \beta_i]e^{\gamma_i t} + [\gamma_i - \kappa_i - \lambda_i \beta_i]} \right] \sqrt{\alpha_i + \beta_i X_i} \quad (38)$$

That is the volatility is level dependent - which of course is not the case for the Gaussian model from section 3.1. We also have the yield-curve (spot rate) volatility structure as:

$$\sigma_R(t, T) = \frac{1}{D_{T-t}} \sigma_P(t, T) \quad (39)$$

where  $D_{T-t}$  is the duration at time  $t$  for a zero-coupon bond that matures at time  $T$ .

### 4. Converting Affine Term Structure models into Yield-Factor models

For the general class of Affine models investigated in section 3.1 and section 3.2 we have the following expression for the  $T$ -period spot-rate:

$$R(t, T) = -\frac{C(t, T)}{T - t} - \frac{D(t, T)^T}{T - t} X_t \quad \text{for all } T > t \quad (40)$$

Let us now suppose that we wish to determine the yield on  $n$  different securities, with maturity dates  $N = [T_1, T_2, \dots, T_n]$ , for  $T_1 > t$ , and we will also assume that  $n > m$  - that is, there are more securities than state-variables/Wiener-processes<sup>48</sup>.

#### Definition 1

**In a yield-factor model we have that the yield calculated from equation 40 must be identical to the actual yield for a given maturity date  $T_i$ , ie the parameters must be chosen such that the bond price satisfies the following constraints:**

---

<sup>47</sup> Other examples are given in Appendix B.

<sup>48</sup> Remember namely that we have assumed - which is normally the case in the literature (except Madsen (1998) and Bhar and Chiarella (1995)) - that the number of state-variables coincides with the dimension of the Brownian motion.

- The case of stochastic spot-rate, market price of risk and volatility

$$\begin{aligned}
 C(t, T_i) = D_j(t, T_i) = 0, \text{ for } j \neq i \\
 \text{and} \\
 D_i(t, T_i) = (T_i - t) \text{ for } i \in [1, 2, \dots, n]
 \end{aligned}
 \tag{41}$$

Under Definition 1 we have that the m-state-variables - which in equation 40 are left unspecified - now are represented by adhoc specification of m distinct maturity dates.

More precisely we have that the state-variables are related to the yield-factors through the following expression:

$$X_t = - \left( \frac{D(t, V)}{V - t} \right)^{-1} \left[ R(t, V) + \frac{C(t, V)}{V - t} \right]
 \tag{42}$$

where V is an m-dimensional vector of maturity dates, and  $V \subset \mathbb{N}$ .

The result in formula 42 relies on the fact that  $\frac{D(t, V)}{V - t}$  is non-singular. In this connection it is straightforward to prove that it is always possible to find a set of distinct maturities V such

that  $\frac{D(t, V)}{V - t}$  is non-singular when the model is non-degenerate<sup>49</sup>.

The Estimating of an Affine multi-factor model using the yield-factor approach has been investigated by Chen and Scott (1993), Pearson and Sun (1994) and Duffie and Singleton (1996).

The differences between adopting the yield-factor approach in the estimation of multi-factor yield-curve models and the approach where the state-variables are left unspecified can be formulated as follows:

The selection of m distinct maturites in the yield-factor model as state-variables is of course bound to be arbitrary. From a practical point of view this however has some important implications:

Firstly, it is in this case possible to relate each of the factor-loadings directly to observable economic measures. It is namely very important from a hedging point of view that the factors that drive the dynamic in the yield-curve are observable in the market. Strictly speaking by observable we mean that the state-variable is priced/quoted in the market. .

---

<sup>49</sup> By non-degenerate we mean, that no state-variables are redundant. In other words, we cannot find a change of variables that allows for the existence of fewer state-variables.

## A 3-Factor Model for the Yield-Curve Dynamics

---

Secondly, even though it might be possible to assume that the unspecified state-variables can be recognized as being a fairly good representation of some observable economic variables - this is generally not enough from a hedging point of view. For example, if now one of the state-variables is the unconditional mean in the SDE for the spot rate<sup>50</sup>, then the “long-rate” would here be the relevant observable variable to relate this state-variable to - but the question is, which long-rate?

That is, even though the selection of yields as state-variables in the yield-factor model by nature is arbitrary it makes connection with available and relevant/observable market information straightforward.

It is from this possible to conclude that by estimating a multi-factor yield-curve model using the general approach of unspecified state-variables, we encounter the same problems as when we employ the PCA techniques for determining the appropriate number of factors that drive the dynamics in the yield-curve - namely that applying appropriate economic meaning to the factor-loadings/state-variables is not straightforward<sup>51</sup>.

### 5. A three-factor model for the dynamic in the yield-curve

In order to utilize the information in the yield-curve when estimating the parameters, we generally need a multi-factor model that is fairly tractable. This limits us to models in the class of Affine yield-curve models<sup>52</sup>. Furthermore if we are interested in using the models for pricing and hedging purposes<sup>53</sup> we need models that are fairly tractable. What I mean by tractable - is that we wish to be in a situation where it is relatively easy to price a zero-coupon bond, which means either a true closed-form solution or a fairly easily solved semi-closed form solution.

From the discussion in section 2 we have decided to specify our 3-factor model in the following way:

---

<sup>50</sup> In the Beaglehole and Tenney (1991) double-decay model, which has been estimated by Lund (1997) under the assumption of unspecified state-variables - this is the kind of result that could be obtained here.

<sup>51</sup> It is however worth emphasising that in PCA models the dynamic part of the factor-structure is not analysed whereas this is exactly the purpose in multi-factor yield-curve models, see Madsen (1998).

<sup>52</sup> Another candidate model could be the SAINTS model from Constantinides (1992), which is a non-linear model but where a closed form solution is available in a multi-factor setting - this will however not be pursued here.

<sup>53</sup> Which must be considered to be the ultimate goal.

## - The case of stochastic spot-rate, market price of risk and volatility

---

$$\begin{aligned}
 dr_t &= [a - \kappa r_t - \Gamma]dt + \sqrt{\sigma_0 + \sigma_1 r_t} dW_1 \\
 d\Gamma &= [b - \kappa_2 \Gamma]dt + \sigma_2 dW_2 \\
 d\sigma_0 &= [c - \kappa_3 \sigma_0]dt + \sigma_3 \sqrt{\sigma_0} dW_3
 \end{aligned}
 \tag{43}$$

where we assume non-correlation between the Brownian Motions.

The process for the volatility might appear strange - let me for that reason explain the motivation. We assume that the volatility in the spot-rate consists of two independent parts. The first part is purely stochastic and is incorporated for capturing the persistence in volatility. In this connection we assume that  $\sigma_0$  follows a square-root process - that is  $\sigma_0$  is restricted to being positive<sup>54</sup>. The second part takes on the other hand care of the level-effect in the process for the spot-rate. As is normally the case in the literature we also assume that the market price of risk for volatility is not priced in the market.

The assumption of a stochastic market price of risk in the process for the spot-rate has the following implications:

- First, we have removed the interactions of parameters in the drift-specification
- Secondly, we are now able to observe the mean-reversion parameter directly

Of course assuming that the market price of risk can be both positive and negative introduces a positive probability for getting negative interest rates - as will be apparent from section 6 this did not give us any problems at all, that is, no action had to be taken on that account.

Even though we assume no correlation between the Wiener processes in the three-factor model there is interaction between the processes, as the process for the spot-rate depends on the SDEs for the market price of risk and the volatility.

Let us now assume that the price of a zero-coupon bond can be expressed as:

$$P(t, T) = e^{A(\tau) + B(\tau)r + C(\tau)\Gamma + D(\tau)\sigma_0} ; \quad \tau = T - t
 \tag{44}$$

Using a standard arbitrage-argument we can express the PDE - that all contingent claims which are only a function of  $[\tau, r, \Gamma, \sigma_0]$  have to satisfy with respect to their boundary conditions - as<sup>55</sup>:

---

<sup>54</sup> This of course means that we have to fulfill the so-called Feller condition - ie  $\sigma_3^2 < 2\kappa_3\theta_3$

<sup>55</sup> I will interchangeably use the notation P and P(t, T) - whenever appropriate.

## A 3-Factor Model for the Yield-Curve Dynamics

---

$$\begin{aligned}
 P_t + P_r a - P_r \kappa r - P_r \Gamma + \frac{1}{2} P_{rr} \sigma_0 + \frac{1}{2} P_{rr} \sigma_1 r \\
 + P_{\Gamma} b - P_{\Gamma} \kappa_2 \Gamma + \frac{1}{2} P_{\Gamma\Gamma} \sigma_2^2 \\
 P_{\sigma_0} c - P_{\sigma_0} \kappa_3 \sigma_0 + \frac{1}{2} P_{\sigma_0 \sigma_0} \sigma_3^2 \sigma_0 - rP = 0
 \end{aligned} \tag{45}$$

Where we have that the boundary condition for a zero-coupon bond is:  $P(T,T) = 1$ .

Under the assumption that the solution can be expressed as in formula 44 the solution to the PDE in formula 45 can be obtained by solving the following system of ODE's:

$$\begin{aligned}
 A_{\tau} + B_{\tau} a + D_{\tau} c + \frac{1}{2} C^2 \sigma_2^2 &= 0 \\
 B_{\tau} r - B_{\tau} \kappa r + \frac{1}{2} B^2 \sigma_1 r - r &= 0 \\
 C_{\tau} \Gamma - B_{\tau} \Gamma - C_{\tau} \kappa_2 \Gamma &= 0 \\
 D_{\tau} \sigma_0 - D_{\tau} \kappa_3 \sigma_0 + \frac{1}{2} B^2 \sigma_0 + \frac{1}{2} D^2 \sigma_3^2 \sigma_0 &= 0
 \end{aligned} \tag{46}$$

A closed form solution is not available for this system of ODE's - but they are easily solved numerically<sup>56</sup>.

Even though we cannot express the price of a zero-coupon bond analytically for the three-factor model - we can fairly easily calculate the factor-loadings. This indicates that the model even though it is untractable still can be useful as a practical hedging tool. Using the model for the pricing of options and other contingent claims is however not straightforward - whether this is manageable from a practical point of view I will leave for future research.

## 6. Estimation procedure

In principle we can estimate the parameters using the Kalman filter, as the 3-factor model from formula 43 is Affine, which means that it can be written in state space form.

Under the assumption that the measurement errors are additive and normally distributed we can express the measurement equation of the state space model as follows:

---

<sup>56</sup> For that purpose we use a re-written version of the ODEINT routine from Press, Teukolsky, Vetterling and Flannery (1994).

- The case of stochastic spot-rate, market price of risk and volatility

$$R(x,\tau) = -\frac{A(\psi,\tau)}{\tau} - \frac{B(\psi,\tau)}{\tau}r(x) - \frac{C(\psi,\tau)}{\tau}\Gamma(x) - \frac{D(\psi,\tau)}{\tau}\sigma_0(x) + \varepsilon_x \quad (47)$$

*where*  
 $\varepsilon_x \approx N(0,H(\psi))$

where the vector  $\psi$  contains the set of parameters for the state space model. Furthermore we have that  $\tau$  is an N-dimensional vector of time-to-maturities.

As we are estimating the continuous time model using discrete observations, the transition equation has to be derived from the exact discrete-time distribution of the state variables. In abstract form we can express the transition equation as:

$$X_x = c_x(\psi) + \Phi_x(\psi)X_{x-1} + \eta_x(\psi) \quad (48)$$

*where*  
 $\eta_x(\psi) \approx N(0,V_x(\psi))$   
*and*  
 $X_x = [r(x), \Gamma(x), \sigma_0(x)]$

where  $c_x(\psi)$  is an N-dimensional vector,  $\Phi_x(\psi)$  is an N x 3 dimensional matrix and  $\eta_x(\psi)$  is an 3 x 3 dimensional matrix. Inspection of equation 48 reveals that the transition equation is given as the solution to the multi-dimensional SDE for the state-variables. In general we have that the SDE for the state variables in the Affine yield-curve models can be expressed as follows using matrix-notation<sup>57</sup>:

$$dX_t = (b + BX_t)dt + \Sigma S(X_t)d\tilde{W}_t \quad (49)$$

which alternatively can be written as:

$$X_s = e^{-B(s-t)}X_t + \int_t^s e^{-B(s-v)}adv + v(t,s) \quad (50)$$

*where*

$$v(t,s) = \int_t^s e^{-B(s-v)}\Sigma S(X_v)d\tilde{W}_v$$

In the Kalman filter algorithm - which consists of a sequence of prediction and update steps - we need the conditional mean and conditional covariance for the state variables<sup>58</sup>.

As Ito stochastic integrals are martingales we have that the conditional mean of  $v(t,s)$  is zero.

---

<sup>57</sup> See formula 19 and 31.

<sup>58</sup> See Harvey (1993).

## A 3-Factor Model for the Yield-Curve Dynamics

---

From this - under the assumption that  $B$  is non-singular - the conditional mean follows directly from formula 50:

$$E[X_s|X_t] = e^{-B(s-t)}X_t + B^{-1}[I - e^{-B(s-t)}]a \quad (51)$$

The conditional covariance is defined as<sup>59</sup>:

$$Cov[X_s|X_t] = E[v(s,t)v(s,t)^T|X_t] = \int_t^s e^{-B(s-v)}\Sigma E[S(X_v)|X_t]\Sigma^T e^{-B(s-v)^T}dv \quad (52)$$

where this equation in the Gaussian case degenerates into the integral equation for  $\Omega(\tau)$  from formula 24.

Under the assumption of a Gaussian distribution of the state variables it is well known that the linear Kalman filter is optimal<sup>60</sup> - in the sense, that we obtain the exact likelihood function from the prediction error decomposition, ie<sup>61</sup>:

$$Log_L = -\frac{1}{2}\sum_{x=1}^L \log|F_x| - \frac{1}{2}\sum_{x=1}^L v_x^T F_x^{-1} v_x \quad (53)$$

However, in our case we have that the conditional covariance in the transition equation depends on lagged state-variables, which means that the linear Kalman filter is no longer optimal as we do not obtain the exact likelihood function.

In the estimation of multi-factor CIR specifications for the dynamic in the yield-curve, Chen and Scott (1995) and Duan and Simonato (1995) propose a Gaussian QML approach that is based on the linear Kalman filter.

Their approach is as follows:

The transition equation is derived from the first and second conditional moments of the state-variables, that is they are obtained from formula 51 and 52. More precisely we have that  $V_x(\psi) = Cov[X_x|X_{x-1}]$ .

The linear Kalman filter is now modified in two ways. First, as  $\eta_x$  now depends on lagged state-variables,  $V_x(\psi)$  are evaluated at  $\hat{X}_{x-1}$  when computing  $\hat{\eta}_{x-1}$  in the prediction step (see

---

<sup>59</sup> This formula has among others been derived by Lund (1997) and Duan and Simonato (1995).

<sup>60</sup> Here we implicitly assume that the errors in the measurement equation are also normally distributed.

<sup>61</sup>  $F_x$  and  $v_x$  is derived from the Kalman filter, see Appendix C where the steps in the linear Kalman filter recursion is shown.

## - The case of stochastic spot-rate, market price of risk and volatility

formula 6 in Appendix C). Secondly, as the state-variables (at least a subset of them) are restricted to be non-negative, and such a restriction is not taken into account in the linear update step (see formula 7 (line 1) in Appendix C) - we need to adjust the linear update step. Chen and Scott and Duan and Simonato suggest in this case the replacement of negative values for the state-variables with a zero (0) (assuming that the state-variable in question is restricted to be non-negative).

Apart from these changes the Kalman filter recursions are identical to the formulas shown in Appendix C for the linear Kalman filter. The QML estimation criterion is in this case identical to the prediction error decomposition from formula 53, apart from the fact that we need to scale it by  $\frac{2}{L}$ .

However as mentioned by Duan and Simonato (1995) and Lund (1997) the conditions for consistency for the QML estimator<sup>62</sup> are not satisfied by the exponential Affine state-space model. In this connection Lund (1997) proposes a modified QML-estimator that achieves consistency but at the cost of ignoring certain aspects of the yield-curve dynamics. More precisely Lund proposes to ignore the dependency of the state-variables in the transition equation.

This means that if we wish to use the Kalman filter algorithm when estimating Affine yield-curve models of the general kind (that is, we do not restrict ourselves to Gaussian models) we need to decide what is the worst of the following two things:

- A QML-estimator that is not consistent
- Treating the transition equation as if the model were Gaussian

These observations make it compelling to abandon the Kalman filter (and QML) for non-Gaussian Affine term structure models. Besides being inconsistent, the QML-estimator is inefficient compared to maximum likelihood.

Other approaches can however be used which provide consistent estimators which also attain full asymptotic efficiency. One candidate is the Bayesian approach. Another possibility is the simulation based EMM procedure from Dai and Singleton (1997). Still another candidate is the Indirect Inference methodology from Gourieroux, Monfort and Renault (1993). Compared to the Kalman filter algorithm these estimation procedures are very computationally involved - for that reason I will not pursue these techniques here but instead leave them for further research.

As mentioned earlier I will utilize the yield-factor approach - that is I will perform a change of variables, which means I will map the state-variables into observed yields on zero-coupon

---

<sup>62</sup> Sufficient conditions for the QML estimator is  $E_{x-1}[\mathbf{v}_x] = \mathbf{0}$  and  $E_{x-1}[\mathbf{v}_x \mathbf{v}_x^T] = \mathbf{F}_x$   
Bollerslev and Wooldridge (1992).

## A 3-Factor Model for the Yield-Curve Dynamics

---

bonds. Mapping the state-variables into observed state-variables, which then by construction are observed without measurement errors has the following important implication:

- In this case we can derive the exact maximum likelihood function, which is identical to the relationship in equation 53. The estimation procedure is then in principle identical to the estimation procedure which in general is used in connection with GARCH-models, see Greene (1993, section 19.7.1)

Compared to the Kalman filter algorithm, the yield-factor approach is just as efficient seen from a computational perspective. From this we can deduce that the “cost” of subjectively selecting 3-maturities that are observed without measurement errors is that we get an efficient estimation procedure.

### 6.1 Estimation of the 3-factor Yield-factor model

Estimation of the unknown parameter vector  $\psi = [a, \kappa, \sigma_1, b, \kappa_2, \sigma_2, c, \kappa_3, \sigma_3]$  is performed using weekly data on the Danish Bond Market from the period 2 January 1990 to 30 June 1998. The zero-coupon bond yields we will use in order to obtain information across the yield-curve consist of the following nine (9) time-to-maturities  $\tau = [0.25, 0.5, 1, 2, 3, 4, 5, 10, 12.5]$  - which are chosen to be constant over the time-period.

We have selected the following 3 state-variables:

- The short-rate represented by the 1 month yield (the Short-factor)
- The spread between the long-rate and the short-rate (the Slope-factor), where the long-rate is the 15 year rate - that is the state-variable is being defined as:  $R(15) - R\left(\frac{30}{360}\right)$
- The Butterfly-factor, defined as follows:  $2R(7.5) - \left[ R(15) + R\left(\frac{30}{360}\right) \right]$

The implications for selecting these three state-variables is that the 1 month rate, the 7.5 year rate and the 15 year rate are observed without measurement errors.

The evolution in the observed factors<sup>63</sup> over the time-period are shown below in figure 1.

---

<sup>63</sup> Investigation of these equations reveals that the difference between the slope-factor and the butterfly-factor is equal to the following yield-spread  $2[R(15) - R(7.5)]$ .

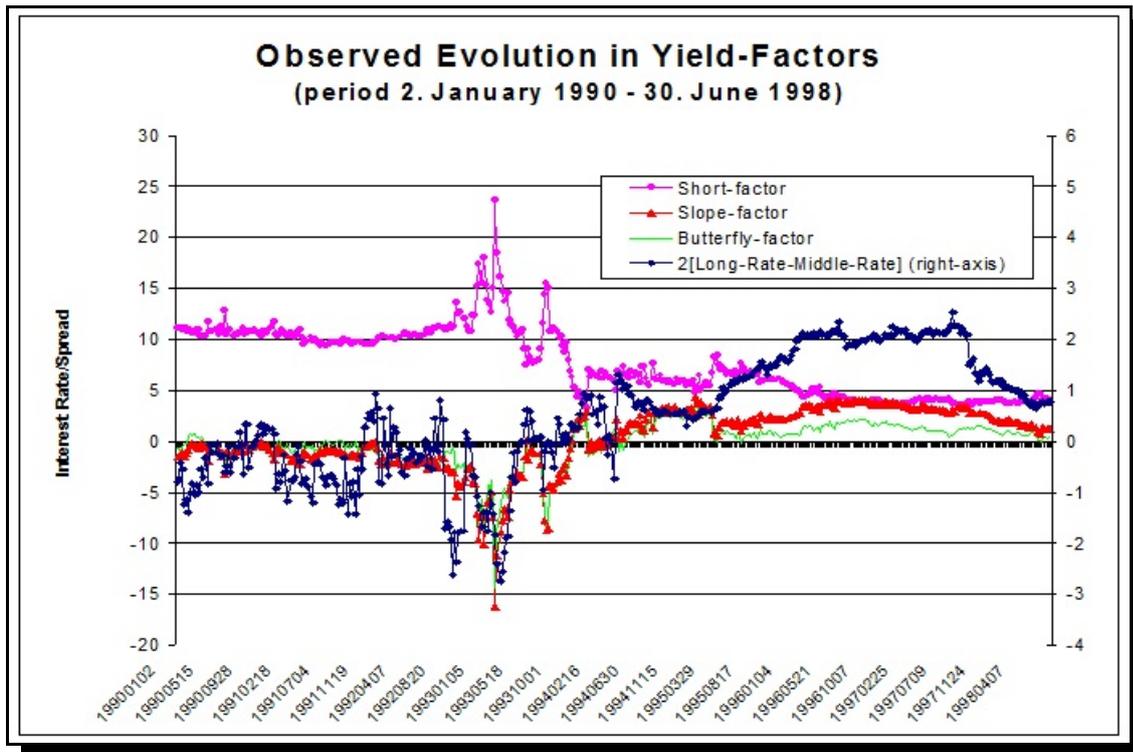


Figure 1

By observing figure 1 it may be concluded that over the period 2 January 1990 - 30 June 1998 the yield-curve has had a lot of different shapes, ranging from approximately flat, to generally downward-and upward sloping.

Before continuing let me present a few stylized facts about the Danish Bond Market represented by our  $\tau$ -vector of time-to-maturity dates: [0.25,0.5,1,2,3,4,5,10,12.5].

**Table 1:** Some Statistics on the Danish Bond Market  
(for the period 2 January 1990 - 30 June 1998)

|                           | 3-Month Rate | 6-Month Rate | 1-Year Rate | 2-Year Rate | 3-Year Rate | 4-Year Rate | 5-Year Rate | 10-Year Rate | 12.5-Year Rate |
|---------------------------|--------------|--------------|-------------|-------------|-------------|-------------|-------------|--------------|----------------|
| <b>Mean</b>               | 7,38         | 7,16         | 7,04        | 7,10        | 7,24        | 7,37        | 7,49        | 7,83         | 7,91           |
| <b>Variance</b>           | 51,87        | 24,38        | 18,06       | 14,08       | 11,04       | 8,87        | 7,44        | 4,72         | 4,14           |
| <b>Skewness</b>           | 0,19         | 0,07         | 0,03        | 0,01        | 0,02        | 0,03        | 0,04        | -0,07        | -0,12          |
| <b>Kurtosis</b>           | 1,73         | 1,47         | 1,48        | 1,55        | 1,62        | 1,70        | 1,79        | 2,33         | 2,52           |
| <b>ACF<sup>64</sup> 1</b> | 0,99         | 0,99         | 0,99        | 0,99        | 0,99        | 0,99        | 0,99        | 0,98         | 0,98           |

<sup>64</sup> ACF x - represents the autocorrelation-function for x-lags.

## A 3-Factor Model for the Yield-Curve Dynamics

|               |      |      |      |      |      |      |      |      |      |
|---------------|------|------|------|------|------|------|------|------|------|
| <b>ACF 5</b>  | 0,95 | 0,96 | 0,96 | 0,96 | 0,96 | 0,95 | 0,95 | 0,93 | 0,92 |
| <b>ACF 10</b> | 0,90 | 0,92 | 0,93 | 0,92 | 0,91 | 0,90 | 0,89 | 0,85 | 0,84 |
| <b>ACF 20</b> | 0,84 | 0,86 | 0,85 | 0,83 | 0,80 | 0,78 | 0,76 | 0,68 | 0,66 |
| <b>ACF 50</b> | 0,61 | 0,64 | 0,62 | 0,56 | 0,51 | 0,46 | 0,41 | 0,24 | 0,18 |

In table 1 we have shown some statistics on the yields we intend to fit our model against. From the table it follows that as expected the volatility structure is downward sloping<sup>65</sup>. Another thing worth noting is that the persistence in rates (symbolized by ACF) is higher for short-term bonds than for long-term bonds. We have also performed a goodness-of-fit test for the hypothesis that the individual data-series can be assumed to be normally distributed with a first- and second order moment as the historical distribution. The general conclusion is as follows<sup>66</sup>:

- The normal distribution assumption is generally accepted for number fractile-intervals of no more than 10
- The normal distribution assumption is generally rejected as we increase the number of fractile-intervals

### 6.2 Optimization Procedure

We have estimated our 3-factor model by minimizing the log-likelihood function from formula 53. In this case we have assumed that the variance matrix  $F_x$  is constant for all  $x$ , more precisely we assume that  $F_x$  is of the following form:  $F_x = \sigma^2 \tau I$ , where  $\sigma$  is an unknown parameter that is estimated together with the process parameters,  $\tau$  is the time-to-maturity vector and  $I$  is the identity matrix.

When minimizing the log-likelihood function we have been using different optimization procedures in order to ensure that the estimated parameters are the optimal solution<sup>67</sup>. In this connection we have estimated the model by using the following three methods<sup>68</sup>:

---

<sup>65</sup> In this connection it is worth pointing out that we have calculated Variance, Skewness and Kurtosis in the table by centering them around the mean.

<sup>66</sup> See Appendix E where the test-statistics is shown for respectively 10 and 20 fractile-intervals.

<sup>67</sup> In general when estimating the log-likelihood function with  $m$  unknown parameters (when  $m$  is high) we cannot ensure that the estimate is optimal - we can only render probable the result.

<sup>68</sup> For a good introduction to these techniques see Harvey (1990, section 4) and Greene (1993, section 12).

## - The case of stochastic spot-rate, market price of risk and volatility

---

- Quasi maximum likelihood procedures using the BFGS/DFP<sup>69</sup> updating formula for the Hessian - which as long as the initial Hessian matrix is positive definite will ensure that the Hessian always will be positively definite<sup>70</sup>
- The BHHH algorithm from Berndt, Hall, Hall and Hausman (1974)
- The score-method combined with the information-matrix

To improve efficiency we use the method of Marquardt<sup>71</sup> (1963) which is based on the same principle as the quadratic hill-climbing method to make a slight change in the iterations-scheme in order to ensure a downhill movement.

As we also constrain some of the parameters to lie within certain intervals - we combine these optimization methods with a pivoting technique to handle the criteria that the parameters in the non-linear optimization procedure have to adhere to certain restrictions. For example it is worth mentioning that we constrain the mean-reversion parameters to be non-negative in order to ensure stationarity. Furthermore we restrict the unconditional mean in both the spot-rate process and the volatility process to be non-negative.

For initializing the different algorithms we use different starting values combined with the downhill simplex-method. As is well known the downhill simplex-method is very robust for “bad” starting values and furthermore it only needs the function to be minimized - neither the first or second partial derivatives are required. This makes the downhill simplex-method a powerful method for obtaining starting values for the main optimization procedures.

### 6.3 Estimation Results

In the table below the estimated parameters are shown. In this connection we might mention that in all cases the different optimization procedures converged to approximately the same results<sup>72</sup>:

**Table 2: Parameter Estimates for the 3-factor Affine Yield-Factor model**  
(period 2 January 1990 - 30 June 1998)

---

<sup>69</sup> The DFP and the BFGS methods are inherently identical. The only difference being that the DFP method is an update sequence for the inverse of the Hessian whereas the BFGS is an update sequence for the Hessian, see Hald (1979).

<sup>70</sup> Here I have chosen to initialize the Hessian using the identity matrix which means that the first iteration is carried out by the method of steepest descent.

<sup>71</sup> This is generally a good approach in connection with the BHHH algorithm, as the Hessian here is approximated by the Jacobian which might not always be positively definite. In connection with the score-method it has also proved to increase performance.

<sup>72</sup> The standard-errors are calculated using the exact Hessian matrix at the optimal parameter estimates.

## A 3-Factor Model for the Yield-Curve Dynamics

|  | Parameter Value | Standard-Error <sup>73</sup> |
|--|-----------------|------------------------------|
| <b>Spot-Rate: mean-reversion</b><br>- $\kappa$               | 6.7301          | 0,249632                     |
| <b>Spot-Rate: unconditional mean</b> - $\theta$              | 0.0476          | 0,009203                     |
| <b>Spot-Rate: volatility (level-related)</b> - $\sigma_1$    | 0.0093          | 0,000831                     |
| <b>Market-Price-of-Risk: mean-reversion</b> - $\kappa_2$     | 0.0223          | 0,001516                     |
| <b>Market-Price-of-Risk: unconditional mean</b> - $\theta_2$ | -0.0467         | 0,008391                     |
| <b>Market-Price-of-Risk: volatility</b> - $\sigma_2$         | 0.0055          | 0,001466                     |
| <b>Volatility: mean-reversion</b><br>- $\kappa_3$            | 0.2594          | 0,015942                     |
| <b>Volatility: unconditional mean</b> - $\theta_3$           | 0.0073          | 0,001176                     |
| <b>Volatility: volatility</b> - $\sigma_3$                   | 0.0012          | 0,000865                     |

The most striking features of this table are:

- The spot-rate has a strong degree of mean-reversion
- The process for the market-price of risk is close to being non-stationary, as the mean-reversion parameter is fairly small
- The unconditional mean in the process for the market-price of risk is negative
- The standard-errors are much smaller than the estimated parameters and the parameters are all significant

In the literature for multi-factor term structure models a general result is also that the spot-rate has a high degree of mean-reversion, see for example Duan and Simonato (1995), Chen and Scott (1995) and Andersen and Lund (1997). As mentioned in section 2 a general result from

---

<sup>73</sup> In connection with the standard errors it is of importance to point out the following. They should not be taken as literal as the big difference between the estimated parameter and the standard error is not only due to efficiency, but also (maybe more) related to the fact that the exact Hessian is not robust in the Newey-West sense - that is we (probably) have a certain degree of autocorrelation and heteroskedasticity. This is also the explanation why we have not included the t-statistics in table 2 as we are led to believe that the statistic is not t-distributed. But despite these short-comings, the parameters appear to be fairly significant.

## - The case of stochastic spot-rate, market price of risk and volatility

---

the literature is also that the level-related process is close to being non-stationary - this is also what we observe.

In table 3 below we show some statistics for the model-generated interest-rate series which are also shown in table 1:

**Table 3:** Some Statistics for the model-generated interest rate series  
(for the period 2 January 1990 - 30 June 1998)

|                 | 3-<br>Month<br>Rate | 6-<br>Month<br>Rate | 1-Year<br>Rate | 2-Year<br>Rate | 3-Year<br>Rate | 4-Year<br>Rate | 5-Year<br>Rate | 10-<br>Year<br>Rate | 12.5-<br>Year<br>Rate |
|-----------------|---------------------|---------------------|----------------|----------------|----------------|----------------|----------------|---------------------|-----------------------|
| <b>Mean</b>     | 7,51                | 7,20                | 7,02           | 7,08           | 7,22           | 7,36           | 7,48           | 7,84                | 7,91                  |
| <b>Variance</b> | 75,50               | 36,67               | 26,55          | 17,60          | 12,24          | 9,24           | 7,54           | 4,75                | 4,17                  |
| <b>Skewness</b> | 0,33                | 0,16                | 0,10           | 0,09           | 0,07           | 0,06           | 0,04           | -0,06               | -0,11                 |
| <b>Kurtosis</b> | 2,28                | 1,80                | 1,69           | 1,67           | 1,68           | 1,73           | 1,80           | 2,33                | 2,50                  |
| <b>ACF 1</b>    | 0,98                | 0,98                | 0,98           | 0,98           | 0,99           | 0,99           | 0,99           | 0,98                | 0,98                  |
| <b>ACF 5</b>    | 0,91                | 0,93                | 0,93           | 0,93           | 0,94           | 0,94           | 0,94           | 0,93                | 0,91                  |
| <b>ACF 10</b>   | 0,86                | 0,88                | 0,88           | 0,87           | 0,88           | 0,88           | 0,88           | 0,85                | 0,83                  |
| <b>ACF 20</b>   | 0,77                | 0,80                | 0,80           | 0,78           | 0,78           | 0,76           | 0,75           | 0,68                | 0,65                  |
| <b>ACF 50</b>   | 0,51                | 0,53                | 0,53           | 0,50           | 0,47           | 0,44           | 0,40           | 0,24                | 0,18                  |

Comparing the results for the model-generated interest rate series from table 3 with the actual interest rate series from table 1 it can be seen that they have approximately the same characteristics.

We have performed the following statistical test in order to analyse the relationship between the actual data and the model-generated data<sup>74</sup>:

- The Median test - that is a test to see whether two random samples could have come from the same frequency distribution
- The Wilcoxon-Mann-Whitney U-test - that is a test to see whether two random samples could come from two populations with the same mean<sup>75</sup>
- The Spearman rank correlation test - that is a test to determine the

---

<sup>74</sup> We restrict ourselves to the use of distribution-free tests - because we do not want to impose any distribution assumptions on the interest data series. We have in that connection selected 4 different non-parametric tests, because the distribution-free tests are not as strong as tests that rely on particular distribution assumptions. For additional information on these statistical tests see Kanji (1995).

<sup>75</sup> This test is known as the non-parametric t-test.

## A 3-Factor Model for the Yield-Curve Dynamics

- significance between two series
- The Kolmogor-Smirnov goodness of fit test - that is a test to see whether two samples could have been randomly selected from the same population/distribution

The test-statistics for each of the 9-series are shown below in table 4:

**Table 4:** Distribution-free tests for the model generated interest rate series  
(for the period 2 January 1990 - 30 June 1998)

|   | 3-Month Rate | 6-Month Rate | 1-Year Rate | 2-Year Rate | 3-Year Rate | 4-Year Rate | 5-Year Rate | 10-Year Rate | 12.5-Year Rate |
|---|--------------|--------------|-------------|-------------|-------------|-------------|-------------|--------------|----------------|
| <b>Combined Median</b>  | 6,79         | 6,88         | 6,86        | 7,05        | 7,20        | 7,39        | 7,54        | 8,00         | 8,06           |
| <b><math>\chi^2</math>-Statistics (1-degree of Freedom)</b>     | 1,50         | 2,67         | 0,67        | 0,02        | 0,02        | 0,07        | 0,00        | 0,00         | 0,00           |
| <b>1-<math>\chi^2</math> probability</b>                        | 0,2201       | 0,1021       | 0,4137      | 0,8916      | 0,8916      | 0,7853      | 1,0000      | 1,0000       | 1,0000         |
| <b>U-Statistics</b>   | 87887        | 91580        | 92135       | 92436       | 92538       | 92342       | 92444       | 92665        | 92579          |
| <b>Standardized Normal Z</b>                                    | -1,366       | -0,356       | -0,204      | -0,122      | -0,094      | -0,147      | -0,120      | -0,059       | -0,082         |
| <b>Significance Level</b>                                       | 0,0859       | 0,3609       | 0,4192      | 0,4515      | 0,4626      | 0,4414      | 0,4524      | 0,4765       | 0,4671         |
| <b>Spearman Rank Correlation</b>                                | 0,98         | 0,96         | 0,97        | 0,98        | 0,99        | 0,99        | 1,00        | 1,00         | 1,00           |
| <b>Computed Normal Z Value</b>                                  | 20,25        | 20,00        | 20,14       | 20,37       | 20,51       | 20,63       | 20,67       | 20,72        | 20,72          |
| <b>Significance Level</b>                                       | 0,0000       | 0,0000       | 0,0000      | 0,0000      | 0,0000      | 0,0000      | 0,0000      | 0,0000       | 0,0000         |
| <b>Largest Absolute Difference (10 Cells)</b>                   | 0,0510       | 0,0766       | 0,0418      | 0,0278      | 0,0441      | 0,0255      | 0,0162      | 0,0070       | 0,0093         |
| <b><math>\chi^2</math> approximation</b>                        | 2,2459       | 5,0534       | 1,5035      | 0,6682      | 1,6752      | 0,5615      | 0,2274      | 0,0418       | 0,0742         |
| <b>1-<math>\chi^2</math> probability (2 degrees of Freedom)</b> | 0,3253       | 0,0799       | 0,4715      | 0,7160      | 0,4328      | 0,7552      | 0,8925      | 0,9793       | 0,9636         |

## - The case of stochastic spot-rate, market price of risk and volatility

From table 4 we conclude:

- We cannot reject the hypothesis that the median for the model generated interest rate series is equivalent to the median for the actual interest rate series - this is true for all maturities
- We cannot reject the hypothesis that the mean for the model generated interest rate series is equivalent to the mean for the actual interest rate series - this is true for all maturities
- We cannot reject the hypothesis that the model generated interest rate series are highly correlated with the actual interest rate series - this is true for all maturities
- We cannot reject the hypothesis that the model generated interest rate series and the actual interest rate series come from the same distribution - this is true for all maturities

From this we conclude that the model seems to be able to capture the dynamic in the yield-curve.

It is however worth pointing out that it seems as though it is the short maturities which the model has most difficulty in capturing. We have no explanation<sup>76</sup> for this phenomenon - but perhaps by selecting another rate as a proxy for the short-rate we could change it - this however is left for further research.

By making a change of variables we map the “true” state-variables into new state-variables that are constructed in such a way that the subjectively defined yield-factors are observed without measurement errors. In this connection it could be interesting to see how these implied state-variables evolve over time, this is shown below in figure 2:

---

<sup>76</sup> The correlation matrix does not give any indication of why this could be the case, as the structure of the correlation matrix is as would be expected. The correlation matrix has not been included in the paper but can be obtained from the author.

## A 3-Factor Model for the Yield-Curve Dynamics

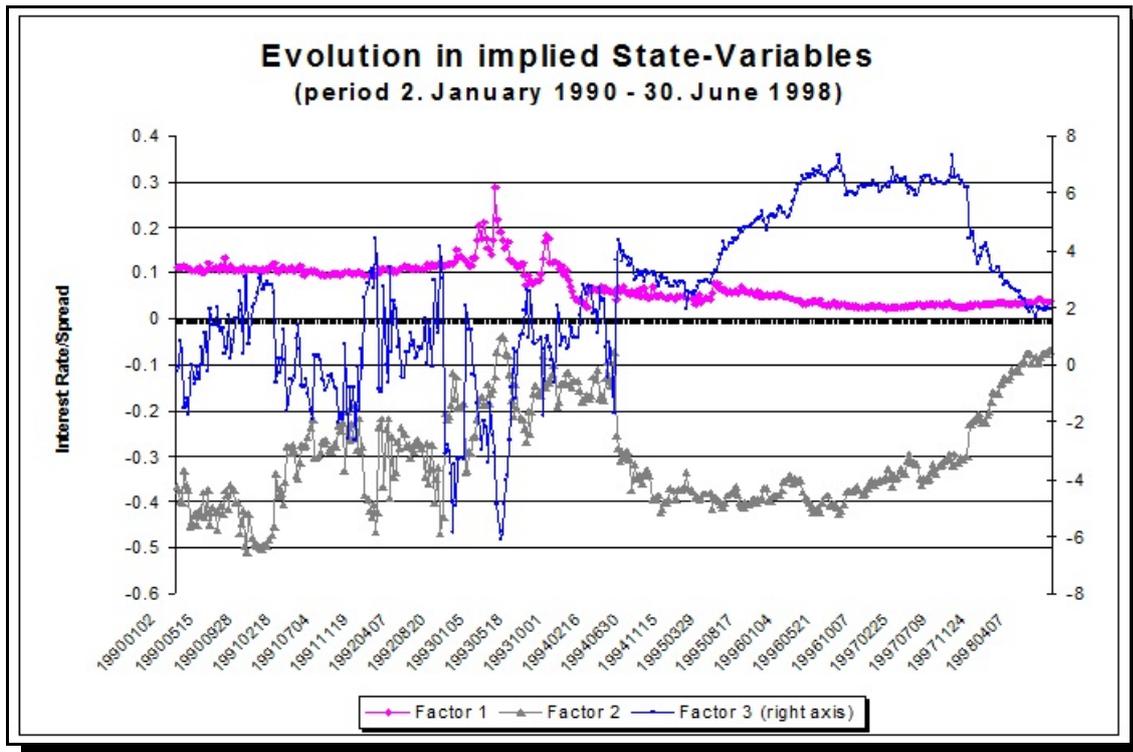


Figure 2

Comparing figure 1 with figure 2 it can be observed that the implied factor 1 is nearly equal to yield-factor 1<sup>77</sup>. Furthermore we can conclude that the implied factor 3 is very closely correlated to the difference between the Butterfly-factor and the Slope-factor<sup>78</sup>.

Deducing the way in which the implied factor 2 is related to the yield-factors is far from straightforward. One reason is probably that the information that is contained in the implied factors 1 and 3 with respect to the yield-curve dynamics are almost identical to the information contained in the yield-factors. However, even though the implied factors 1 and 3 manage to capture the main part of the combined dynamics in the yield-factors - they do not capture the level in rates, see figure 2 in Appendix D. From this we can conclude (as expected) that the implied factor 2 is related to the level of interest rates.

An interesting conclusion that seems logical to draw now is:

- The dynamic in the yield-curve is approximately captured by a 2-factor model - but an additional state-variable is needed to capture the level of yields.

<sup>77</sup> See Appendix D.

<sup>78</sup> See Appendix D.

## - The case of stochastic spot-rate, market price of risk and volatility

If this conclusion is true (which the analysis here seems to indicate) then it also seems logical to assume/expect that the process for the level-factor should be close to non-stationary.

### 7. The effect of State-Variables on the shape of the Term Structure of Interest Rates

In this section we will investigate the relationship between the state-variables and the effects on the shape of the yield-curve.

Before doing this it is natural to show which kind of factor-loadings the model implies, this is shown below in figure 3:

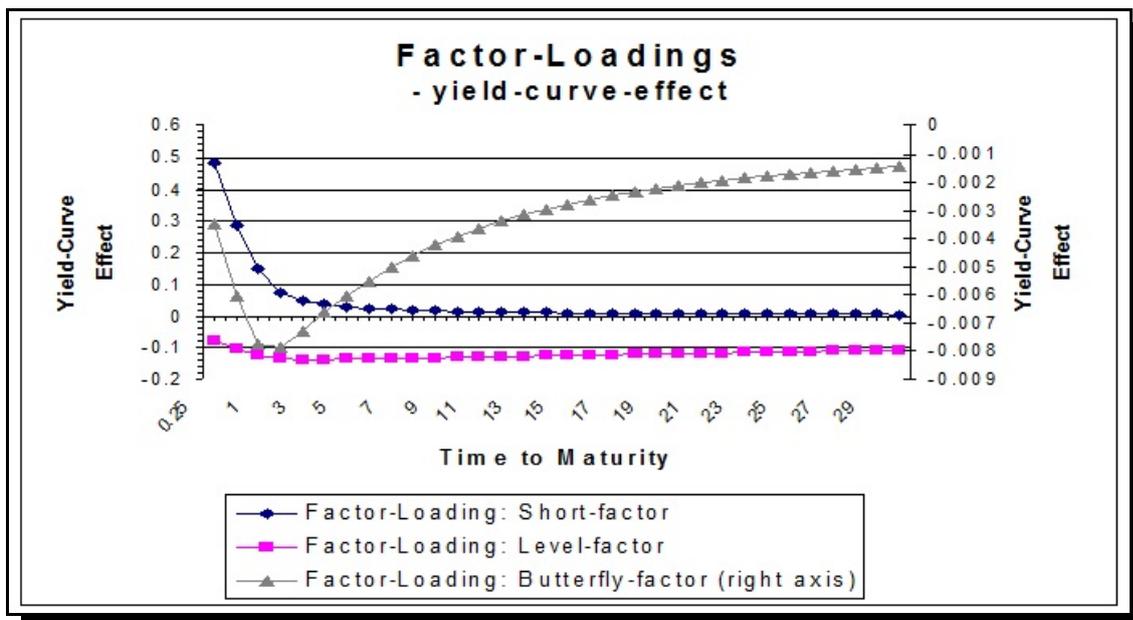


Figure 3

From figure 3 we can deduce that the 3-implied state-variables can be identified as a Short-factor<sup>79</sup>, a Level-factor and a Curvature-factor.

In the following three figures we have shown how the yield-curve is effected by different values of the implied state-variables.

---

<sup>79</sup> Sometimes the Short-factor is referred to as the Steepness-factor, see for example Litterman and Scheinkman (1988).

## A 3-Factor Model for the Yield-Curve Dynamics

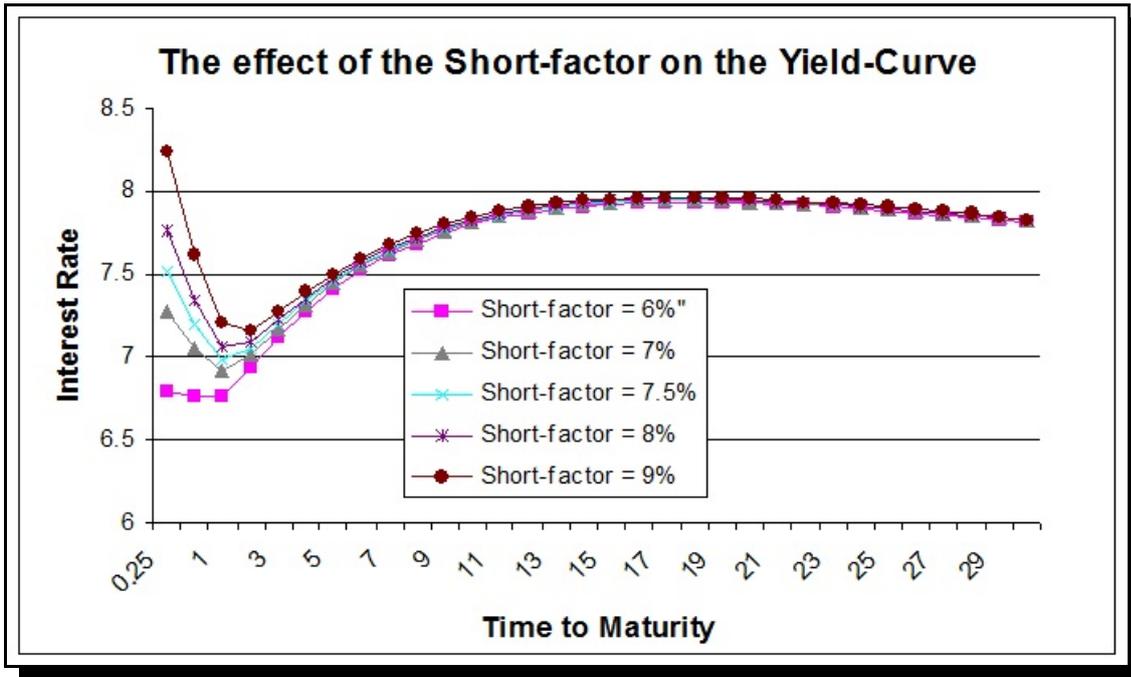


Figure 4

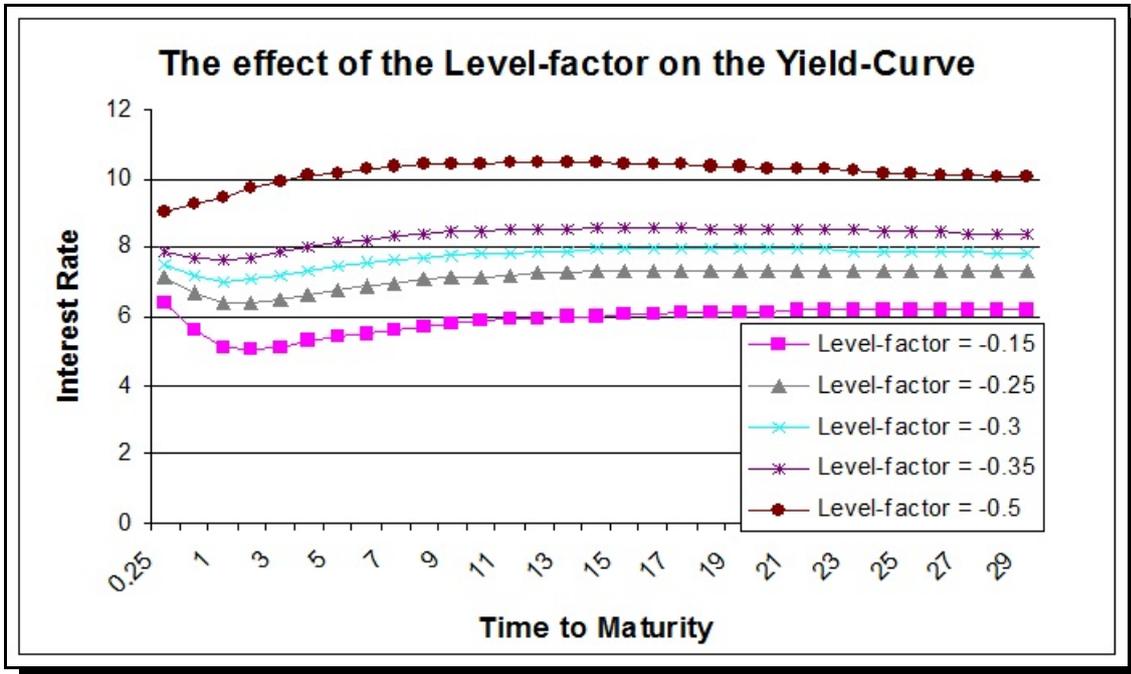


Figure 5

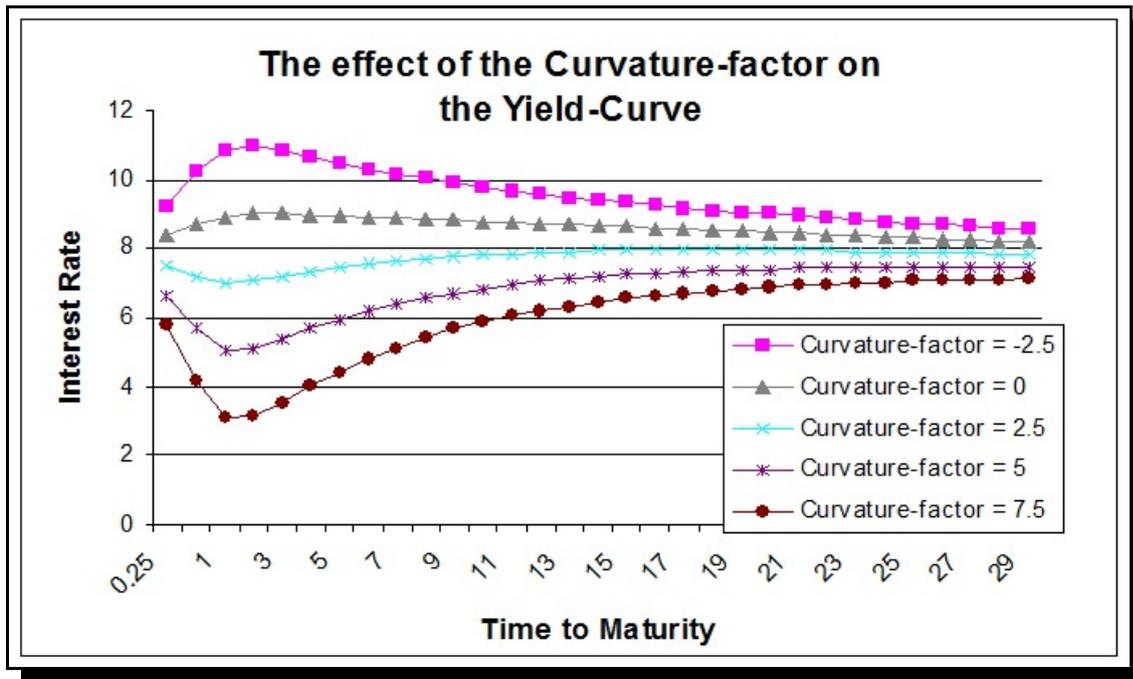


Figure 6

These results are in line with the results reported in for example Litterman and Scheinkman (1991) where 3-dominant factors for the evolution were identified using PCA.

From this we can also conclude the following:

Initially we decided that the 3 yield-factors should be the short-rate, the slope, and the butterfly-spread, see section 6.1. Given our 3-factor model from equation 43 we then implied 3 state-variables using the technique from section 4. These 3 implied state-variables interestingly enough turned out as a Short-factor (Steepness-factor), a Level-factor and a Curvature-factor - with a nearly non-stationary Level-factor. That this should be the case could not be expected, but the result support the fact that (as mentioned before) the dynamic in the yield-curve is governed by three dominant factors - a Short-factor, a Level-factor and a Curvature-factor. As pointed out in Madsen (1998) if there is a non-stationary component in interest rate data then this will turn out as a Level-factor using PCA.

Futhermore these 3-factors seem to pop-up<sup>80</sup> no-matter what stochastic processes are assumed, and how the process parameters are estimated.

All in all this suggests that maybe there actually is a non-stationary component in interest rate data. Offcourse, the analysis here does not prove that - but many factors point in that direction.

<sup>80</sup> See the discussion in section 2.

## A 3-Factor Model for the Yield-Curve Dynamics

---

### 8. Prediction Analysis - Forecasting the evolution in the Yield-Curve

From the results in section 6 these new 3-factor models give a very good description of the dynamic in the yield-curve. All our statistical tests indicated that there is a close relationship between the actual historical interest data series and the model-generated interest rate data series. Among other things we could not reject the hypothesis that the model-generated interest rate series have been selected from the same distribution as the actual interest rate series.

From a practical point of view it is however very important that the model also has some kind of predictable power, because if this is not the case, the factor-loadings are of limited use. Among other things, this renders the model inadequate for use as a tool for hedging yield-curve exposures - that is for risk-managements purpose.

We have for that purpose used the model to predict the evolution in the 6-month interest rate, the 5-year rate and the 10-year rate over the period 1 July 1998 - 30 September 1998 on a daily basis.

The prediction performed will be a “weak prediction” - in the sense that I will not try to forecast the state-variables - but instead take the state-variables as given. By given I mean I will use the observed yield-factors over the period 1 July - 30 September 1998 in order to try to predict the evolution in other yields over the same period.

If the model is to be useful for the purpose of hedging then we need a strong degree of correlation between the yield-curves forecasted by the model given the observed evolution in the yield-factors and the actual observed evolution.

If we also wish to utilize the model for pricing purposes we need to be able to predict the level of interest rates. Given the fact that the Level-factor is close to being non-stationary this is probably too much to hope for - and will for that reason not be pursued here.

In figure 7-9 we have compared the model-generated interest rate with the actual interest rate data for the following maturity dates: 6-months, 5-years and 10-years:

- The case of stochastic spot-rate, market price of risk and volatility

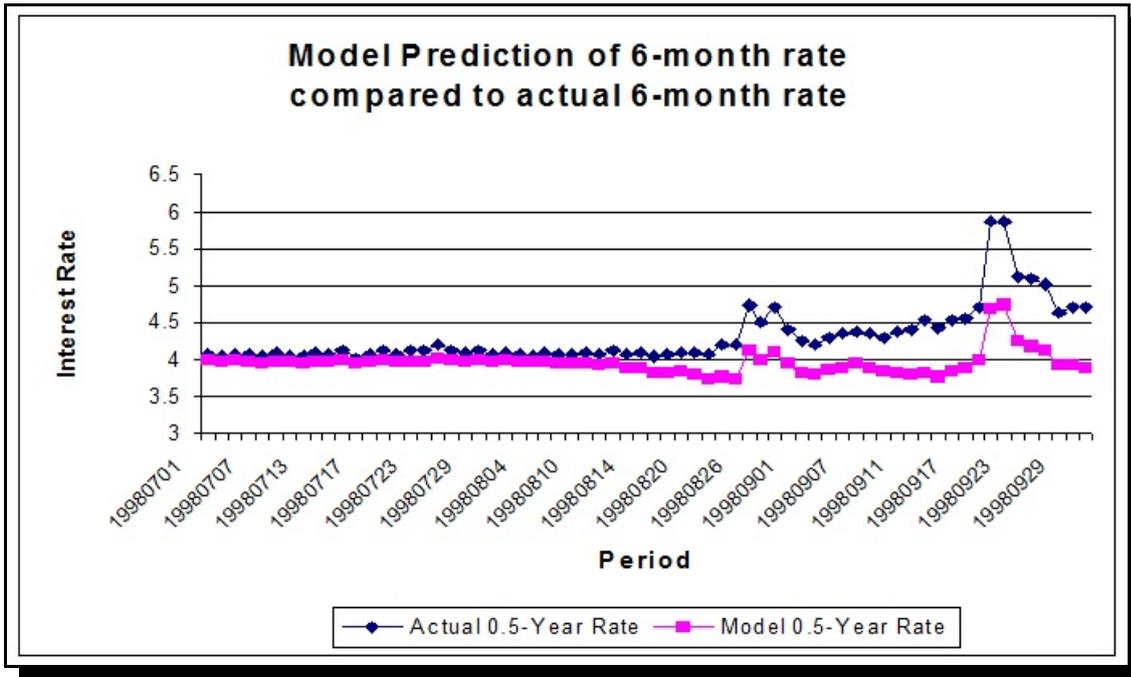


Figure 7

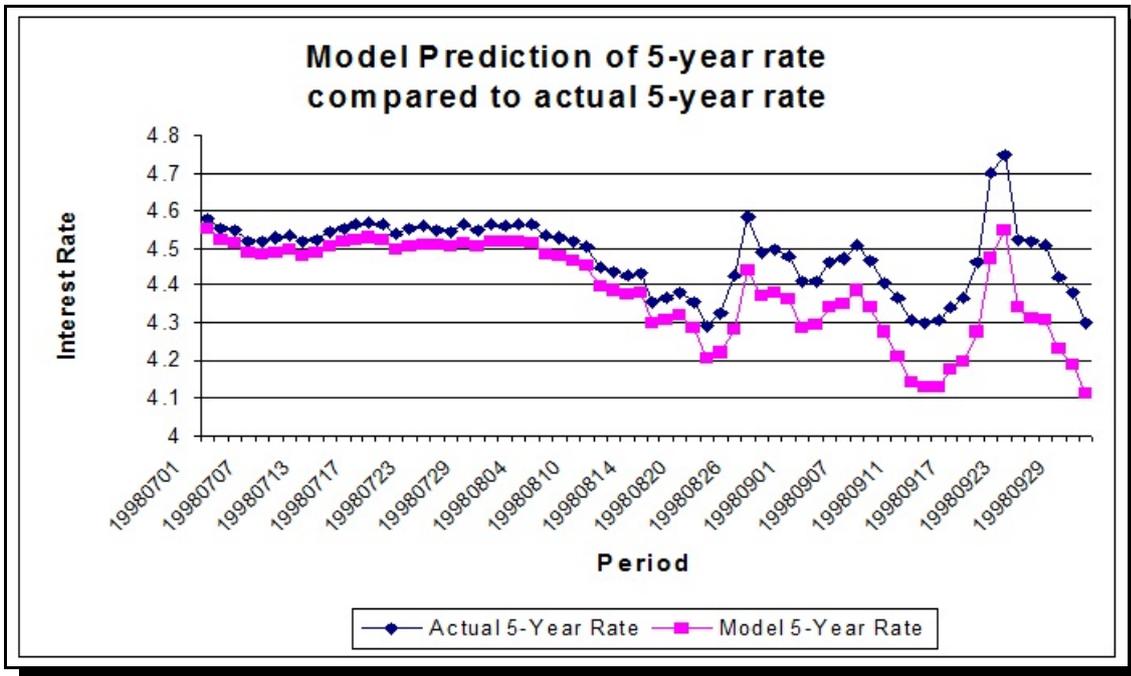


Figure 8

## A 3-Factor Model for the Yield-Curve Dynamics

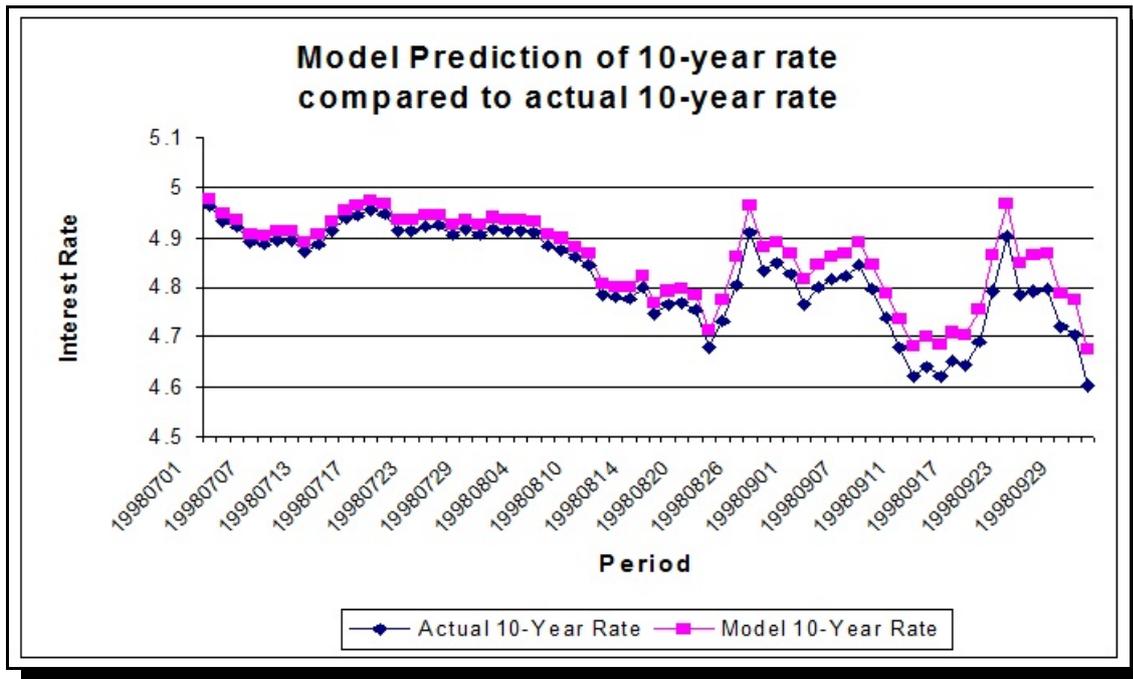


Figure 9

From the figures it appears that the model does manage to capture the dynamic but has some problems with the level of interest rates.

In order to analyse more precisely the extent to which the actual interest rate series and model-generated interest rate series generally agree with each other, we have carried out an ordo-regression - and can report the following results:

### Regression for the 6-month interest rate

| Est. Slope | Sigma  | t-value  | Lower confidence interval | Upper confidence interval |
|------------|--------|----------|---------------------------|---------------------------|
| 1,0299     | 0,0050 | 206,9676 | 1,0200                    | 1,0398                    |

The theoretical t-value at the 0.025 fractile and 65 degrees of freedom = 1.998

### Regressions Analysis Table

| Source            | SS (sum of squares) | DF (degrees of freedom) | MS (mean-sum of squares) | Computed F-value |
|-------------------|---------------------|-------------------------|--------------------------|------------------|
| <b>Total</b>      | 1236,6875           | 66                      | 18,7377                  |                  |
| <b>Regression</b> | 1234,8137           | 1                       | 1234,8137                | 42835            |
| <b>Residual</b>   | 1,8737              | 65                      | 0,0288                   |                  |

- The case of stochastic spot-rate, market price of risk and volatility

|   |       |
|---|-------|
| <b>The multiple correlations coefficient</b>      | 0,999 |
| <b>Degrees of explanation (R<sup>2</sup>)</b>     | 99,80 |
| <b>Regression Significance<sup>81</sup> (95%)</b> | 1,000 |

Regression for the 5-year interest rate

| <b>Est. Slope</b> | <b>Sigma</b> | <b>t-value</b> | <b>Lower confidence interval</b> | <b>Upper confidence interval</b> |
|-------------------|--------------|----------------|----------------------------------|----------------------------------|
| 1,0208            | 0,0018       | 583,2461       | 1,0173                           | 1,0243                           |

The theoretical t-value at the 0.025 fractile and 65 degrees of freedom = 1.998

Regressions Analysis Table

| <b>Source</b>     | <b>SS (sum of squares)</b> | <b>DF (degrees of freedom)</b> | <b>MS (mean-sum of squares)</b> | <b>Computed F-value</b> |
|-------------------|----------------------------|--------------------------------|---------------------------------|-------------------------|
| <b>Total</b>      | 1326,3720                  | 66                             | 20,0966                         |                         |
| <b>Regression</b> | 1326,1186                  | 1                              | 1326,1186                       | 3,40e5                  |
| <b>Residual</b>   | 0,2534                     | 65                             | 0,0039                          |                         |

|   |       |
|---|-------|
| <b>The multiple correlations coefficient</b>  | 1     |
| <b>Degrees of explanation (R<sup>2</sup>)</b> | 100   |
| <b>Regression Significance (95%)</b>          | 1,000 |

Regression for the 10-year interest rate

| <b>Est. Slope</b> | <b>Sigma</b> | <b>t-value</b> | <b>Lower confidence interval</b> | <b>Upper confidence interval</b> |
|-------------------|--------------|----------------|----------------------------------|----------------------------------|
| 0,9926            | 0,0005       | 2014,1681      | 0,9917                           | 0,9936                           |

The theoretical t-value at the 0.025 fractile and 65 degrees of freedom = 1.998

<sup>81</sup> Significance is the area under the curve from 0 to the computed F-value.

## A 3-Factor Model for the Yield-Curve Dynamics

---

Regressions Analysis Table

| Source            | SS<br>(sum of squares) | DF (degrees of<br>freedom) | MS (mean-sum<br>of squares) | Compute<br>d<br>F-value |
|-------------------|------------------------|----------------------------|-----------------------------|-------------------------|
| <b>Total</b>      | 1537,0229              | 66                         | 23,2882                     |                         |
| <b>Regression</b> | 1536,9982              | 1                          | 1536,9982                   | 4,06e5                  |
| <b>Residual</b>   | 0,0246                 | 65                         | 0,0004                      |                         |

|   |       |
|---|-------|
| <b>The multiple correlations coefficient</b>  | 1     |
| <b>Degrees of explanation (R<sup>2</sup>)</b> | 100   |
| <b>Regression Significance (95%)</b>          | 1,000 |

As is well known then, we want to estimate a slope that is close to 1 and at the same time have a high degree of correlation (R<sup>2</sup>). This is because optimally all the observations lie on a straight-line with a 45-degree angle.

From this it appears that the model might be useful as a hedging tool as the correlation between the model generated interest rate data and the actual interest rate data<sup>82</sup> is very high<sup>83</sup>.

However, again we observe that the model clearly has more predictable power for longer maturities than for short maturities.

This is exemplified by figure 10 where we have shown both the actual and model-generated yield-curve from the 30 September 1998. We have in the figure included the volatility structure estimated over the same period.

---

<sup>82</sup> Here represented by the 6-month rate, the 5-year rate and the 10-year rate.

<sup>83</sup> It is actually the case for all maturities that the correlation is fairly high and the slope coefficient is significant and close to 1.

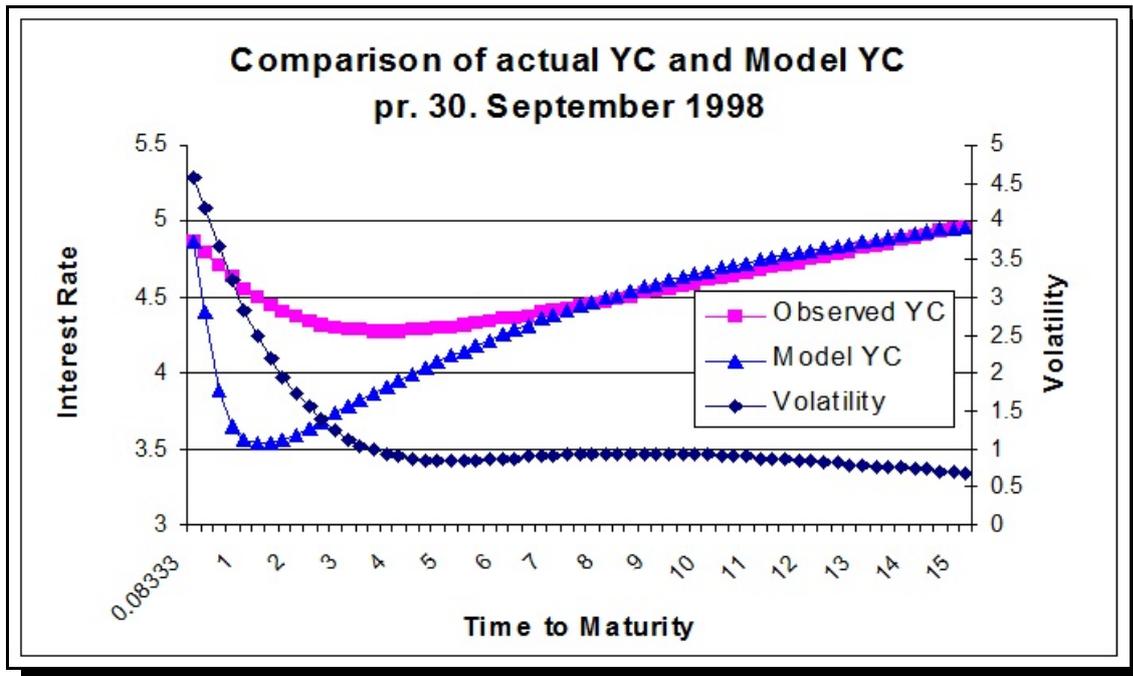


Figure 10

From this figure it is clearly seen that even though the correlation between the evolution in the actual interest rates and the model-generated interest rates is extremely high then the model is not capable of capturing the level of interest rates across the whole yield-curve.

From this we draw the following conclusion:

- Given the state-variables (yield-factors) - which might be “derived” for a particular forecast (macro) of the yield-factors - the model seems to be able to predict the direction of interest rates (the fundamental shape of the yield-curve). This indicates that the model could be appropriate as a hedging tool - that is, the factor-loadings derived are interesting from a risk-mangement perspective. A complete analysis of this lies however outside the scope of this paper and will therefore be left for further research. This conclusion is however further supported by the  $\chi^2$ -test for consistency in a 2x2 table which is reported in Appendix F
- The model is however in adequate for forecasting the actual level of interest rates - which in a sense is not surprising remembering that the Level-factor is close to being non-stationary. This observation makes the model in appropriate for the pricing of interest rate contingent claims

### 9. Conclusion

In this paper we have suggested a new 3-factor model that is a member of the linear Affine class of term structure models. We have specified the model as belonging to the Affine class because it makes the solution of the PDE (the system of ODEs) relatively straightforward even when no closed form solution is available - which happened to be the case for our specification.

As the model belongs to the Affine class of yield-curve models - it is a spot-rate model - the spot-rate is more precisely a linear combination of the state-variables. Our 3-state-variables have been selected as:

- The spot-rate - which is assumed to follow an SDE equal to the extended CIR-model from Duffie and Kan (1996)
- The level in dependent part of the diffusion-term in the process for the spot-rate is assumed to follow a square-root process
- The market price of risk is stochastic - and is assumed to follow an Ornstein-Uhlenbeck process

From this it follows that the volatility in the spot-rate consist of two independent parts. The first part is purely stochastic and is incorporated for capturing the persistence in volatility. The second part takes on the other hand care of the level-effect in the process for the spot-rate.

The assumption of a stochastic market price of risk in the process for the spot-rate has the following implications:

- First, we have removed the interactions of parameters in the drift-specification
- Secondly, we are now able to observe the mean-reversion parameter directly

Of course, assuming that the market price of risk can be both positive and negative introduces a positive probability of obtaining negative interest rates - but no problems were encountered on that account.

The model was estimated by using the yield-factor approach - that is we converted the state-variables in such a way that these yield-factors were observed without measurement error. The yield-factors we selected were:

- The short-rate represented by the 1-month rate (the Short-factor)
- The Slope-factor - the spread between the long-rate (15-year rate) and the short-rate
- The Butterfly-factor, defined as 2 times the middle-rate (7.5-year rate) minus the sum of the long-rate and the short-rate

## - The case of stochastic spot-rate, market price of risk and volatility

Mapping the state-variables into observed state-variables (yield-factors) has the following important implication from an estimation point of view:

- In this case we can derive the exact maximum likelihood function. The estimation procedure is then in principle identical to the estimation procedure which in general is used in connection with GARCH-models, see Greene (1993, section 19.7.1)

Compared to the Kalman filter algorithm the yield-factor approach is just as efficient seen from a computational perspective and the estimator is more efficient. The reason for this is that for general specification of Affine term structure models the QML-estimator is not consistent - it will actually only be consistent if the transition equation is treated as though it were Gaussian.

Using the yield-factor approach allows us to use the information across the whole yield-curve when estimating the parameters that specify the SDEs - while at the same time keeping the computational burden at a reasonable level. In this connection it is worth emphasizing that this actually is the first time in the literature that stochastic volatility models for the dynamic in the yield-curve have been estimated using information from the whole yield-curve - and not only for the short-rate.

The estimation of the model was performed using Danish bond-data (9 different maturity dates) for the period 2 January 1990 - 30 July 1998 and our estimation results can be summarized as:

- The spot-rate has a strong degree of mean-reversion
- The process for the market-price of risk is close to being non-stationary, as the mean-reversion parameter is fairly small
- The unconditional mean in the process for the market-price of risk is negative
- The standard-errors are much smaller than the estimated parameters (at least 5 times smaller) and the parameters are all significant

We also tested the model generated interest rate series against the actual (observed) interest rate series and found for all maturity dates that we could not reject the hypothesis that both samples had been randomly selected from the same distribution. Furthermore this was more pronounced for longer maturity dates than for shorter maturity dates.

We performed an analysis of the implied state-variables - that is the state-variables that arise when converting the model into a yield-factor model, and can report the following observation in that connection:

- The implied spot-rate was almost identical with the 1-month yield (yield-factor 1) - both in level and dynamic
- The implied dynamic in the volatility was almost identical to the difference

## A 3-Factor Model for the Yield-Curve Dynamics

---

- between the Butterfly-factor and the Slope-factor  
The implied market price of risk was clearly a Level-factor

From this we concluded:

- The dynamic in the yield-curve is approximately captured by a 2-factor model - but an additional state-variable is needed to capture the level of yields.

From the literature it is generally reported that there are three dominant factors that govern the dynamic in the term structure of interest rates, see among others Litterman and Scheinkman (1991). These are a Short-factor (sometimes called a Steepness-factor), a Level-factor and a Curvature-factor. This also turned out to be the case for our 3-factor model.

Finally we considered the prediction of yield-curve movements given knowledge about the state-variables (the yield-factors). From this analysis we can report the following:

- Given the state-variables (yield-factors) - the model seems to be able to predict the direction of interest rates ( $R^2 > 95\%$ ). This indicates that the model could be appropriate as a hedging tool - that is the factor-loadings derived are interesting from a risk-management perspective.
- The model is however inadequate for forecasting the actual level of interest rates - which in a sense is not surprising remembering that the Level-factor is close to being non-stationary. This observation makes the model inappropriate for the pricing of interest rate contingent claims

There is however an indication of the existence of a non-stationary component in interest rate data - which has been overlooked/disregarded in the literature. Firstly, it is generally reported in the literature for multi-factor yield-curve models that the Level-factor turns out close to being non-stationary, see among others Zheng (1993) and Chen and Scott (1995). Secondly, as pointed out in Madsen (1998) then a non-stationary component in interest rate data will - using PCA - give rise to a dominant factor which is a Level-factor.

Further analysis on multi-factor yield-curve models are of course required but we think the analysis performed gives some new insight into the dynamics in the term structure.

The model developed here and its possible use as a practical hedging tool needs to be more deeply analysed. The results were however encouraging.

Still other representations for the dynamic could be considered. We believe however that a 3-factor model of the Affine class still might be a feasible solution - models in that framework and their empirical implications are namely still in their early stages, so it is probably - as the results here indicate - too early to dismiss them.

When I point out that we probably only need a 3-factor model - I mean that the dynamic

## - The case of stochastic spot-rate, market price of risk and volatility

should be driven by just 3 Brownian motions. A more feasible way to proceed instead of adding new Brownian motions is probably to add more state-variables. This is an interesting line of research which has not been pursued in the literature.

The reason why this could be interesting is that the analysis in this paper shows that it is possible to get a very good description/fit of the historical evolution in the yield-curve. However the flexibility in the possible shapes of the yield-curve which can be derived from these kinds of models is probably too rich. Because of the high correlation between interest rates the observed yield-curves are usually limited to shapes with a maximum of one hump - which usually occur in the short end. A feasible solution to this observation could be to tie the yield-curve together at more points - but further research is needed to address that issue.

It is also important to decide what our ultimate goal is when specifying a multi-factor model:

- Is it to be used for risk-management?
- Is it to be used for the pricing of interest rate contingent claims?

Clearly the second criteria requires more from the model than the first criteria. Most importantly the second criteria cannot be obtained if we have a Level-factor that is close to being non-stationary.

## A 3-Factor Model for the Yield-Curve Dynamics

---

### References:

Ait-Sahalia (1995) "Testing Continuous-Time Models of the Spot Interest Rate", working paper University of Chicago, July 1995

Andersen and Lund (1997) "Estimating Continuous-Time Stochastic Volatility Models of the Short-Term Interest Rate", *Journal of Econometrics*, vol. 77, page 343-377

Andersen and Lund (1997) "Stochastic Volatility and Mean Drift in the Short-Rate Diffusion: Sources of Steepness, Level and Curvature in the Yield Curve", working paper Northwestern University and Aarhus School of Business, February 1997

Babbel (1988) "Interest Rate Dynamics and The Term Structure, A Note", *Journal of Banking and Finance* 12, page 401-417

Balduzzi, Das and Foresi (1997) "The Central Tendency: A Second Factor in Bond Yields", NBER, working paper 6325, December 1997

Barone, Cuoco and Zautnik (1991) "Term Structure Estimation using the Cox, Ingersoll and Ross Model: The Case of Italian Treasury Bonds", *Journal of Fixed Income* 1, page 87-95

Beaglehole and Tenney (1991) "General Solutions of some Interest Rate-Contingent Claim Pricing Equations", *Journal of Fixed Income* 3, page 69-83

Beaglehole and Tenney (1992) "Corrections and additions to "A nonlinear equilibrium model of the term structure of interest rates", *Journal of Financial Economics* 32, page 345-353

Bhar and Chiarella (1995 no. 53) "Transformation of Heath, Jarrow-Morton Models to Markovian Systems" working paper University of Technology, Sydney, December 1995

Bollerslev (1986) "Generalized Autoregressive Conditional Heteroskedasticity", *Journal of Econometrics* 31, page 307-327

Bollerslev and Wooldridge (1992) "Quasi-Maximum Likelihood Estimation and Inference in Dynamics Models with Time-Varying Covariances", *Econometric Reviews* 11, page 143-172

Brailsford and Maheswaran (1997) "The Dynamics of the Australian Short-Term Interest Rate", working paper November 1997

Berndt, Hall, Hall and Hausman (1974) "Estimation and Inference in Nonlinear Structural Models", *Annals of Economic and Social Management* 3/4, page 653-665

Brennan and Schwartz (1979) "A Continuous Time Approach to The Pricing of Bonds", *Journal of Banking and Finance* 3, page 133-155

## - The case of stochastic spot-rate, market price of risk and volatility

Brennan and Schwartz (1980) "Conditional Predictions of Bond Prices and Returns", *Journal of Finance*, vol 35, no. 2, page 405-419

Brennan and Schwartz (1982) "An Equilibrium Model of Bond Pricing and a Test of Market Efficiency", *Journal of Financial and Quantitative Analysis* vol 17, no. 3, page 301-329

Brenner, Harjes and Kroner (1993) "Another Look at Alternative Models of the Short-Term Interest Rate", working paper University of Arizona, September 1993

Brown and Dybvig (1986) "The Empirical Implications of the Cox, Ingersoll, Ross Theory of the Term Structure of Interest Rates", *Journal of Finance* vol 41, no. 3, page 617-632

Brown and Schaefer (1994) "The Term Structure of Real Interest Rates and the Cox, Ingersoll and Ross Model", *Journal of Financial Economics* 35, page 3-42

Buraschi (1994) "The Nominal Term Structure, Stock Prices and Consumption Dynamics with Nonneutral Inflation", working paper University of Chicago, October 1994

Caverhill and Strickland (1992) "Money Market Term Structure Dynamics and Volatility Expectations", working paper University of Warwick, June 1992

Chan, Karolyi, Longstaff and Sanders (1992a) "An Empirical Comparison of Alternative Models of the Short-Term Interest Models", *Journal of Finance* vol 47, no. 3, page 1209-1227

Chan, Karolyi, Longstaff and Sanders (1992b) "The Volatility of Japanese Interest Rates: A Comparison of Alternative Term Structure Models", *Pacific-Basin Capital Markets Research*, vol 2, page 119-136

Chen (1995) "A Three-Factor Model of the Term Structure of Interest Rates", working paper Federal Reserve Board, July 1995

Chen and Scott (1993) "Maximum Likelihood Estimation for a Multifactor Equilibrium Model of the Term Structure of Interest Rates", *Journal of Fixed Income* vol 3, no. 3, page 14-32

Chen and Scott (1995) "Multi-Factor Cox-Ingersoll-Ross Models of the Term Structure: Estimates and Tests From a Kalman Filter Model", working paper Rutgers University and University of Georgia, January 1995

Constantinides (1992) "A Theory of the Nominal Term Structure of Interest Rates", *The Review of Financial Studies* 5, page 531-552

Cox, Ingersoll and Ross (1985) "A Theory of the Term Structure of Interest Rates", *Econometrica* 53, page 363-385

## A 3-Factor Model for the Yield-Curve Dynamics

---

Dahl (1989) "Variations in the term structure and controlling interest-rate risk exposures" (in Danish), Finans Invest, February 1989

Dai and Singleton (1997) "Specification Analysis of Affine Term Structure Models", NBER, working paper 6128, August 1997

Dietrich-Campbell and Schwartz (1986) "Valuing Debt Options: Empirical Evidence", *Journal of Financial Economics* 16, page 321-343

Duan and Simonato (1995) "Estimating and Testing Exponential-Affine Term Structure Models by Kalman Filter", working paper CIRANO, 95s-44, October 1995

Duffie and Kan (1996) "A Yield-Factor Model of Interest Rates", Working Paper Stanford University

Duffie and Singleton (1996) "An Econometric Model of the Term Structure of Interest Rate Swap Yields", working paper Stanford University and NBER, August 1996

Engle (1982) "Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation", *Econometrica* vol 50, no. 4, page 987-1007

Fischer and Zechner (1984) "Diffusion Process Specifications for Interest Rates", *Lecture Notes in Economics and Mathematical Systems*, vol 227, page 64-73

Gallant and Tauchen (1996) "Which Moments to Match?", *Econometric Theory* 12, page 657-681

Garbade (1986) "Modes of Fluctuations in Bonds Yields - an Analysis of Principal Components", Bankers Trust, Money Market, June 1986

Geyer and Pichler (1995) "A state-space approach to estimate the term structure of interest rates: some empirical evidence", Paper presented at the EFA meeting - Milano 1995

Giannopoulos (1995) "The Theory and Practice of ARCH Models", working paper University of Westminster, November 1995

Golub and Van Loan (1993) "Matrix Computations", The John Hopkins University Press, second edition 1993

Greene (1993) "Econometric Analysis", 2nd Edition, Macmillan Publishing Company 1993

Gourieroux, Monfort and Renault (1993) "Indirect Inference", *Journal of Applied Econometrics* 8, page 85-118.

Hald (1979) "Updating formulas in Quasi-Newton methods" (in Danish), Department for

## - The case of stochastic spot-rate, market price of risk and volatility

Numerical Analysis, report no. NI-79-03

Harvey (1990) "The Econometric Analysis of Time Series" Philip Allan Publishing Company 1990

Harvey (1992) "Forecasting, Structural Time Series Models and the Kalman Filter", Cambridge University Press 1992

Heath, Jarrow and Morton (1991) "Bond Pricing and the Term Structure of Interest Rates: A new Methodology for Contingent Claim Valuation", working paper, Cornell University

Honore (1998) "Five Essays on Financial Econometrics in Continuous-Time Model", Ph.D Dissertation, The Aarhus School of Business, 20. March 1998

Hull and White (1994) "Numerical Procedures for Implementing Term Structure Models II: Two-Factor Models", Journal of Derivatives vol. 2, no. 2, page 37-48

Hørdahl (1997) "Financial Volatility and Time-Varying Risk Premia", Lund Economic Studies number 69, September 1997

Jackwerth (1996) "Recovering Risk Aversion from Option Prices and Realized Returns", working paper no. RPF-265, University of California at Berkeley, September 1995

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

Jamshidian (1990) "Bond and Option Evaluation in the Gaussian Interest Rate Model", Working Paper, Merrill Lynch

Jiang (1997) "Nonparametric Modelling of U.S. Interest Rate Term Structure Dynamics and Implications on the Prices of Derivative Securities", working paper University of Groningen, September 1997

Kanji (1995) "100 Statistical Tests", SAGE Publications 1995

Karatzas and Schreve (1988) "Brownian Motion and Stochastic Calculus", Springer-Verlag New York Inc.

Karoui and Lacoste (1992) "Multifactor Models of the term structure of interest rates", Working Paper University of Paris

Langetieg (1980) "A Multivariate Model of the Term Structure", Journal of Finance 35, page 71-97

Litterman and Scheinkman (1988) "Common factors affecting Bond Returns", Goldman,

## A 3-Factor Model for the Yield-Curve Dynamics

---

Sachs & Co., Financial Strategies Group, September 1988

Longstaff (1989) "A non-linear general equilibrium model of the term structure of interest rates", *Journal of Financial Economics* 23, page 195-224

Longstaff and Schwartz (1990) "Interest-rate volatility and the term structure: A two-factor general equilibrium model", Working Paper Ohio State University November 1990

Lund (1997) "Five Essays in Financial Econometrics and the Term Structure of Interest Rates", Ph.D. dissertation Aarhus School of Business, July 1997

Madsen (1994) "Estimating the Yield-Curve - an Implicit Volatility Approach" (in Danish), working paper Realkredit Danmark, August 1994

Madsen (1995) "The APT-model and variations in the Term Structure of Interest Rates" (in Danish), working paper Realkredit Danmark, October 1995

Madsen (1998) "Term-Structure Dynamics and the Determination of State-Variables - a Multi-Factor Approach", working paper Svenska Handelsbanken, May 1998

Marquardt (1963) "An algorithm for least squares estimation of nonlinear parameters", *Journal of Society of Industrial Applied Mathematics* 11, page 431-441

Marsh and Rosenfeld (1983) "Stochastic Processes for Interest Rates and Equilibrium Bond Prices", *Journal of Finance* vol 38, no. 2, page 635-650

Moler and Stewart (1973) "An Algorithm for Generalized Matrix Eigenvalue Problems", *SIAM Journal of Numerical Analysis* 10, page 241-256

Newey and West (1987) "A Simple Positive Semi-Definite Heteroskedasticity and Autocorrelation Consistent Covariance Matrix", *Econometrica* 55, page 703-708

Nowman (1997) "Gaussian Estimation of Single-Factor Continuous Time Models of the Term Structure of Interest Rates", *Journal of Finance* 52, page 1695-1706

Pearson and Sun (1994) "Exploiting the Conditional Density in Estimating the Term Structure: An Application to the Cox, Ingersoll and Ross Model", *Journal of Finance* 49, page 1279-1304

Press, Teukolovski, Vetterling and Flannery (1994) "Numerical Recipes in C", Cambridge University Press, second edition

Ruiz (1994) "Quasi-maximum likelihood estimation of stochastic volatility models", *Journal of Econometrics* 63, page 289-306

## - The case of stochastic spot-rate, market price of risk and volatility

Schwartz and Schaefer (1984) "A Two-Factor model of the Term Structure of Interest Rates: An Approximate Analytical Solution", *Journal of Financial and Quantitative Analysis* 19, page 413-424

Schwartz and Schaefer (1987) "Time-Dependent Variance and the Pricing of Bond Options", *Journal of Finance* vol 42, nr. 5, side 1113-1128

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

Singh (1995) "Estimation of Multifactor Cox, Ingersoll and Ross Term Structure Model: Evidence on Volatility Structure and Parameter Stability", *Journal of Fixed Income* vol. 5, no. 2, page 8-28

Tanizaki (1996) "Nonlinear Filters - Estimation and Applications", Springer-Verlag Berlin 1996

Uhrig and Walter (1994) "Term Structure of Interest Rates, Consistent Risk Attitudes, and Preference-Free Option Prices", Working Paper University of Mannheim July 1994

Van Loan (1978) "Computing Integrals Involving the Matrix Exponential", *IEEE Transactions on Automatic Control* 23, page 395-404

Vasicek (1977) "An Equilibrium Characterization of the Term Structure", *Journal of Financial Economics* 5, page 177-188

Vasicek and Fong (1991) "Interest Rate Volatility as a Stochastic Factor", working paper Gifford Fong Associates February 1991

Vetzal (1997) "Stochastic volatility, movements in short term interest rates, and bond option values", *Journal of Banking and Finance* 21 (1997), page 169-196

Zheng (1993) "Testing Term Structure Models of Interest Rates with Unspecified Dynamic Factors", working paper First National Bank of Chicago, December 1993

## Appendix A

Let us now consider the following augmented SDE:

$$\begin{aligned}
 dZ_t &= \begin{bmatrix} dR(t,T) \\ dX_t \end{bmatrix} = \begin{bmatrix} 0 \\ b \end{bmatrix} + \begin{bmatrix} 0 & \omega^T \\ 0 & B \end{bmatrix} \begin{bmatrix} R(t,T) \\ X_t \end{bmatrix} dt + \begin{bmatrix} 0 \\ S \end{bmatrix} d\tilde{W}_t \\
 \text{for} & \\
 dZ_t &= (c + KZ_t) + Td\tilde{W}_t \\
 &\text{and} \\
 b &= a - S\lambda^T
 \end{aligned} \tag{54}$$

where  $c$  is an  $m+1$  vector,  $K$  is an  $(m+1) \times (m+1)$  matrix and  $T$  is an  $(m+1) \times m$  matrix of diffusion coefficients. Furthermore  $Z_t$  is an  $(m+1)$  vector of state-variables  $[R(t,T), X_t]$ .

From equation 54 we can deduce that the first element in the vector of state-variables  $R(t,T)$

can be expressed as:  $R(t,T) = \omega^T \int_t^T x_s ds$  - that is  $R(t,T)$  is the  $T$ -period spot rate.

The process in formula 54 is normally distributed and the expected value and variance can be written as:

$$\begin{aligned}
 E[Z_T | Z_t] &= e^{-K(T-t)} Z_t + \int_t^T e^{-K(s-t)} c ds \\
 V[Z_T | Z_t] &= \int_t^T e^{-C(s-t)} T T^T e^{-C(s-t)^T} ds
 \end{aligned} \tag{55}$$

The conditional expected value and variance for  $R(t,T)$  can now be found as the first element in the vector  $E[Z_T | Z_t]$  and the (1,1) element in the variance matrix  $V[Z_T | Z_t]$ .

From this we can deduce that the expected value of  $R(t,T)$  ( $E[R(t,T) | X_t]$ ) can be expressed as:

- The case of stochastic spot-rate, market price of risk and volatility

$$\begin{aligned}
 E[R(t,T)|X_t] &= \omega^T \int_t^T \left[ e^{-B(s-t)} X_t + \int_v^s e^{-B(s-v)} b dv \right] ds \\
 &= \omega^T Q A^{-1} [I - e^{-A(T-t)}] Q^{-1} X_t + \omega^T \left[ - \int_t^T e^{-B(s-t)} b B^{-1} ds + \int_t^T b B^{-1} ds \right] \\
 &= \omega^T Q A^{-1} [I - e^{-A(T-t)}] [Q^{-1} X_t - Q^{-1} B^{-1} b] + \omega^T B^{-1} b (T-t)
 \end{aligned} \tag{56}$$

where A is a matrix with B's eigenvalues in the diagonal and Q is a matrix with B's eigenvectors in the columns. Formula 3 can be recognized as being identical to formula 32 in Langetieg (1980), except that the eigenvalue decomposition has been performed here.

The first term in line 55 in formula 56 can be expressed as (disregarding the Boolean vector  $\omega^T$ ):

$$\int_t^T e^{-B(s-t)} ds = B^{-1} [I - e^{-B(T-t)}] = Q A^{-1} [I - e^{-A(T-t)}] Q^{-1} \tag{57}$$

where the equality between the second and third term in formula 57 (under the assumption that B is non-singular) relies on the calculation rule for the matrix exponential, as it follows from here that if the matrix B can be expressed as  $Q A Q^{-1}$  - then we can express  $e^{B(\tau)}$  as  $Q e^{A(\tau)} Q^{-1}$ .

As the matrix exponential to a diagonal matrix is equal to the exponential of the elements in the diagonal it follows that an eigenvalue decomposition is an efficient way to calculate the matrix exponential.

There is however one case where the eigenvalue decomposition - for evaluating the integral of the matrix exponential - is not possible, and that is if B is non-singular. The reason for this is that a non-singular B matrix will produce one or more eigenvalues that are zero.

As mentioned by Beaglehole and Tenney (1991) it is of course more general to assume that one or more of the eigenvalues can be a complex number - an eigenvalue decomposition is however also possible in that case.

The variance for R(t,T) ( $V[R(t,T)|X_t]$ ) can be written as follows:

## A 3-Factor Model for the Yield-Curve Dynamics

---

$$\begin{aligned}
 V[R(t,T)|X_t] &= \omega^T \int_t^T QA^{-1}(I - e^{-A(s-t)})Q^{-1}SS^TQA^{-1}(I - e^{-A(s-t)})Q^{-1}\omega ds \\
 &= \omega^T \int_t^T QA^{-1}e^{-A(s-t)}Q^{-1}SS^TQA^{-1}e^{-A(s-t)}Q^{-1}\omega ds \\
 &\quad - \omega^T \int_t^T QA^{-1}e^{-A(s-t)}Q^{-1}SS^TQA^{-1}Q^{-1}\omega ds \\
 &\quad - \omega^T \int_t^T QA^{-1}Q^{-1}SS^TQA^{-1}e^{-A(s-t)}Q^{-1}\omega ds \\
 &\quad + \omega^T \int_t^T QA^{-1}Q^{-1}SS^TQA^{-1}Q^{-1}\omega ds
 \end{aligned} \tag{58}$$

which can be written as:

$$\begin{aligned}
 V[R(t,T)|X_t] &= \omega^T \int_t^T QA^{-1}e^{-A(s-t)}Q^{-1}SS^TQA^{-1}e^{-A(s-t)}Q^{-1}\omega ds \\
 &\quad - \omega^T [QA^{-1}A^{-1}e^{-A(T-t)}Q^{-1} - (BB)^{-1}]SS^T(B^{-1})^T\omega \\
 &\quad - \omega^T B^{-1}SS^T[QA^{-1}A^{-1}e^{-A(T-t)}Q^{-1} - (BB)^{-1}]^T\omega \\
 &\quad + \omega^T B^{-1}SS^T(B^{-1})^T\omega(T-t)
 \end{aligned} \tag{59}$$

Where formula 59 can be seen to be identical to equation 33 in Langetieg (1980). The only thing that needs to be worked out is the integral in line 1, Langetieg (1980 footnote 23) suggests an eigenvalue decomposition in order to solve the integral in line 1 - this will however not be pursued here. Instead I will leave the integral in abstract form.

As we have that the price of a zero-coupon bond can be expressed in term of the R(t,T) spot rate we have:

$$P(t,T) = e^{\left[ -E[R(t,T)|X_t] + \frac{1}{2}V[R(t,T)|X_t] \right]} \tag{60}$$

We have now established a semi-analytical expression for the bond-price in a multi-dimensional Gaussian term-structure model (given B is non-singular) - where correlation between the Wiener processes is obtained when  $SS^T$  is not a diagonal matrix.

### Appendix B

Given equation 28 and 30 in the main text we can rewrite formula 29 as:

$$\begin{aligned}\frac{dD(\tau)}{d\tau} &= \frac{1}{2} \sum_{i=1}^m D_i(\tau)^2 \beta_i - BD(\tau)^T - \sum_{i=1}^m \lambda_i D_i(\tau) \beta_i - \omega \\ \frac{dC(\tau)}{d\tau} &= \frac{1}{2} \sum_{i=1}^m D_i(\tau)^2 \alpha_i + D(\tau)^T a - \sum_{i=1}^m \lambda_i D_i(\tau) \alpha_i\end{aligned}\quad (61)$$

As B is a diagonal matrix this can be expressed as:

$$\begin{aligned}\frac{dD_i(\tau)}{d\tau} &= \frac{1}{2} D_i(\tau)^2 \beta_i - \kappa_i D_i(\tau) - \lambda_i D_i(\tau) \beta_i - 1 \quad \text{for } i = [1, 2, \dots, m] \\ \frac{dC_i(\tau)}{d\tau} &= \frac{1}{2} D_i(\tau)^2 \alpha_i + D_i(\tau) \alpha_i - \lambda_i D_i(\tau) \alpha_i \quad \text{for } i = [1, 2, \dots, m]\end{aligned}\quad (62)$$

That is the solution to the PDE can be expressed as 2m ODEs.

Because of the diagonality for the mean-reversion matrix B we only need to figure out the solution for the i'th process - I will therefore in the derivation of the solution for  $C_i(\tau)$  and  $D_i(\tau)$  disregard the subscript.

Which respect to  $D$ <sup>84</sup> we need to solve the following integral:

$$\int_0^\tau \frac{1}{\left[ \frac{1}{2} \beta D^2 - D[\kappa + \lambda \beta] - 1 \right]} dD = \tau \quad (63)$$

From this we can see that there exist three different solutions depending on the nature of the roots of the polynomial in the denominator.

We will however only pursue the case  $[\kappa + \lambda \beta]^2 > 2\beta$   
rewrite 63 as:

<sup>85</sup>. Under this assumption we can

---

<sup>84</sup> We will use the notation C, D and C(τ), D(τ) whenever applicable.

<sup>85</sup> This is also the only case that has been pursued in the literature.

## A 3-Factor Model for the Yield-Curve Dynamics

---

$$\tau = \frac{1}{\gamma} \ln \left[ \frac{r_2 [D - r_1]}{r_1 [D - r_2]} \right]$$

*where*

$$r_1 = \frac{-d + \gamma}{\beta}$$

$$r_2 = \frac{-d - \gamma}{\beta}$$

*and*

$$d = \kappa + \lambda\beta$$

$$\gamma = \sqrt{d^2 + 2\beta}$$
(64)

That is D can be written as:

$$\tau = \frac{1}{\gamma} \ln \left[ \frac{\frac{-d - \gamma}{\beta} \left[ D - \frac{d - \gamma}{\beta} \right]}{\frac{-d + \gamma}{\beta} \left[ D + \frac{d + \gamma}{\beta} \right]} \right]$$

$$\Rightarrow D(\tau) = - \frac{2[e^{\gamma\tau} - 1]}{[\gamma + \kappa + \lambda\beta]e^{\gamma\tau} + [\gamma - \kappa - \lambda\beta]}$$
(65)

We are now ready to solve the second ODE - that is, figure out the analytical expression for C( $\tau$ ).

With respect to C we need to evaluate the following integral-equation:

$$C(\tau) = \frac{1}{2} \alpha \int_0^\tau D(s)^2 ds + [a - \lambda\alpha] \int_0^\tau D(s) ds$$
(66)

The second integral in formula 66 can be evaluated analytically using the substitutions-rule which yields:

$$\int_0^\tau D(s) ds = \frac{2}{\beta} \ln \left[ \frac{2\gamma e^{\frac{1}{2}[\gamma + \kappa + \lambda\beta]\tau}}{[\gamma + \kappa + \lambda\beta]e^{\gamma\tau} + [\gamma - \kappa - \lambda\beta]} \right]$$

$$= \ln \left[ \frac{2\gamma e^{\frac{1}{2}[\gamma + \kappa + \lambda\beta]\tau}}{[\gamma + \kappa + \lambda\beta]e^{\gamma\tau} + [\gamma - \kappa - \lambda\beta]} \right]^{\frac{2}{\beta}}$$
(67)

## - The case of stochastic spot-rate, market price of risk and volatility

---

In order to facilitate a solution to the first integral in formula 66 we rewrite it as:

$$\int_0^{\tau} D(s)^2 ds = r_1^2 \left[ \int_0^{\tau} \left[ 1 + \frac{4\gamma}{\beta[r_2 e^{\gamma s} - r_1]} \right] ds + \int_0^{\tau} \frac{4\gamma^2}{\beta^2[r_2 e^{\gamma s} - r_1]^2} ds \right] \quad (68)$$

The solution to the first integral in equation 68 can be expressed as:

$$r_1^2 \int_0^{\tau} \left[ 1 + \frac{4\gamma}{\beta[r_2 e^{\gamma s} - r_1]} \right] ds = \frac{r_1[\beta r_1 - 4\gamma]}{\beta} \tau - \frac{4r_1}{\beta} \ln \left[ \frac{r_2 - r_1}{r_2 e^{\gamma \tau} - r_1} \right] \quad (69)$$

In order to figure out a solution to the last integral in formula 68, we do the following.

First we have that:

$$\begin{aligned} g(\tau)^2 &= [r_2 e^{\gamma \tau} - r_1]^2 = r_1^2 - r_2 e^{\gamma \tau} [r_1 - r_2 e^{\gamma \tau}] - r_1 r_2 e^{\gamma \tau} \\ &\quad \text{where} \\ g(\tau) &= r_2 e^{\gamma \tau} - r_1 \end{aligned} \quad (70)$$

which means:

$$\begin{aligned} \frac{1}{g(\tau)^2} &= \frac{g(\tau)^2 + r_2 e^{\gamma \tau} [r_1 - r_2 e^{\gamma \tau}] + r_1 r_2 e^{\gamma \tau}}{r_1^2 g(\tau)^2} \\ &= \frac{1}{r_1^2} - \frac{g'(\tau)}{\gamma r_1^2 g(\tau)^2} - \frac{g'(\tau)}{\gamma r_1 g(\tau)^2} \end{aligned} \quad (71)$$

This expression can be integrated straight away. Firstly, the second term is the differential coefficient of a constant times  $\ln[g(\tau)]$ , secondly, the third term is the differential coefficient of a constant times  $\frac{1}{g(\tau)}$ .

From this we deduce that the second term in formula 68 can be written as:

$$\left( \frac{2\gamma r_1}{\beta} \right)^2 \left[ \frac{\tau}{r_1^2} - \frac{1}{\gamma r_1 g(\tau)} - \frac{1}{\gamma r_1 [r_1 - r_2]} + \frac{1}{\gamma r_1^2} \ln \left[ \frac{r_2 - r_1}{g(\tau)} \right] \right] \quad (72)$$

## A 3-Factor Model for the Yield-Curve Dynamics

---

Formula 31 in the main text now follows directly from formula 65, 67, 69 and 72.

A few things are worth emphasising here.

### Case $\alpha = 0$

In this case the specification degenerates into the CIR model, and from here it follows that  $D(\tau)$  is given by formula 65 and  $C(\tau)$  by formula 67.

### Case $\beta = 0$

In this case the specification degenerates into the Vasicek model - this has the implication that  $D(\tau)$  is found through a simplified version of formula 65.

More precisely we have that  $D(\tau)$  is defined by the following ODE:

$$\frac{dD(\tau)}{d\tau} = -\kappa D(\tau) - 1 \quad (73)$$

where the solution to this ODE is given in formula 19 in the main text. The solution to  $C(\tau)$  is given by:

$$\frac{dC(\tau)}{d\tau} = \frac{1}{2}D(\tau)^2\alpha + D(\tau)a - \lambda D(\tau)\sqrt{\alpha} \quad (74)$$

The solution to these ODE's (with respect to equation 73) is given in formula 19 in the main text for  $\alpha = \sigma^2$  and  $a = \kappa\theta^{86}$ .

### Case $a = 0$

In this case we still have that  $D(\tau)$  is given by formula 65. But now  $C(\tau)$  is determined completely by equation 69 and 72 - though taking into account that  $a = 0$ .

### Case $\kappa = 0$

In this case  $D(\tau)$  is given by a simplified version of the ODE in equation 62, ie:

$$\frac{dD(\tau)}{d\tau} = \frac{1}{2}D(\tau)^2\beta - 1 \quad (75)$$

This can be recognized as being a special case for the solution given in formula 65. The expression for  $D(\tau)$  can easily be derived from equation 75, using the technique applied for deriving formula 65 - this is left for the reader.

Other combinations can of course also be recognized as being incorporated in this fairly general formulation, for example  $a = 0$  and  $\beta = 0$  - this results in a Gaussian model with the

---

<sup>86</sup> Here we have expressed the market price of risk as in the Vasicek model.

## - The case of stochastic spot-rate, market price of risk and volatility

---

unconditional mean equal to 0 (zero).

It is however sufficient to mention that the derivation here incorporates most of the models examined in the literature.

### Appendix C

In a Gaussian yield-curve model we have that the dynamics of the vector of state-variables  $X_t$  is governed by the following SDE:

$$dX_t = (a + BX_t)dt + SdW_t \quad (76)$$

where  $a$  is an  $m \times 1$  vector,  $B$  is an  $m \times m$  matrix. In the general case we have that the volatility matrix is found by Cholesky factorization of the variance-covariance matrix and that the  $m$ -Brownian motions are correlated. Furthermore we assume that the spot rate is given by a linear combination of the state-variables, ie:  $r_t = \omega^T X_t$ .

From section 3.1 in the main text we have that the conditional mean and conditional covariance matrix for the model in formula 76 can be expressed as follows:

$$\begin{aligned} E[X_s|X_t] &= e^{-B(s-t)}X_t + \int_t^s e^{-B(s-v)}a dv \\ V[X_s|X_t] &= \int_t^s e^{-B(s-v)}SS^T e^{-B^T(s-v)} dv \end{aligned} \quad (77)$$

## A 3-Factor Model for the Yield-Curve Dynamics

---

Which can be evaluated analytically as shown in section 3.1.

Let us now assume that we wish to estimate the unknown parameter vector - which we will denote  $\psi$  - that governs the dynamic of the dimensional SDE from formula 76. For that purpose we will use the yields on a unique set of  $N$  zero-coupon bonds. The observed yields at time  $t_x$ , for  $x = 1, 2, \dots, L$ , are denoted  $R(x, \tau)$  - that is  $R(x, \tau)$  is an  $N$ -dimensional vector of yields on zero-coupon bonds with a maturity date equal to  $\tau$ , for  $\tau = \tau_1, \tau_2, \dots, \tau_N$ . We also assume that  $N > m$  - that is there are more observations for each  $x$  than the number of state-variables.

### C.1 The state space form

Under the usual assumption that the measurement errors are additive and normally distributed the measurement equation has the following form:

$$R(x, \tau) = d(\psi) + Z(\psi)X_x + \varepsilon_x$$

*where*

$$\varepsilon_x \approx N(0, H(\psi)) \tag{78}$$

where the  $i$ 'th row of the matrices  $d$  ( $N \times 1$ ) and  $Z$  ( $N \times m$ ) are given by  $-\frac{A(\tau_i)}{\tau_i}$  and  $-\frac{B(\tau_i)^T}{\tau_i}$

- and where  $A(\tau)$  and  $B(\tau)$  can be derived from equation 24 in the main text.

As we are estimating the continuous time model using discrete observations the transition equation has to be derived from the exact discrete-time distribution of the state variables. In the Gaussian case, this distribution follows immediately from the solution to the SDE, ie. equation 77.

More precisely we have from equation 77 that the transition equation is a first order Markov process with Gaussian innovations, ie:

$$X_x = c_x(\psi) + \Phi_x(\psi)X_{x-1} + \eta_x(\psi)$$

*for*

$$c_x(\psi) = \int_{t_{x-1}}^{t_x} e^{-B(t_x - s)} a ds$$

$$\Phi_x(\psi) = e^{-B(t_x - t_{x-1})}$$

$$\eta_x(\psi) \approx N(0, V_x(\psi))$$

$$V_x(\psi) = \int_{t_{x-1}}^{t_x} e^{-B(t_x - s)} S S^T e^{-B(t_x - s)^T} ds \tag{79}$$

where we have that the matrices  $c_x(\psi)$ ,  $\Phi_x(\psi)$  and  $\eta_x(\psi)$  are time-varying unless  $\delta = t_x - t_{x-1}$

## - The case of stochastic spot-rate, market price of risk and volatility

---

is a constant for all  $x$ . The relations derived here will however be formulated for the general case - ie  $t_x - t_{x-1}$  is not equal for all  $x$ .

### C.2 The linear Kalman Filter

As the model represented by formula 78 and 79 is Gaussian, see Harvey (1993 section 3.7.1), then the estimates of the state variables will be optimal in the MSE (mean square error) sense.

In the following, we will call  $\hat{X}_{x|x-1}$  and  $\hat{X}_x$  respectively the optimal estimator for the vector of state variables based on information up to and at the points in time  $t_{x-1}$  and  $t_x$ . The optimal estimator is given by the conditional mean of  $X_x$ , and will be denoted  $E_{x-1}[\cdot]$  and  $E_x[\cdot]$  and

The prediction step is given by:

$$\hat{X}_{x|x-1} = E_{x-1}[X_x] = c_x(\psi) + \Phi_x(\psi)\hat{X}_{x-1} \quad (80)$$

which has an MSE matrix identical with:

$$\begin{aligned} \Sigma_{x|x-1} &= E_{x-1}[(X_x - \hat{X}_{x|x-1})(X_x - \hat{X}_{x|x-1})^T] \\ &= \Phi_x(\psi)\Sigma_{x-1}\Phi_x^T(\psi) + V_x(\psi) \end{aligned} \quad (81)$$

In the update step the additional information given by  $R(x,\tau)$  is used in order to achieve a more precise estimate for  $X_x$ , so that:

$$\begin{aligned} \hat{X}_x &= E_x[X_x] = \hat{X}_{x|x-1} + \Sigma_{x|x-1}Z(\psi)^TF_x^{-1}v_x \\ \Sigma_x &= \Sigma_{x|x-1} - \Sigma_{x|x-1}Z(\psi)^TF_x^{-1}Z(\psi)\Sigma_{x|x-1} = [\Sigma_{x|x-1}^{-1} + Z(\psi)^TH^{-1}Z(\psi)]^{-1} \\ &\text{for} \\ v_x &= R(x,\tau) - d_x(\psi) - Z(x,\psi)\hat{X}_{x|x-1} \\ F_x &= Z(\psi)\Sigma_{x|x-1}Z(\psi)^T + H \end{aligned} \quad (82)$$

where this new estimate of  $X_x$  is called the filtered estimate.

The main purpose of the Kalman filter is to extract information about  $X_x$  from the observed zero-coupon rates. The Kalman filter algorithm can however also be used to calculate the exact likelihood function by using the prediction error decomposition technique.

If we disregard a constant, the log-likelihood function is given by:

$$\log_L(R(1,\tau),R(2,\tau),\dots,R(L,\tau);\psi) = -\frac{1}{2}\sum_{x=1}^L \text{Log}|F_x| - \frac{1}{2}\sum_{x=1}^L v_x^TF_x^{-1}v_x \quad (83)$$

## A 3-Factor Model for the Yield-Curve Dynamics

---

In order to save calculation time, the determinant and the inverse of  $F_x$  can be calculated more efficiently using the following expression (see Harvey (1993) page 108):

$$\begin{aligned}
 F_x^{-1} &= H^{-1} - H^{-1}Z(\psi)\Sigma_x Z(\psi)^T H^{-1} \\
 &\text{and} \\
 |F_x| &= |H| |\Sigma_{xx}^{-1}| |\Sigma_x^{-1}|
 \end{aligned}
 \tag{84}$$

This saves time particularly when  $L$  is (much) larger than  $m$ , which is generally the case.

In order to start the Kalman filter algorithm we need an initial estimate of the vector of state variables  $X_0$  and its MSE  $\Sigma_0$ . If the process for the state variable is stationary, its unconditional mean and covarians matrix are used, which can be found, respectively, by letting  $t_{-1} \rightarrow -\infty$  in the expression for  $[\Phi_x(\psi), c_x(\psi)]$  and  $V_x(\psi)$  in formula 79, see also Harvey (1993) page 121. If, however, some of the state variables are not stationary, the diffuse-prior technique can be used, see Harvey (1993) section 3.4.3.

Appendix D

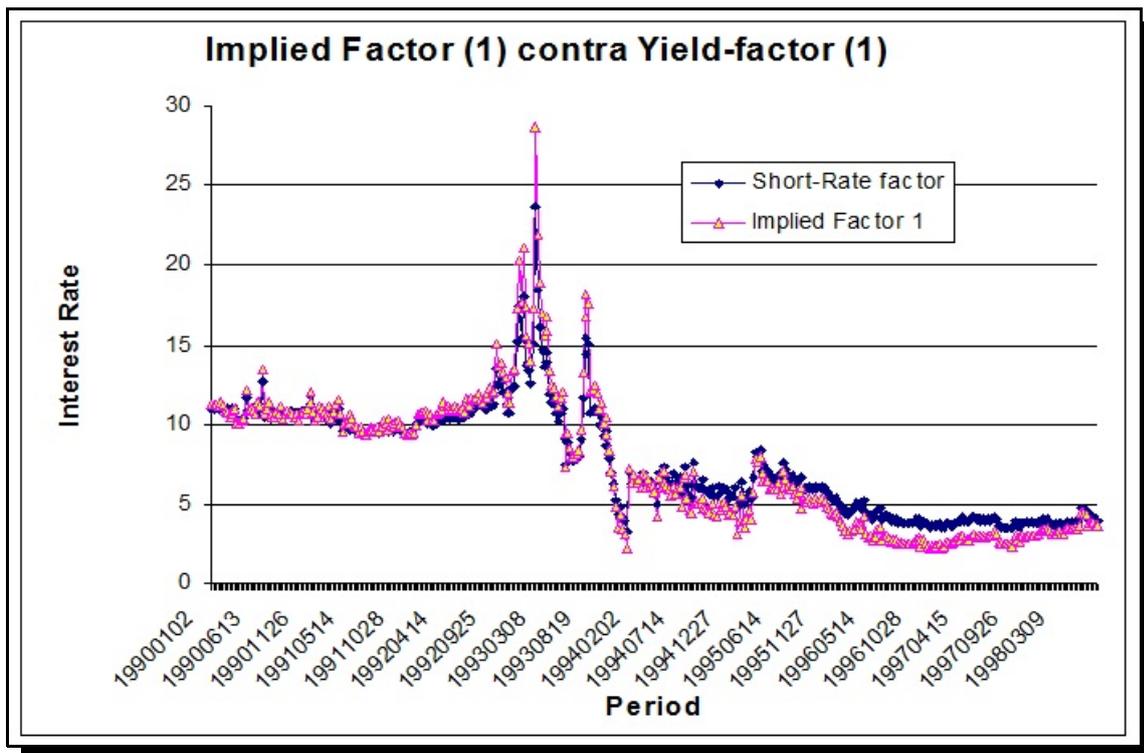


Figure 1

## A 3-Factor Model for the Yield-Curve Dynamics

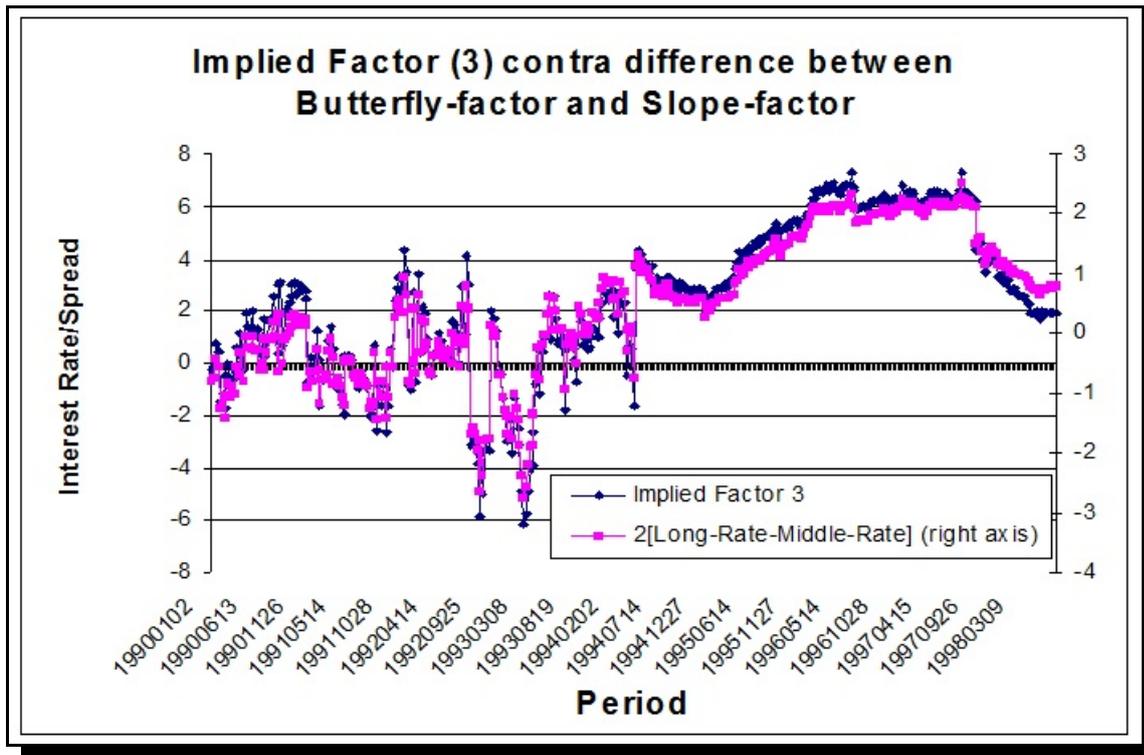


Figure 2

### Appendix E

Normal Distribution tests for the observed Interest Rate Series  
(period 2 January 1990 - 30 June 1998)

|                     | $\chi^2$ | Degree of Freedom | Significance | $\chi^2$ | Degree of Freedom | Significance |
|---------------------|----------|-------------------|--------------|----------|-------------------|--------------|
| <b>3-Month Rate</b> | 3592,76  | 10                | 0,3876       | 4796,42  | 20                | 0,0000       |
| <b>6-Month Rate</b> | 3727,12  | 10                | 0,4846       | 5454,05  | 20                | 0,0000       |
| <b>1-Year Rate</b>  | 4073,24  | 10                | 0,6789       | 6406,61  | 20                | 0,4050       |
| <b>2-Year Rate</b>  | 3409,44  | 10                | 0,2339       | 5307,60  | 20                | 0,0000       |
| <b>3-Year Rate</b>  | 3355,40  | 10                | 0,1839       | 4560,71  | 20                | 0,0000       |
| <b>4-Year Rate</b>  | 3490,52  | 10                | 0,3050       | 4344,11  | 20                | 0,0000       |
| <b>5-Year Rate</b>  | 3819,50  | 10                | 0,5440       | 4517,51  | 20                | 0,0000       |
| <b>10-Year Rate</b> | 4192,33  | 10                | 0,7295       | 4821,16  | 20                | 0,0000       |

- The case of stochastic spot-rate, market price of risk and volatility

---

|                       |         |    |        |         |    |        |
|-----------------------|---------|----|--------|---------|----|--------|
| <b>12.5-Year Rate</b> | 4287,85 | 10 | 0,7649 | 4555,53 | 20 | 0,0000 |
|-----------------------|---------|----|--------|---------|----|--------|

### Appendix F

In order to further test the prediction power of the model we have constructed the following test table:

|   | <b>Rates move Up</b>                 | <b>Rates move Down</b>                 | <b>Total</b>      |
|---|--------------------------------------|--|-------------------|
| <b>Model predict Rates to move Up</b>   | Model predict UP and Actual UP (a)   | Model Predict UP and Actual DOWN (b)   | a + b             |
| <b>Model predict Rates to move Down</b> | Model predict DOWN and Actual UP (c) | Model Predict DOWN and Actual DOWN (d) | c + d             |
| <b>Total</b>                            | a + c                                | b + d                                  | n = a + b + c + d |

The test statistic is:

$$\chi^2 = \frac{(n - 1)[ad - bc]^2}{(a + b)(a + c)(b + d)(c + d)} \quad (85)$$

## A 3-Factor Model for the Yield-Curve Dynamics

---

Where this is to be compared with the theoretical value of  $\chi^2$  with 1 degree of freedom. If  $\chi^2$  exceeds the critical value the null hypothesis of independence between samples and classes is rejected. From this we can conclude that if  $\chi^2$  calculated from formula 1 is higher than the theoretical value - then we have a high degree of correlation between the two-samples. It is easily seen from the table that this is obtained if a and d are "much" larger than c and b.

The theoretical value at the 5% confidence level with 1 degree of freedom is 3.84 - which mean that for values of  $\chi^2$  calculated using formula 1 with exceeds 3.84, then we cannot reject that the predicted interest rate data are highly correlated with the actual interest data.

Prediction power tests for the correlation between  
the Model-Generated Interest Rate Series and the Actual Interest Rate Series  
(period 1 July 1998 - 30 September 1998)

|                       | $\chi^2$ |
|-----------------------|----------|
| <b>3-Month Rate</b>   | 60,18    |
| <b>6-Month Rate</b>   | 43,24    |
| <b>1-Year Rate</b>    | 20,84    |
| <b>2-Year Rate</b>    | 19,89    |
| <b>3-Year Rate</b>    | 30,73    |
| <b>4-Year Rate</b>    | 42,32    |
| <b>5-Year Rate</b>    | 56,50    |
| <b>10-Year Rate</b>   | 60,17    |
| <b>12.5-Year Rate</b> | 60,17    |

Which indicate a high degree of correlation between the predicted interest rate data and the observed evolution.