Technical University of Denmark

DTU

Empirical yield-curve dynamics, scenario simulation and risk-measures

Madsen, Claus

Publication date:
2012

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

Link back to DTU Orbit

Citation (APA):
Madsen, C. (2012). Empirical yield-curve dynamics, scenario simulation and risk-measures. Department of Management Engineering, Technical University of Denmark.

DTU Library
Technical Information Center of Denmark

EMPIRICAL YIELD-CURVE DYNAMICS, SCENARIO SIMULATION AND RISK-MEASURES

Claus Madsen (cmads@man.dtu.dk)
October 2012

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

Using Principal Component Analysis (PCA) we show that it takes a 4 factor model to explain the variation in the term structure of interest rates over the period from the beginning of 2002 to early-2012. These 4 factors can be called a Slope factor, a Short-Curvature Factor, a Short Factor and a Curvature Factor.

Using the methodology of Heath, Jarrow and Morton (1990) we now specify a 4-factor model for the dynamic in the term structure of interest-rates.

This 4-factor model is afterwards being extended to have a stochastic volatility part, which we assume is to be modelled with a GARCH(1,1) process. The resulting 4-factor yield-curve model belongs to the class of USV (unspanned stochastic volatility) models, as the volatility part is un-correlated with the PCA model for the variation in the yield-curve.

Our Risk-Model relies on the scenario simulation procedure of Jamshidian and Zhu (1997). The general idea behind the scenario simulation procedure is to limit the number of portfolio evaluations by using the factor loadings derived in the first part of paper and then specify particular intervals for the Monte Carlo simulated random numbers and assign appropriate probabilities to these intervals (states).

Our overall conclusion is the following:

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

Keywords: Multi-factor models, PCA, empirical yield-curve dynamics, APT, VaR, ETL, Risk-Model, Stochastic Volatility, Monte Carlo simulation, scenario simulation, non-linear securities - Danish MBBs

Empirical Yield-Curve Dynamics, Scenario Simulation and Risk-Measures[1]

1.   Introduction

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

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

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

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

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

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

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

---

[1] I thank Kostas Giannopoulos for comments to the GARCH estimation and VaR in general.

directly comparable on a risk adjusted basis

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

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

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

The paper is organized as follows: In section 2 we will specify the relationship between the shift function and price sensitivities in a fairly general way. After that in section 3 we will specify a multi-factor term structure model in the Heath, Jarrow and Morton framework and show how the volatility structure is related to the shift function.

Section 4 and 5 will be focusing on the estimation of the non-parametric volatility structure (the shift function) using PCA. After that, in section 6, we will compare price sensitivities in the traditional one-factor duration model with price sensitivities derived from the empirical 4-factor yield-curve model.

The rest of the paper (section 7) will concentrate on scenario simulation and VaR/ETL. We will here start with a short introduction to VaR/ETL and at the same time discuss some of the different approaches that has been proposed in the litterature.

Next we will turn our attention to a practical example using a simple portfolio of government bonds. For that portfolio we will compare the scenario simulation model with the full Monte Carlo simulation procedure.

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

2.   The traditional approach - Duration models

The traditional approach to calculating the sensitivities of interest-rate-contingent claims as a function of changes in the initial term structure is to apply a fairly basic assumption - namely

that term structure movements only appear as additive shifts to the initial term structure.

The price of a coupon bond can generally be expressed as follows:

$$P^k(t,T) = \sum_{j=1}^{n} F_j e^{-R(t,T_j)(T_j - t)} \tag{1}$$

Where R(t,T) is the initial term structure and F_j is the j'th cash flow of the bond.

An expression of the marginal change in this bond, assuming that the term structure movements are defined by a shift function S(t,T), can be formulated as follows[2]:

$$
\begin{aligned}
dP^k(t,T) &= \frac{dP^k(t,T)}{d\tau}\Delta\tau + \frac{dP^k(t,T)}{d\lambda}\lambda + \frac{1}{2}\frac{d^2P^k(t,T)}{d\lambda^2}\lambda^2 \\
&= -\Theta_o^k\Delta\tau - k^k(S(t,T))\lambda + \frac{1}{2}Q^k(S(t,T))\lambda^2 \qquad for\ \lambda = 1
\end{aligned}
\tag{2}
$$

Where $k^k(S(t,T))$ is the price risk, $Q^k(S(t,T))$ is the curvature, Θ_o^k is the time sensitivity of the bond, i.e. the bond theta, $\lambda S(t,T)$ is the size of the impact on the initial term structure, τ is the time to maturity, $\Delta\tau$ is the chosen time-change unit for which the time sensitivity is desired to be computed, and S(t,T) is a specific term structure shift function - as of now left unspecified.

2.1   The relationship between the initial term structure and the shift function

Initially the price of a bond is given by formula 1. Following the impact on the initial term structure caused by the shift function S(t,T) the price of this bond can be formulated as:

$$\tilde{P}^k(t,T) = \sum_{j=1}^{n} F_j e^{-[R(t,T_j) + S(t,T_j)](T_j - t)} \tag{3}$$

Now define a function f(λ) as follows:

$$f(\lambda) = \sum_{j=1}^{n} F_j e^{-[R(t,T_j) + \lambda S(t,T_j)](T_j - t)} - \sum_{j=1}^{n} F_j e^{-R(t,T_j)(T_j - t)} \tag{4}$$

Where f(λ) is an expression of the change in the bond price when the initial term structure changes from R(t,T) to λS(t,T). In addition, as for $\lambda = 1$, $f(\lambda) = \tilde{P}^k(t,T) - P^k(t,T)$ implies that we want to determine the second-order approximation to f(λ) in point $\lambda = 1$.

The f(λ) function can be written as a second-order Taylor expansion around $\lambda = 0$, as follows:

---

[2] If we disregard yield-curve changes of higher order than 2 and time changes of higher order than 1.

$$f(\lambda) = \frac{df(0)}{d\lambda}\lambda + \frac{1}{2}\frac{d^2f(0)}{d\lambda^2}\lambda^2 \tag{5}$$

As f(0) = 0 (see formula 3).

Based on formula 4, it can be deduced that a computation of $\frac{df(0)}{d\lambda}\lambda$ for $\lambda = 1$, yields the desired first-order approximation, and the corresponding argumentation can be used for the second-order approximation.

Bearing this in mind, the price risk ($k^k(S(t,T))$) and the curvature ($Q^k(S(t,T))$) can be formulated as follows:

$$k^k(S(t,T)) = -\sum_{j=1}^{n} F_j(T_j - t)S(t,T_j)e^{-R(t,T_j)(T_j - t)}$$

$$og \tag{6}$$

$$Q^k(S(t,T)) = \sum_{j=1}^{n} F_j(T_j - t)^2 S(t,T_j)^2 e^{-R(t,T_j)(T_j - t)}$$

Where the traditional approach is to let S(t,T) be equal to 0.01, i.e. an additive shift in the term structure of 1%.

Now it can be seen that S(t,T) = 0.01 causes the sensitivities stated in formulas 2 to degenerate into the traditional key figures in a duration/convexity approach, however extended by including the sensitivity to a shortening of maturity[3].

Thus, in this connection it can be deduced that once the shift function is known (is determined/specified), it is possible to calculate relevant and consistent key figures.

3.   A multi-factor model for the bond return

Let us first recall some properties about the HJM framework as this will be our starting point when specifying a bond return model.

The dynamic in the zero-coupon bond-prices P(t,T), for $t < T \leq \tau$, is assumed to be governed by an Ito process under the risk-neutral martingale measure Q:

---

[3] In the rest of paper I will however only focus on the sensitivity that is a function of the shift-function - thus I will disregard time-sensitivities.

$$\frac{dP(t,T)}{P(t,T)} = rdt - \sum_{i=1}^{m} \sigma_P(t,T;i)d\tilde{W}_i(t)$$
$$for$$
$$\tilde{W}_i(t) = W_i(t) - \Gamma_i(t)$$

(7)

Where we have that P(0,T) is known for all T and P(T,T) = 1 for all T. Furthermore r is the risk-free interest rate, and $\sigma_p(t,T;i)$ represents the bond-price volatility, which can be associated with the i'th Wiener process, where $\tilde{W}_i$ is a Wiener process on (Ω,F,Q), for dQ = ρdP and ρ is the Radon-Nikodym derivative. We also have that $\Gamma_i(t)$ represents the market-price of risk that can be associated with the i'th Wiener process.

In order to derive the following results it is not necessary to assume that $\sigma_p(t,T;i)$ for i = {1,2,...m} is deterministic. It is sufficient to assume that $\sigma_p(t,T;i)$ is bounded, and its derivatives (which are assumed to exist) are bounded.

Formula 7 can be rewritten as:

$$d\ln P(t,T) = \left[r - \frac{1}{2}\sum_{i=1}^{m} \sigma_P^2(t,T;i) \right]dt - \sum_{i=1}^{m} \sigma_P(t,T;i)d\tilde{W}_i(t)$$

(8)

The solution to this process can be expressed as:

$$\ln P(t,T) = \ln P(0,T) + \int_0^t \left[r(s) - \frac{1}{2}\sum_{i=1}^{m} \sigma_P^2(s,T;i) \right]ds - \sum_{i=1}^{m} \int_0^t \sigma_P(s,T;i)d\tilde{W}_i(s)$$

(9)

and:

$$0 = \ln P(0,t) + \int_0^t \left[r(s) - \frac{1}{2}\sum_{i=1}^{m} \sigma_P^2(s,t;i) \right]ds - \sum_{i=1}^{m} \int_0^t \sigma_P(s,t;i)d\tilde{W}_i(s)$$

(10)

Where equation 10 follows from the horizon condition that P(T,T) = 1.

The drift in the process for the bond-price - r in formula 7 - can now be eliminated if we consider the difference between the process defined in formula 9 and the process that follows from the horizon condition (formula 10), ie:

$$\ln P(t,T) = \ln\frac{P(0,T)}{P(0,t)} - \sum_{i=1}^{m}\int_{0}^{t}\frac{1}{2}[\sigma_P^2(s,T;i) - \sigma_P^2(s,t;i)]ds$$

$$- \sum_{i=1}^{m}\int_{0}^{t}[\sigma_P(s,T;i) - \sigma_P(s,t;i)]d\tilde{W}_i(s) \qquad (11)$$

The process for the forward-rates, can also be derived - namely by using formula 11, ie:

$$r^F(t,T) = r^F(0,T) + \sum_{i=1}^{m}\int_{0}^{t}\sigma^F(s,T;i)\sigma_P(s,T;i)ds + \sum_{i=1}^{m}\int_{0}^{t}\sigma^F(s,T;i)d\tilde{W}_i(s) \qquad (12)$$

Where $\sigma^F(t,T;i)$ is defined as $\dfrac{\partial\sigma_P(t,T;i)}{\partial T}$, and can be recognized as being a measure for the forward rate volatility.

We furthermore assume that the volatility function satisfies the usual identification hypothesis,

that $\begin{pmatrix} \sigma^F(t,T_1) \\ \sigma^F(t,T_2) \\ . \\ . \\ \sigma^F(t,T_m) \end{pmatrix}$ is non singular for any t and any unique set of maturities $[T_1, T_2, ... T_m]$.

The spot-rate process is easily found from the process for the forward rate, ie:

$$r(t) = r^F(0,t) + \sum_{i=1}^{m}\int_{0}^{t}\sigma^F(s,t;i)\sigma_P(s,t;i)ds + \sum_{i=1}^{m}\int_{0}^{t}\sigma^F(s,t;i)d\tilde{W}_i(s) \qquad (13)$$

That is, the spot-rate process is identical to the forward-rate process, except that in formula 13 we have simultanous variation in the time and maturity arguments.

It may be seen from formulas 12 and 13 that the process for the interest rates is fully defined by the initial yield-curve and the volatility structure, which is precisely the main result of the Heath, Jarrow and Morton (1991) model framework.

From this we can deduce that there are the following relationships between the bond price volatility structure and the shift-function:

$$S(t,T) = \sum_{i=1}^{m} \frac{\sigma_P(t,T;i)}{T-t} = \sum_{i=1}^{m} \sigma_R(t,T;i) \tag{14}$$

It can now be seen that the shift function is identical to the volatility structure of the spot rate structure. In this connection, it can therefore be deduced that the traditional 1-factor duration model given by additive impacts on the term structure of interest rates can be formulated as follows:

$$S(t,T) = \sigma_R \tag{15}$$

Where this formulation of the spot rate volatility structure is identical to the continuous-time version of the Ho and Lee model.

This result can be seen to clash with Ingersoll, Skelton and Weil (1978), who postulate that under the no arbitrage assumption the term structure of interest rates cannot change additively without the term structure of interest rates being flat. Thus, it can be concluded that this assertion is not valid, where this is also shown by Bierwag (1987b), however in a completely different framework

In determining the shift function S(t,T) there are, in principle, two methods, as was also shown by Heath, Jarrow and Morton (1990), namely to either make a pre-defined specification of the functional shape of the volatility structures (the implicit method)[4], or to estimate them historically - that is determine the volatility structure empirically[5].

The first method is the principle used in option theory and will not be the approach used in this paper; instead I intend to use historical data to determine the volatility structures - that is a non-parametric approach.

4.   Principal Component Analysis (PCA)

It is now assumed that the shift function is to be determined by considering the historically observed movements in the term structure of interest rates; in addition, it is assumed that the estimation length (i.e. the number of times the term structure of interest rates is to be observed in the estimation) is equal to L, where l = {1,2,....,L} is the l'th observed variation of the term structure of interest rates. Furthermore, only a finite number of points on the term structure of interest rates are used. The number of points/interest rates is P, where p = {1,2,...,P} is the p'th interest rate and where the terms to maturity of these P interest rates cover the entire maturity spectrum from t-T.

---

[4] See for example Amin and Morton (1993).

[5] This is not to be understood in the sense that it is not possible to estimate the parameters that describe the functional shape of the volatility structures by using historical data, as this is of course possible; see Heath, Jarrow and Morton (1990) "Contingent Claim Valuation with a Random Evolution of Interest Rates".

It is now assumed that the shift function S(t,T) can be written as a linear factor model, as follows[6]:

$$S(t,T) = \sum_{i=1}^{m} v_p(t,T;i)dF_l(t,T;i) + \varepsilon_p(t,T) \qquad (16)$$

Where $v_p(t,T;i)$ is a function consisting of m-independent risk factors, where a risk factor is defined for each p (i.e. each point/interest rate used in the estimation); in matrix form $v_p(t,T;i)$ can therefore be understood as being defined by P-rows and m-columns. In addition $dF_l(t,T)$ represents the change in the risk factors $v_p(t,T)$ across the entire estimation length L, i.e. in matrix form $dF_l(t,T)$ is therefore defined as a matrix with m-rows and L-columns. Furthermore, $\varepsilon_p(t,T)$ is an error element that is assumed to be normally distributed, so that $\varepsilon_P(t,T) \sim N(0,V)$ where it is assumed that the error elements are independent of the interest rate variation, so that V is a diagonal variance matrix, which can be formulated as follows in matrix form:

$$\varepsilon_p(t,T) \sim N(0,V)$$
$$for$$
$$V = \begin{bmatrix} V_1 & 0 & 0 & 0 & 0 \\ 0 & V_2 & 0 & 0 & 0 \\ 0 & 0 & V_3 & 0 & 0 \\ . & . & . & . & . \\ 0 & 0 & 0 & 0 & V_P \end{bmatrix} \qquad (17)$$

That is, $\varepsilon_p(t,T)$ is a PxP matrix. Where the more V deviates from the 0 matrix, the less correctly will the model describe the original data material.

The model constructed here can also be formulated as a function of the bond yield, as follows:

$$\ln\left[\frac{P(t + \alpha,\tau + \alpha)}{P(t,\tau)}\right] = \sum_{i=1}^{m} v_p(t,\tau;i)dF_l(t,\tau;i) + \omega_p(t,\tau)$$
$$for \ all \ \tau \ \varepsilon \ [T_j,T], \ for \ j \ \varepsilon \ [1,2....,P-1]$$
$$and \ \alpha \ \varepsilon \ \Delta l, \ for \ l \ \varepsilon \ [1,2,....,L]$$
$$\qquad (18)$$

Where it can be seen that this model formulation is identical to the Ross (1976) APT model (The Arbitrage Pricing Theory), with the quantity $v_p(t,\tau;i)$ being better known as the i'th factor loading.

The difference between how we have formulated the APT model and the traditional APT model is that we have no factor-independent Rate-Of-Return (apart from the residual element).

---

[6] This is of course not the only formulation of the shift function imaginable; for a brief review of models related to this formulation, please refer to Appendix A.

We focus namely on modelling/estimating the total Rate-Of-Return of the bonds, and not the excess rate-of-return, where the excess rate-of-return is defined as the Rate-Of-Return that exceeds the risk-free interest rate.

In addition, it should be mentioned at this point that we have not made any kind of explicit definition of the individual factors F; the explanation is that this analysis will not focus on estimating spot rate volatility structures that have a specific pattern, but on the other hand on identifying the number of linear independent parameters that explain the variation in the term structure of interest rates historically.

In connection with the estimation of the model, we have used formula 16 as my starting point. The model is estimated by constructing a matrix of historical variations in the term structure of interest rates, after which the loading matrix $v_p(t,\tau;i)$ and the factor values $dF_l(t,\tau;i)$ are estimated using the Principal Component Analysis (PCA).

The underlying idea of PCA is to analyze the correlation structure (the correlation matrix), that is, the starting point is to find the correlation matrix on the basis of the matrix of historical interest changes, after which it is standardized in such a way that the diagonal consists of 1's, which means that the dispersion matrix is equal to its correlation matrix. The principal factor solution to the estimation problem is then:

$$K\Lambda^{\frac{1}{2}} = [\sqrt{\lambda_1}k_1, \sqrt{\lambda_2}k_2, \ldots, \sqrt{\lambda_m}k_m] \qquad \textbf{(19)}$$

where $[\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_m]$ and $[k_1, k_2, \ldots, k_m]$ respectively repre correlation matrix and the associated standardized orthogonal eigenvectors. In addition, K is a matrix containing the eigenvectors by columns and Λ is a diagonal matrix with the eigenvalues in the diagonal. As the correlation matrix is an estimate of the squared loading matrix, the squared loading matrix is defined by formula 19.

In this formulation it is assumed that the number of factors is known in advance, namely m. If this is not the case, the factors that are attributable to the highest eigenvalues are selected until there is a satisfactory description of the data material, where Kaiser's theorem suggests that all the factors that have an eigenvalue higher than 1 are to be chosen.

As a factor loading is a vector of correlation coefficients, this means that the best interpretable factor loadings are achieved when they are either close to 0 (zero) or 1[8]. This can be achieved by rotating the factors, the main rule being that a new estimate for the loading matrix can be obtained - without changing the explanatory degree at row level, or for that matter at column

---

[7] In this connection we have used Jacobi's algorithm to solve the eigenvalue problem.

[8] It can therefore also be concluded that the individual loadings relate to the explanatory degree, in fact, the explanatory degree of the individual factors is given by the squared loadings.

level - by multiplying the principal factor solution by an orthogonal matrix. In this connection we have chosen to use Kaiser's Varimax method for rotating the factors.

The generation of the factor values is the last point missing. Where this can be done in the following way $F = D^T(s)L[L^TL]^{-1}$, for L being the loading matrix and D(s) being the or interest rate variation matrix (with the number of rows equal to P and the number of columns equal to L), however in a standardized form, i.e. each column has a mean value equal to 0 (zero) and a standard deviation of 1[9]. In conclusion, it should be stressed that these considerations regarding PCA are taken from Harman (1967).

Lastly, with respect to the model defined in formulas 16 and 17, the individual factor loadings are assumed to be constant, whereas the factor values (factor scores) are time-dependent.

This indicates that - assuming that the residual element is negligible - the factor values can be understood as being a time-dependent weighting parameter which, in principle, can be fixed at 1% when risk parameters are to be calculated. Where this means that the shift function can be formulated as follows:

$$S(t,T) \approx \sum_{i=1}^{m} \frac{v_p(t,T;i)}{100} \qquad (20)$$

Which results in term structure shifts being measured in terms of standard deviations, which is also a relevant calculation unit when bearing formula 14 in mind. The determination of risk parameters in this multi-factor model will be discussed in more detail in section 6.

5.   Estimation of the Volatility Structure

The analysis period has been selected to cover every Wednesday over the period 2 January 2002 - 2 January 2012, based on the yield-curve estimated using the methodology of B-Splines, see for example Webber and James (2000).

We are using the following vector of maturity dates as our key-maturity dates: [0.083,0.25,0.5,1,2,3,4,5,7.5,10,12.5,15,20,25,30] - that is P = 15.

In the table below are shown the ordered eigenvalues for each of the 15 potential factors.

Eigenvalue	Degree of Explanation
9,66	64,37
2,04	78,00
1,47	87,81
1,04	94,74

---

[9] In the construction of this expression of the factor values, the residual matrix has been disregarded, as will appear. This means that it is implicitly assumed that so many factors are being selected that the entire variation in the data material has been described.

0,27	96,56
0,23	98,09
0,10	98,76
0,07	99,23
0,05	99,54
0,03	99,72
0,02	99,87
0,01	99,93
0,01	99,99
0,00	99,99
0,00	100,00

According to Kaiser's theorem one should select all the factors that have an eigenvalue higher than 1, which in our case means the first 4 eigenvalues. Selecting 4-factors will explain approximate 95% of the total variation in the yield-curve over the analysed time-period - this is assumed to be sufficient.

Figure 1 below show the estimated factor loadings (for the 4 most important factors) before the varimax rotation and figure 2 show the factor loadings after we have performed the varimax rotation:

Figure 1   **Estimated Factor Loadings - Original**
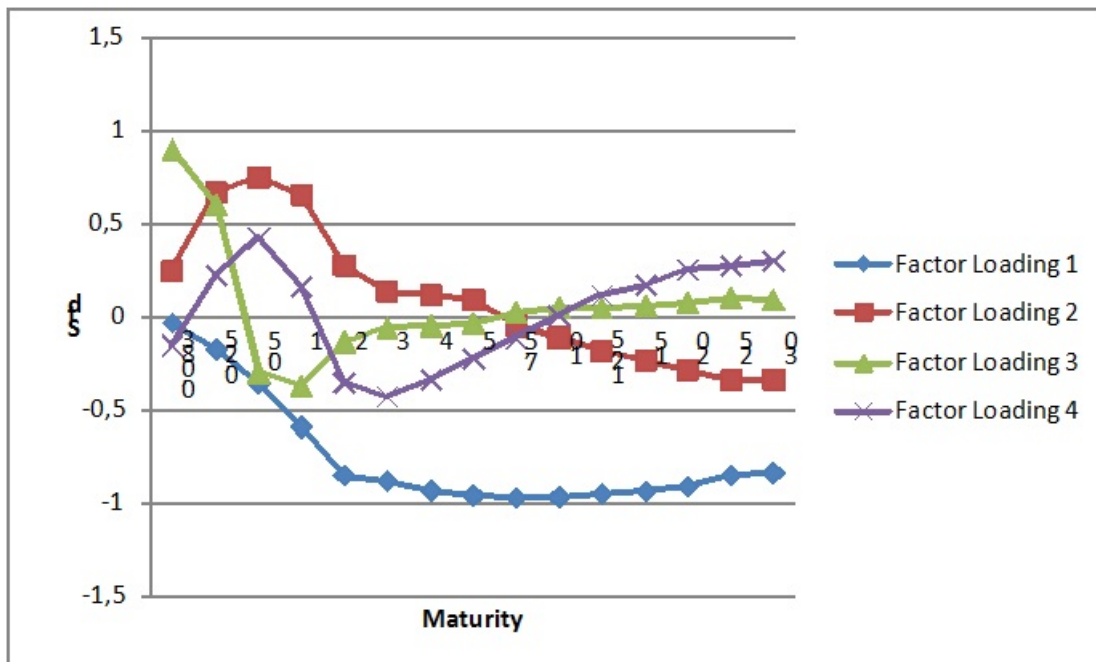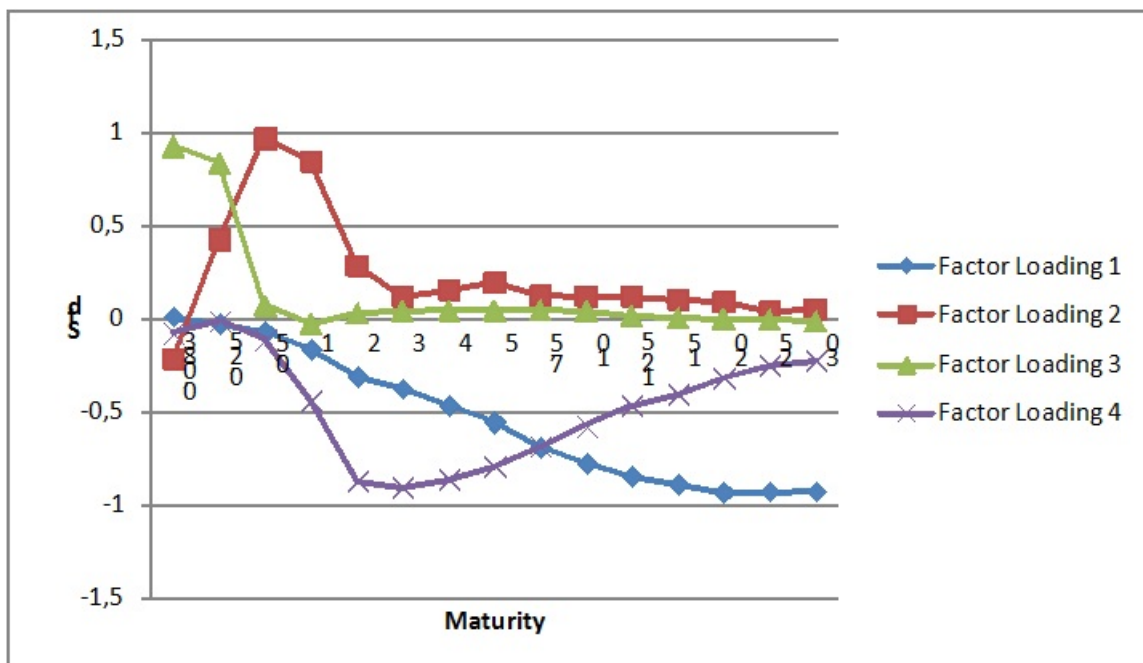   **(2. January 2002 - 2. January 2012)**

Figure 2   **Estimated Factor Loading - after Varimax Rotation**
(2. January 2002 - 2. January 2012)

In figure 3 below we have shown the degree of explanation for each of the factors[10]:

Figure 3   **Degree of Explanation (2. January 2002 - 2. January 2012)**

---

[10] From now on when we use the phrase factor loading we refer to the factor loadings that are obtained after the varimax rotation.
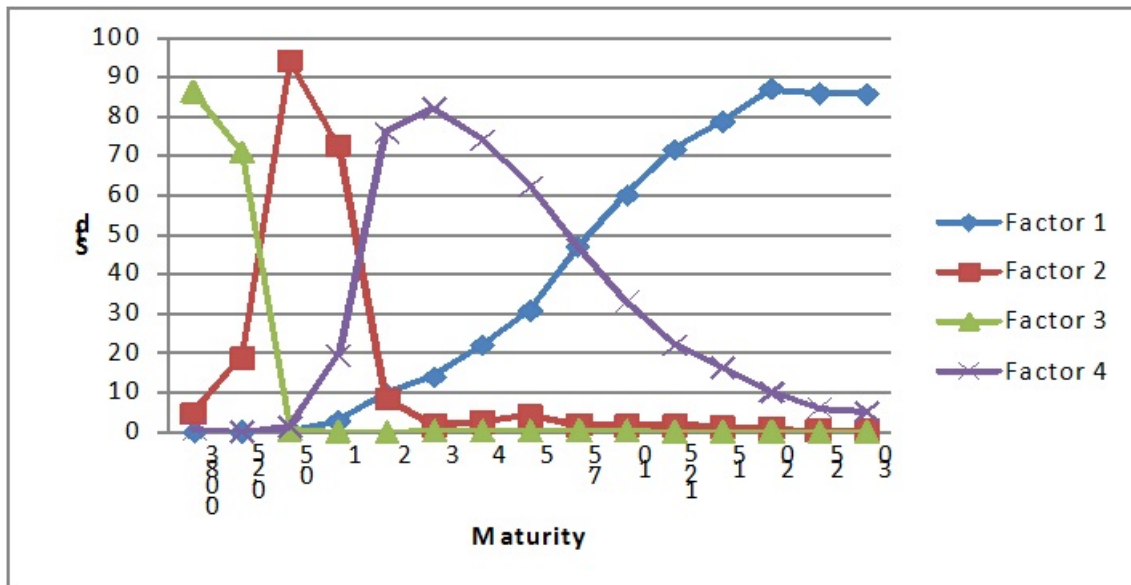
These 4 factors can be interpreted as: a slope factor, a short-curvature factor, a short factor and a curvature factor. It is here interesting to mention, that no matter how we slice and dice our data, this meaning that if we for example select the period from 2. January 2002 to 2. January 2008 we still get the same 4 factors.

As can be seen in the graphs above, the result we obtain are somewhat different from what normally has been reported in the litterature, for example Dahl (1989) using Danish data, Litterman and Scheinkmann (1988) and Garbade (1986) using American data, and Caverhill and Strickland (1992) using English data, which conclude that empirically 3 factors exist that describe the dynamics of the term structure of interest rates.

Lord and Pelsser (2006) concluded the following in their study of the fact that traditionally 3-factors are observed, namely: A Level Factor, a Slot Factor and a Curvature Factor:

"...we analysed the so-called level, slope and curvature pattern one frequently observes when conducting a principal components analysis of term structure data. A partial description of the pattern is the number of sign changes of the first three factors, respectively zero, one and two. This characterisation enables us to formulate sufficient conditions for the occurrence of this pattern by means of the theory of total positivity. The conditions can be interpreted as conditions on the level, slope and curvature of the correlation surface. In essence, the conditions roughly state that if correlations are positive, the correlation curves are flatter and less curved for larger tenors, and steeper and more curved for shorter tenors, the observed pattern will occur.

As a by-product of these theorems, we prove that if the correlation matrix is a Green's or Schoenmakers-Coffey matrix, level, slope and curvature is guaranteed. An unproven conjecture at the end of this paper demonstrates that at least slope seems to be caused by two stylised empirical within term structures: the correlation between two contracts or rates

decreases as a function of the difference in tenor between both contracts, and the correlation between two equidistant contracts or rates increases as the tenor of both contracts increases.

...we can conclude that the level, slope and curvature pattern is part fact, and part artefact. It is caused both by the order and positive correlations present in term structures (fact), as well as by the orthogonality of the factors and the smooth input we use to estimate our correlations (artefact)."

So the reason for the difference we observe here, is probably given by the fact that we are employing a B-Spline estimation algorithm (which are very flexible and able to capture both the long end of the yield-curve and the short end simultanously), while the other studies use (if it is mentioned at all) less flexible functional forms for the yield-curve, like for example the Nelson and Siegel (1988) model or extensions hereof like the Svensson (1994) model.

To investigate that observation we have re-done the estimations of the PCA-model for the same period 2. January 2002-2. January 2012, but now for the following set of interest-rate maturities: [1,2,3,4,5,7.5,10,12.5,15,20,25,30]. If our observation is correct we would expect to get 3 factors, namely: A Level Factor, a Slot Factor and a Curvature Factor. The result are shown below:

Estimated Factor Loadings (No Rotation) and a subset of maturities (2. January 2002 - 2. January 2012)
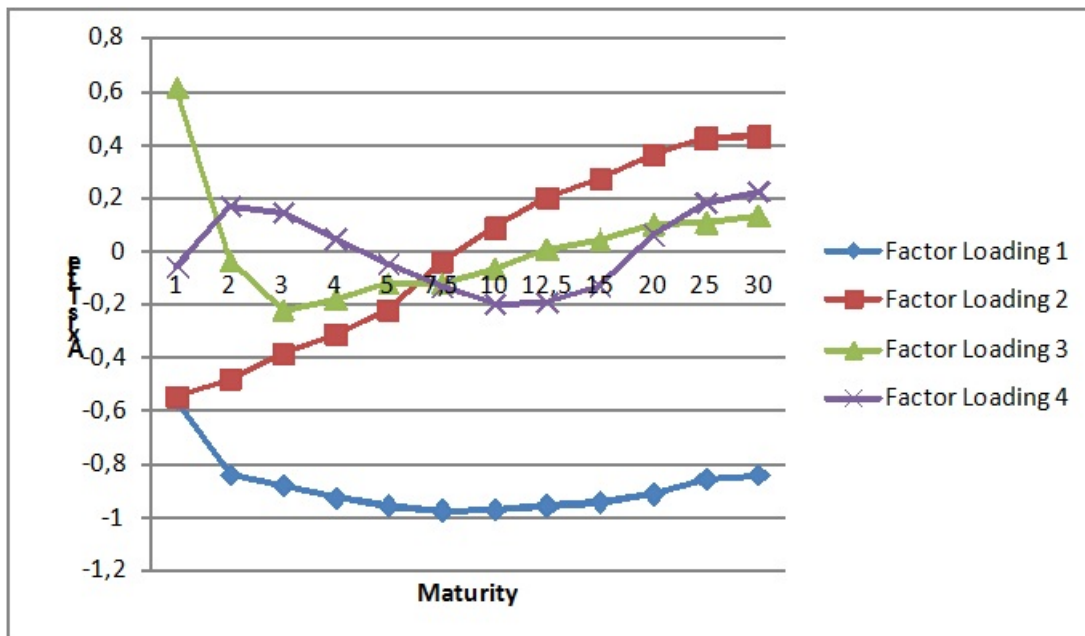
The first 5 eigenvalues are:

Eigenvalue	Degree of Explanation
9,52434	79,36953
1,44311	91,39544

0,54039	95,89866
0,25575	98,02987
0,09443	98,81678

From the above we can actually see that now the results obtained are identical with what is normal reported in the litterature - namely 3-factors that explan more than 95% of the variation in the yield-curve which can be interpreted as: A Level Factor, a Slope Factor and a Curvature Factor.

6. Measuring of Risk in a multi-factor shift function model

The shift function in this 4-factor term structure model can now be formulated as follows:

$$S(t,T) = \frac{v_1 dw_1 + v_2 dw_2 + v_3 dw_3 + v_4 dw_4}{100}; \quad for \ dw_1 = dw_2 = dw_3 = dw_4 = 1 \quad (21)$$

where v_m, for m = {1,2,3,4} is the vector of factor loadings and dw_m is the vector of factor scores.

In this connection, the vector sensitivities can be formulated as follows for the coupon bond $P^k(t,T)$:

$$\frac{\partial P^k(t,T)}{\partial dw_m} \frac{1}{P^k(t,T)} = -\frac{\sum_{j=1}^{n} F_j(T_j - t)v_m(j)e^{-R(t,T_j)(T_j - t)}}{P^k(t,T)}; \quad for \ m \ \varepsilon \ [1,2,3,4] \quad (22)$$

where this can be regarded as the factor duration.

The factor convexity has the following form:

$$\frac{\partial^2 P^k(t,T)}{\partial dw_m^2} \frac{1}{P^k(t,T)} = \frac{\sum_{j=1}^{n} F_j(T_j - t)^2 v_m(j)^2 e^{-R(t,T_j)(T_j - t)}}{P^k(t,T)}; \quad for \ m \ \varepsilon \ [1,2,3,4] \quad (23)$$

These two equations deserve a few comments.

To fully understand these relations we first note that if we disregard v_m, for each m, then the equations degenerate into the portion of returns of the bond which result from a unit change (a standard deviation, S(t,T) = 1%) in the whole yield-curve. The factor model on the other hand tells us that a unit change to the factor does not change the yield-curve by one percent - but by v_m percent.

15

From this we can deduce that in order to find the total impact on bond returns from a factor change, we need to scale by a weight that is exactly equal to the appropriate factor loading.

6.1 A practical example - risk factors in the 4-factor term structure model compared with the risk factors in the traditional 1-factor duration model

In order to illustrate the difference between the traditional approach in calculating price sensitivities (i.e. S(t,T) = 0.01), and the price sensitivities generated by this 4-factor model the table below shows the risk calculated using modified duration and the factor duration for a wide range of Danish Government bonds.

Table 1:   Duration and Factor Sensitivities (2. January 2012)

sin	Instrument-Name	Mod. Duration	Factor 1 Dur	Factor 2 Dur	Factor 3 Dur	Factor 4 Dur
DK0009922593	4 Ink st 2012	-0,869	-0,119	0,768	-0,001	-0,307
DK0009920894	5 Ink st 2013	-1,824	-0,524	0,684	0,043	-1,466
DK0009922833	2 Ink st 2014	-2,812	-1,020	0,432	0,121	-2,523
DK0009921439	4 Ink st 2015	-3,656	-1,632	0,606	0,171	-3,151
DK0009922759	2.5 Ink st 2016	-4,636	-2,477	0,909	0,220	-3,705
DK0009902728	4 Ink st 2017	-2,800	-1,242	0,667	0,132	-2,217
DK0009921942	4 Ink st 2017	-5,352	-3,126	0,964	0,256	-4,052
DK0009922403	4 Ink st 2019	-6,930	-4,685	0,965	0,334	-4,708
DK0009922676	3 Ink st 2021	-8,673	-6,484	1,149	0,383	-5,198
DK0009918138	7 Ink st 2024	-9,524	-7,577	1,263	0,236	-4,941
DK0009922320	4.5 Ink st 2039	-18,076	-16,151	1,268	-0,002	-5,517

The most interesting to note in this table is the fact that the degree of sensitivity a given bond has to a particular (risk) factor has got nothing to do with the importance of this (risk) factor - this issue will be discussed later in section 7.4.

One can from the table see that this 4-factor model measure the risk at the relevant segments/parts of the yield-curve where the risks of each securities are more logically attributable than what is the case for the traditional 1-factor duration model. For example the short bullet bond 4 Ink st 2012 is more sensitive to factor 2 than the other 3-factors as the maturity date of that bond occur at a time where factor 2 has more influence than the other 3-factors, see Figure 2.

This ends the first part of the paper - namely determining the number of factors that drive the evolution in the yield-curve using an empirical approach.

In the next section we will explain how we - for the current analysis - have defined to setup our Risk-Model framework.

7.   Risk-Models - a Short Survey

VaR is probably the most important development in risk management. This methodology has been specially designed to measure and aggregate diverse risky positions across an entire institution using a common conceptual framework. Even though these measures come under difference disguises e.g. Banker Trusts Capital at Risk (CaR), J.P. Morgans Value at Risk (VaR) and Daily Earnings at Risk (DeaR), they are all based on the same foundation. Even though different institutions has come up with their own names the one that seems to be most commonly used is VaR - which is the name we will be using here.

Definition 1:

Risk Capital is defined as the maximum possible loss for a given position (or portfolio) within a known confidence interval over a specific time horizon.

VaR plays three important roles within a modern financial institution:

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

Even though there are some open issues regarding its calculation, VaR is nevertheless a very useful tool for helping management to steer and control diverse risk operation.

The problem with VaR is that there is a variety of different ways to implement the definition of Risk Capital (see Definition 1) each having distinct advantages and weaknesses. In order to get a better grasp of the trade-offs implicit in each method - it is important to understand which kind of components Risk Capital is built around.

Risk Capital comprises of the following two (2) distinct parts:

- The sensitivity of a position's (portfolio's) value to changes in market rates
- The probability distribution of changes in the market rates over a predefined reporting horizon

Given these assumption - VaR is (usually) defined as the maximum loss within the 99% confidence interval over at 10-day time period (Basel II) - for Market Risk calculations[11].

In the following table we have listed the 3 most common methods used when calculating VaR - together with their advantages and disadvantages:

---

[11] Recently there has been talk about changing from the use of VaR to the use of ETL, for more about ETL see below.

Method	RiskMetric™	Delta-Gamma Methods[12]	Simulation Based Methods
Description	Assume that asset returns are normally distributed, implying linear pay-off profiles and normally distributed portfolio returns[13]: $$VaR = \alpha\sqrt{\Delta t}\sqrt{VCV^T}$$	Assume that asset returns are normally distributed, but pay-off profiles are approximated by local second order terms	Approximate probability distribution for asset returns based on simulated rate movements, either historically or model based
Advantage	Simplicity.	Simplicity. Captures second order effects.	Captures local and non-local price movements. Takes into account fat tails, skewness and kurtosis. With Monte Carlo simulation, we have flexibility to select a probability distribution. With historical simulation we do not need to infer a probability distribution.
Disadvantage	Assuming normality of returns ignores fat tails, skewness and kurtosis. Ignores higher order moments in sensitivities. Captures only risks of local movements for linear securities. Does not capture risks of non-local movements.	Assuming normality of returns ignores fat tails, skewness and kurtosis. Does not capture risks of non-local movements.	Computationally expensive - for Monte Carlo simulation. With Monte Carlo simulation then we need to select a probability distribution. With historical simulation then we cannot select a probability distribution.

---

[12] This method is due to Wilson (1994).

[13] $\alpha = 2.54$ if we wish to calculate VaR at the 99% confidence level, Δt = unwind period, V = vector of volatilities and C is the correlation matrix.

It is not directly mentioned in this table but in general it is assumed that the expected return is

zero (0), ie: $E_{t-1}\left[\dfrac{P_t - P_{t-1}}{P_{t-1}}\right] = 0$. This assumption is related to the fact that it is the standa

measure Risk over a "short" time interval, for longer periods (unwinding periods) one should not employ the zero expected return assumption!

The issue that has been mostly widely discussed in the literature with respect to VaR is how to derive appropriate volatility estimates? - see for example Alexander (1996). We will not here discuss that issue but will address it in section 7.3.

Another very important issue in connection with Risk-Calculations is how to derive a reliable correlation matrix? The problem with estimating and modelling the correlation matrix can be summarized as follows:

- Correlations coefficients are highly unstable and their signs are ambiguous
- If for example we have 15 interest rates we need to keep track of (model)

  $$\prod_1^{14} i$$ correlation coefficients

- We need a long data period in order to have enough degrees of freedom to estimate a reliable correlation matrix. There is however no guarantee that the resulting matrix satisfy the multivariate properties of the data - we might even encounter a correlation matrix that is not positive definit

These observations have inspired research to reduce the dimensionality - where a common suggested method is PCA - like the one we performed in section 4 and 5 in this paper. The nice property in this context is - as mentioned in section 5 - that even though the correlation coefficients are highly unstable the factor loadings are extremely stable[14].

Barone-Adesi, Bourgoin and Giannopoulos (1997) has suggested an interesting approach to worst case scenarios. This method is based on historical returns combined with a GARCH approach (actually a AGARCH-model) for forecasting purposes. The procedure does not employ the correlation matrix directly as the correlation is embedded indirectly in the historical simulation procedure - which is a very neat property of their method.

7.1   VaR contra ETL

Artzner, Delbaen, Eber and Heath (1999) showed that VaR is not what is termed a coherent risk-measure. Where the definition of a coherent risk-measure is a risk-measure that satisfies:

---

[14] The Factor ARCH approach of Engle, Ng and Rothschild (1990) is also in this spirit - see Christiansen (1998) on Danish data.

1: Monotonicity: $Y \geq X \Rightarrow rm(Y) \leq rm(X)$
2: Subadditivity: $rm(X+Y) \leq rm(X)+rm(Y)$
3: Positive homogeneity: $rm(hX) = hrm(X)$ for $h > 0$
4: Translational invariance: $rm(X+n) = rm(X) - n$, for some certain amount n

Properties 1, 3 and 4 are well-behavedness properties to rule out strange outcomes. The most important property is the sub-additivity property.

The sub-additivity property tell us that a portfolio made up of subportfolios will have a risk that is not more than the sum of the risks of the subportfolios:

- This reflects an expection that when we aggregrate individual risk, they diversify or, at worst, do not increase – the risk of the sum is always less than or equal to the sum of the risks

This sub-addivity property spells trouble for VaR as VaR is not sub-additive - except by posing particulary restrictions on the profit/loss distribution.

One risk-measure candidate which is coherent is ETL (expected tail loss)[15].

ETL is defined as the average of the worst 100(1-α)% of loses, that is:

$$ETL_\alpha = \frac{1}{1-\alpha} \int_\alpha^1 q_p \, dp$$

In case the profit/loss distribution is discrete ETL can be calculated as:

$$ETL_\alpha = \frac{1}{1-\alpha} \sum_{p=0}^{\alpha} [\text{pth highest loss}] \times [\text{probability of pth highest loss}]$$

From this we can deduce that ETL is the probability-weighted average of the tail losses – which suggest that ETL can be estimated as an average of "tail VaRs".

For the above reason we will be both considering VaR and ETL when calculating risk-measures.

For our implementation of a Risk model we will limit ourselves to domestic interest rate data - but extending to multiple currencies is of course possible, this will however not be treated in this paper.

---

[15] This is sometimes also referred to as Tail-Loss or CvaR (Conditional Value-at-Risk), we however prefer ETL.

Our Risk approach is based on the following basic assumptions:

- We will build our VaR/ETL approach on top of our empirical yield curve dynamics model from section 4 and 5, for elaboration see section 7.1
- We will estimate the volatility structure using the GARCH approach, see section 7.3
- Our approach is a simulation based method which follows along the lines of Jamshidian and Zhu (1997)[16], see section 7.2
- One could say that we combine the methodology of Barone-Adesi, Bourgoin and Giannopoulos (BABG) (1997)[17], with the simulation based approach from Jamshidian and Zhu (JZ) (1997)
- We will impose the restriction that the expected price change in a discount bond is equal to the price change with arises through the combined effect of a shortening of maturity and a movement down the yield-curve. This means that we will assume that the expected yield-curve will be equal to the initial yield-curve

7.1   An Empirical Multi-Factor Model

Before continuing it is worth mentioning that we have the following relationship between dW_i (the Wiener process) and dw_i (the factor scores), for $i = [1,2,3,..,P]$, and P = number of key-rates (maturities) (in our case 12)[18]:

$$dw_i = \frac{1}{\lambda_i}\sum_{p=1}^{P} v_{pi}dW_p \Rightarrow dW_p = \sum_{i=1}^{P} v_{pi}dw_i \qquad (24)$$

this equation arises because $dW_i \approx N(0,1)$ and as we have normalized the correlation matrix[19] then $|v_i|^2 = \lambda_i$ and therefore $dw_i \approx N(0,1)$. That is, the factor scores are assumed to be standardised normal distributed variables.

The equal sign in the last equality in formula 24 becomes an approximation sign for P > 4 - where 4 is the number of factors. As follows we have limited ourselves to a 4-factor model as we are using Kaiser's theorem to select the number of factors, see section 4. Furthermore we

---

[16] Se also Frey (1998).

[17] One might add that the approach from Hull and White (1998) is quite similar to the BAGS model.

[18] For these derivations, we note that if we consider the factor loadings before the varimax rotation - then the correlation matrix is given by KK^T and $K^TK = diag(\lambda)$ - if K is the factor loading matrix. It should here be stressed that if we consider the factor loadings after varimax rotation then K^TK will only be equal to $diag(\lambda)$ if the vector of eigenvalues are being recalculated from the factor loading matrix after rotation.

[19] In the case that we have not normalized the correlation matrix then λ is the vector of eigenvalues for the correlation matrix. This is however only true if we do not perform varimax rotation of the factors.

assume that the approximation error this gives rise to in equation 24 is negligible.

As mentioned above the volatility parameters can be obtained directly from the estimated eigenvalues - that is however not the approach we will employ here. We will instead use a GARCH(1,1) model to estimate the volatility structure and use this model to derive volatility forecasts that is to be used in connection with our VaR/ETL calculations, for more see section 7.3 and section 7.4.

This means that our model setup will belong to the type of models that normally is referred to as unspanned stochastic volatility (USV) models (see for example Casassus, Collin-Dufresne and Goldstein 2005).

A number of papers have investigated whether the existing models can capture the joint dynamics of term structure and plain-vanilla derivatives such as caps and swaptions. Simple, model-independent evidence based on principal component analysis suggests that there are factors driving Cap and Swaption implied volatilities that do not drive the term structure. Collin-Dufresne and Goldstein 2002 call this feature unspanned stochastic volatility, simplified put, it appears that fixed-income derivatives such as Caps and Swaptions cannot be perfectly replicated by trading (even in a very large number of) the underlying bonds. That is, in contrast to the predictions of standard short-rate models - which assume that bonds does span the bond market.

A model that exhibits USV, is a model where bond prices are insensitive to volatility-risk, and hence cannot be used to hedge volatility-risk. In our case this is true because there is no correlation between our GARCH (1,1) model for volatility and our PCA-model for the yield-curve dynamic.

Using the results from section 3 we can now in abstract form express the dynamic in the yield-curve as follows:

$$dR(t,T) \ = \ \mu(T)dt \ + \ \sigma(t,T)S(t,T)dW \qquad (25)$$

where $S(t,T)$ is the shift-function matrix which in our case for all practical purposes can be fixed to have 4 columns. Furthermore we have that $\sigma(t,T)$ is a vector of volatilities. Where as mentioned above the vector of volatilities $\sigma(t,T)$ will be derived from a GARCH-model.

The drift in formula 25 depends on the expectation hypothesis, which means that the process has been written under the original probability measure - the implication is that the model as it is formulated in formula 25 is not appropriate for pricing purposes.

Using formula 24 and because for our purpose it is more convenient (and appropriate) to assume a lognormal process, we can rewrite formula 25 as:

$$\frac{dR(t,T)}{R(t,T)} \ = \ \mu(T)dt \ + \ \sigma(t,T)v_1dw_1(t) \ + \ \sigma(t,T)v_2dw_2(t) \ + \ \sigma(t,T)v_3dw_3(t) \ + \ \sigma(t,T)v_4dw_4(t) \quad (26)$$

or alternatively expressed in integral form:

$$R(t,T) = \mu(T)e^{\sigma(t,T)v_1 dw_1 + \sigma(t,T)v_2 dw_2 + \sigma(t,T)v_3 dw_3 + \sigma(t,T)v_4 dw_4} \tag{27}$$

We have assumed that $\mu(T)$ (for all $T > t$) is equal to the initial yield-curve. That is we assume that the expected yield-curve is equal to the actual yield-curve.

This is a more intuitive assumption than using forward rates when our motive is to generate yield-curve scenarios for future dates that are to be used as inputs in various valuation models. The model setup is nevertheless arbitrage free as the uncertainty of each key rate is governed by its own market price of risk - which implicitly are determined through the expectation hypothesis. Again we stress that we are here working under the actual probability measure and not under the risk-neutral probability measure.

For longer time-horizons, the above formulation is not appropriate, due to the fact that we have omitted mean-reversion in the drift specification - extension to a model which has mean-reversion is however straightforward, but for our purpose the above formulation is sufficient.

7.2   Scenario Simulation

In order to simulate yield-curve movements, we can apply the Monte Carlo method to equation 27. If we now assume that we perform N simulations for each of the 4-factors - then the total number of yield curves at a given point in time will be N^4. For N = 100 we will therefore have 100,000,000 future yield-curves which is (more than) adequate in order to generate a probability distribution of the future yield curves. Another approach is of course to perform a Monte Carlo simulation directly in our 4-factor model, however still a fairly high number of simulations is required in a 4-factor model in order to obtain a "precise" probability distribution - especially as we are interested in the tails of the distribution.

In general, it can be said that "brute force" Monte Carlo will require a substantial amount of portfolio evaluation. This approach will for that reason not be of practical use, if firstly the portfolio is fairly large and secondly if the portfolio contains a fair number of non-linear instruments of moderate complexity, like for example Bermudan Swaptions, flexi-caps and/or Mortgage-Backed-Bonds.

The interesting question is - how can we reduce the computational burden while at the same time having a reasonable specification of the future yield-curves probability distribution?

As shown by Jamshidian and Zhu (1997) this is possible. The argumentation is as follows:

Let us suppose that x is a random variable with distribution P(x). In a Monte Carlo method, we simulate N possible outcomes for the random variable x - where each simulated x has the same probability.

However, we have that numbers x_i's, for i = [1,2,...,N], that falls between x_l and x_u is proportional to the probability $P(x_l < x_i \leq x_u)$. From this it follows that we can select a regio $(x_l, x_u]$ and assign a given probability for all numbers that fall inside this region. If we utilize this procedure, it is possible for example to perform 100 simulations for each factor but we only need to perform the portfolio evaluation at a limited number of states - more precisely at the number of states that equal the number of predefined regions.

A good candidate for a probability distribution is the multinomial distribution[20]. If k + 1 states (ordered from 0 to k) are selected then the probability for a given state i is given by the binomial distribution and can be expressed as:

$$Prob(i) = \frac{2^{-k}k!}{i!(k-i)!} \quad for \ i \ \varepsilon \ [0,1,2,...,k] \tag{28}$$

For this distribution we have that there is a distance of $\dfrac{2}{\sqrt{k}}$ between two adjacent states, and furthermore is the furthest state $\sqrt{k}$ standard deviation away from the center.

Let us now assume that we only need to select nine (9) states for factor 1, five (5) states for factor 2 and three (3) states for factor 3[21]. This gives us all in all 9 x 5 x 3 = 135 scenarios in which we have to evaluate our portfolios - which is independent of the number of Monte Carlo simulations.

Under this assumption we get the following probabilities for each of the states for the 3-factors:

$$Factor \ 1: \frac{1}{252}, \frac{1}{32}, \frac{7}{64}, \frac{7}{32}, \frac{35}{128}, \frac{7}{32}, \frac{7}{64}, \frac{1}{32}, \frac{1}{252}$$
$$Factor \ 2: \frac{1}{16}, \frac{1}{4}, \frac{3}{8}, \frac{1}{4}, \frac{1}{16} \tag{29}$$
$$Factor \ 3: \frac{1}{4}, \frac{1}{2}, \frac{1}{4}$$

As the 3-factors are independent we have that their joint probability is defined by the products of the three marginal probabilities.

There has been discussing in the litterature about the robustness of the JZ approcimation, for

---

[20] See Cox and Rubenstein (1985).

[21] In Jamshidian and Zhu (1997) they select 7,5 and 3 - but as they mention (their footnote 7) the determination of the appropriate number of states for each factor is an empirical question.

example Abken (2000), he states the following: "*The outcomes for the nonlinear test portfolios demonstrate that scenario simulation using low- and moderate dimensional discretizations can give "poor" estimates of VaR. Although the discrete distributions used in scenario simulation converge to their continuous distributions, convergence appears to be slow, with irregular oscillations that depend on portfolio characteristics and the correlation structure of the risk factors.*" This observation is in line which the oscillations seen in the pricing of long dated options using the Cox, Rox and Rubenstein (1979) model, and something one should be aware of....we will address that issue later.

Gibson and Pritsker (2000) have the following comment: "*We show that their method has two serious shortcomings which imply it cannot accurately estimate VaR for some fixed-income portfolios. First, risk factors chosen using principal components analysis will explain the variation in the yield curve, but they may not explain the variation in the portfolio's value. This will be especially problematic for portfolios that are hedged. Second, their discrete distribution of portfolio value can be a poor approximation to the true continuous distribution.*" They instead propose a Grid-based Monte Carlo approach on the (derived) risk-factors for the portfolio in question. A similar approach is advocated by Chishti (1999).

We will look into these issues in section 7.4.

7.3   Stochastic Volatility

As mentioned earlier we are using the GARCH(1,1) model to specify the volatility in our multi-factor yield-curve model. To examplify this we have below in figure 4 shown the volatility structure as of the 2 January 2012 that has been determined using the GARCH(1,1) model:

Figure 4 **GARCH(1,1) Volatility Structure (2. January 2012)**
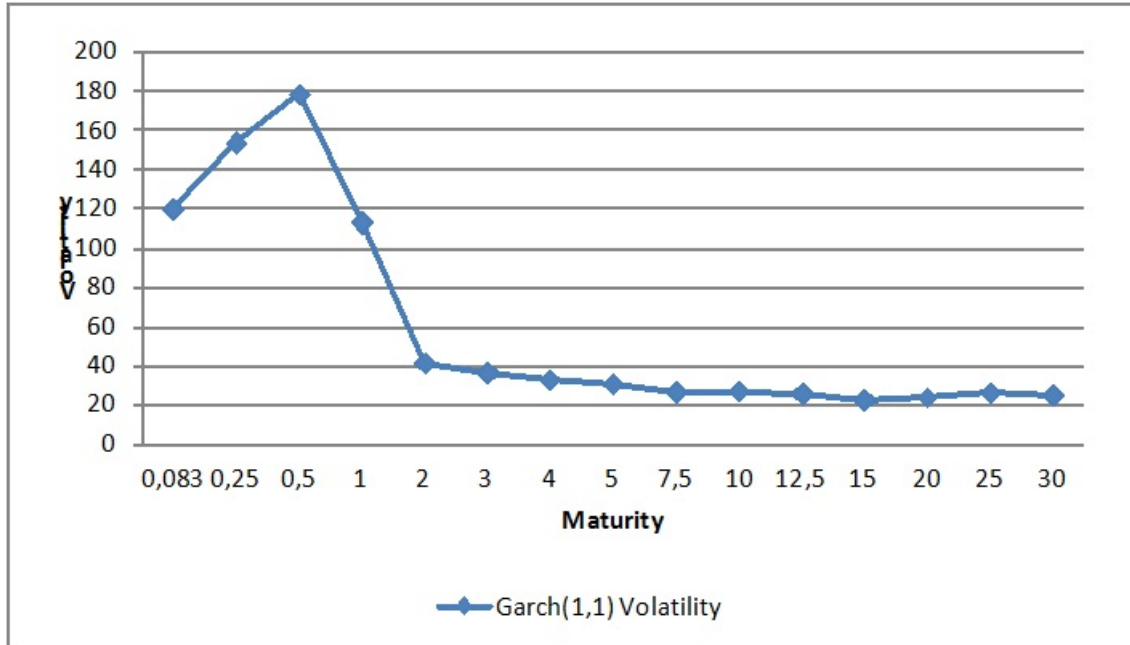
For the interested reader we have in Appendix C briefly discussed the idea behind GARCH and also explained how we have constructed the volatility structure.

7.4 An Illustrative Example

In order to show how the approach works we have constructed the following example:

The portfolio we have selected is the following:

Table 2: **Government Bond Portfolio as of 2. January 2012**

sin	Instrument-Name	Position	Clean Price	Accrued IR	Dirty Price	Value
DK0009920894	5 Ink st 2013	2500000	109,084	0,697	109,781	2.744.518,03
DK0009902728	4 Ink st 2017	1500000	108,000	2,230	110,230	1.653.442,62
DK0009922320	4.5 Ink st 2039	4000000	149,520	0,627	150,147	6.005.881,97
					Total Value:	10.403.842,62

From table 2 it follows that the value of the portfolio as at 2 January 2012 is 10.403.842,62 kr.

We now wish to determine the distribution of the portfolio for a 10-day horizon using the approach we have specified above for our 4-factor model. Given this distribution we can then derive our VaR estimate give our model setup.

As comparison benchmark we have used Monte Carlo to simulate 250.000 yield-curves for our 10-days horizon for our 4 factor-model and calculated the joint probability distribution.

The purpose of this is in order to investigate the approximation error in the Jamshidian and Zhu (1997) model for calculation VaR - and in that connection address some of the issues that has been stated in litterature, as explained by the end of section 7.2.

Before we show the results we need to address one more issue - namely the building of a yield-curve given the price of zero-coupon bonds at the following fixed maturity dates: [0.083,0.25,0.5,1,2,3,4,5,7.5,10,12.5,15,20,25,30].

For that purpose we are using the maximum smoothness approach of Adams and Deventer (1994), see Appendix C for an elaboration.

The results are stated in the table 3 below:

Table 3: Comparison of scenario simulation vs Monte Carlo simulation

|	MC(250.000)	States(9,5,3,3)	States(7,5,3,3)	States(9,7,5,3)	States(9,5,3)
Sample	250000	405	315	945	135
Mean	-0,012	-0,013	-0,013	-0,013	-0,011
Standard Deviation	0,956	1,651	1,450	1,652	1,622
Maximum	4,068	3,005	2,622	3,053	2,645
Minimum	-4,519	-3,299	-2,851	-3,371	-2,858
	States(7,5,3)	States(9,5)	States(13,11,7,5)	States(7,7,7.7)	
Sample	105	45	5005	2401	
Mean	-0,011	-0,011	-0,014	-0,014	
Standard Deviation	1,417	1,634	2,021	1,517	
Maximum	2,257	2,642	3,967	3,062	
Minimum	-2,415	-2,853	-4,539	-3,420	

From the results in table 3 we conclude the following:

- The scenario simulation procedure generates a mean that is very close to the mean derived from the Monte Carlo simulation - this is the case for all the 8 scenario simulations
- The standard deviation in the scenario simulations is in all cases higher than the standard deviation obtained in the Monte Carlo simulation
- We also observe that the range of obtained values (the distance between maximum and minimum) is higher in the Monte Carlo simulation compared to the scenario simulations

Before we start investigating the VaR numbers we have below in figure 5 make a comparison

between the True CDF (as obtained from Monte Carlo) and JZ Approximation.

Figure 5:   Comparison of the True CDF and the JZ Approximation
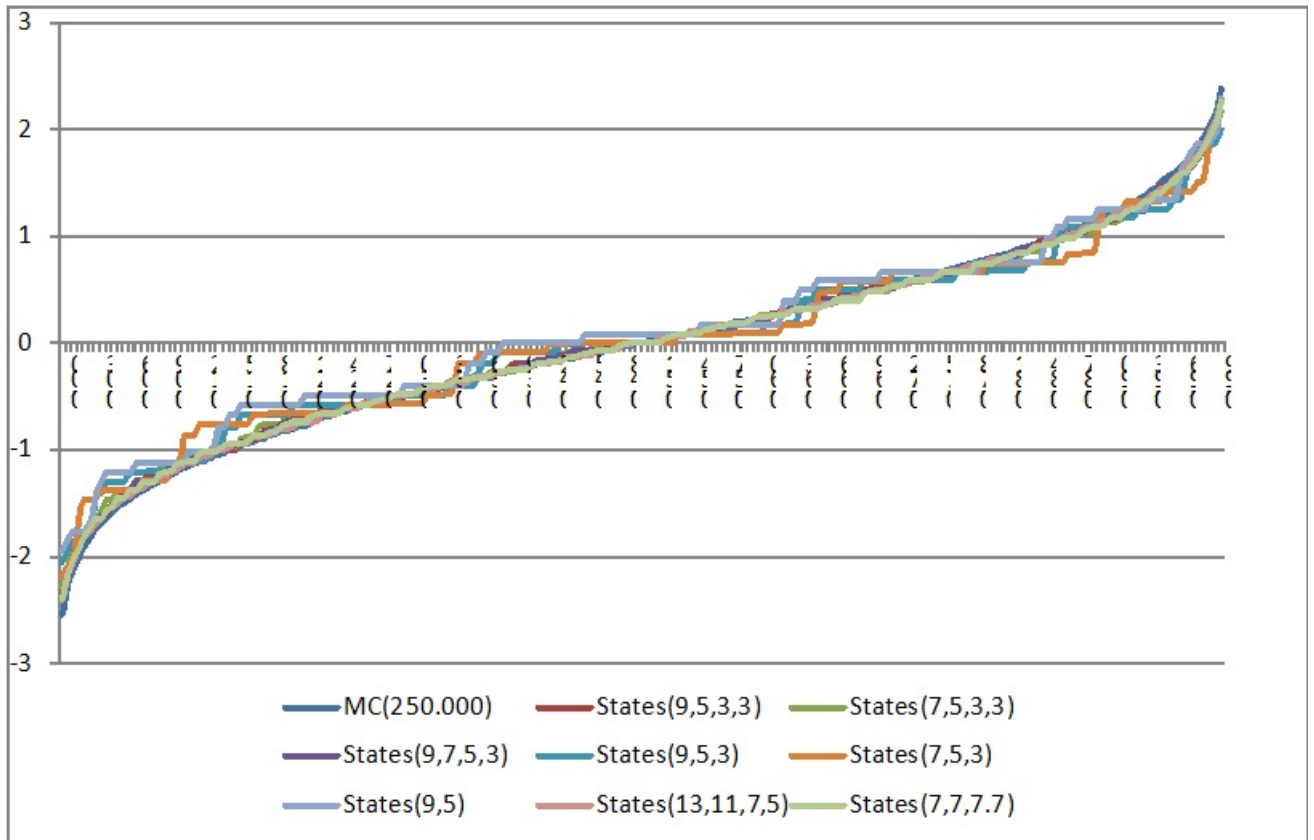
The main conclusion from this figure is the following:

- The less number of factors we select the more oscillation we see
- The less number of states for the factors the more oscillation we see
- We can also deduce that the scenario simulation will under estimate the VaR number (this is in line with the results from Gibson and Pristker (2000)) - the question is how much?

Knowing the probability distribution makes it straightforward to calculate VaR/ETL. This is done below in table 4.

Table 4: A comparison of Risk-Measures

	MC(250.000)	States(9,5,3,3)	States(7,5,3,3)	States(9,7,5,3)	States(9,5,3)
VaR(99) 10-Day	-2,297	-2,200	-2,107	-2,215	-1,965
ETL(99) 10-Day	-2,647	-2,665	-2,654	-2,554	-2,978
	States(7,5,3)	States(9,5)	States(13,11,7,5)	States(7,7,7.7)	
VaR(99) 10-Day	-2,114	-1,873	-2,210	-2,191	
ETL(99) 10-Day	-2,631	-5,173	-2,551	-2,453	

From table 4 we find:

- Yes, VaR will be underestimated using the JZ simulation technique - in our case with an error that lies between 3-17%
- A big surprise is that, as a rule of thumb the ETL estimate is significant better than the VaR estimate. The only exception is the much higher ETL for the case: States(9,5)

In interesting question is how can that be?

In the MC method each simulation has the same probability for occuring, that is not the case in the JZ approach - here the probability for any given observation is defined as the product of being in a given state for each of the factors.

What happens with the JZ in case there is not "enough" states is that some of the observations in the tails will be derived from products of probabilities for being in states for one or more factors where the state is not "far enough" away - in terms of standard deviations - from the centre. In connection to this is that we are bound to get quite a few observations in the tails and that the profit loss distribution will oscillate[22].

To conclude how many states we need to get reliable ETL estimates is real hard to say, but it seems like if we; 1) select a suitable number of states for each factor; 2) that all factors are taken into account and 3) that our sample size is significantly larger than 1/(1-quantile) then we will get fairly reliable ETL estimates[23].

To perform a detailed analysis of this interesting subject lies however outside the boundaries of this paper, we will for the current analysis assume this and will for the rest of the paper be using the JZ model with States(7,5,3,3).

The analysis performed in this section gives rise to the following conclusion:

- The scenario simulation procedure is computationally very superior to the

---

[22] We did some more calculations in order to investigate this phenomenon (like for example States(7,5)) and we see the same problem in all the cases - namely "much" to high ETL number.

[23] A similar observation was done in connection with figure 5.

Monte Carlo simulation procedure
- It is easier with scenario simulation to obtain good estimates for ETL than for the point estimate - VaR. Which is good news as ETL also is a better risk-measure than VaR - see section 7
- There is also evidence that it is appropriate to select 7 x 5 x 3 x 3 states - at least in our case

Some closing remarks.

The scenario simulation procedure is so fast for simple (nearly) linear instruments that it rivals the parametric approach.

Actually to derive the return distribution using the scenario simulation for the case States(7,5,3,3) takes less than 0.1 second all in all, where this includes:

- Determining the states
- Calculating the probability for realising a particular yield curve
- Calculating the 315 yield curves for the 15 maturities[24]
- Estimating 315 yield curves using the maximum smoothness approach of Adams and Deventer (1994)
- Calculating the price distribution of each of the 3-bonds in the portfolio
- Calculation the mean, standard deviation etc
- Calculating VaR/ETL for a given confidence interval

From this we conclude that there is no reason why we should be satisfied with a single number from a risk-management point of view - when we can construct the whole probability distribution with very little effort from a computational perspective.

7.5   Scenario Simulation for Non-linear Portfolios

Of course scenario simulation is even more important for non-linear securities - or more precisely to limit the number of calculations.

We will for that reason turn our attention to some interesting non-linear securities - namely Danish Mortgage-Backed-Bonds[25].

Here we will not focus on a portfolio but instead select a few different MBBs and then use the scenario simulation procedure to calculate the probability distribution for these securities.

---

[24] One might note here, in case we had not been looking at such an extreme time-period as we did in the analysis here, the general result would have been that we only would need somewhat around 3-factors and the most appropriate number of states being 9 x 5 x 3 - these results is because of lack of space omitted here.

[25] For a good introduction to Danish Mortgage-Backed-Bonds, see Karner, Kelstrup and Schelde (1998) and Kelstrup, Madsen and Rom-Poulsen (1999).

However, expanding to a portfolio approach is straightforward as the probability distribution of a portfolio is just defined as the value-weighted sum of the probability distributions for each of the securities in the portfolio.

For the purpose of the following discussion we assume that a pricing model is available. We will however here mention a few of the ingredients that are an integrated part of a pricing model for Danish MBBs (and for that matter American MBBs):

- A yield-curve
- A prepayment model for the behaviour of the debtors - we are here using the Madsen (2005) model
- A yield-curve model, we will here be using the same swap-curve model as introduced in section 5
- A model for the volatility structure, we are implying the volatility parameters from traded caps/floors
- A pricing model that uses the first four ingredients to calculate the price - we are here using the Black and Karisinski (1991) model (Log-Normality assumption) , though calibrated using the trinomial approach of Hull and White (1994)
- An OAS-Model, that is a model that takes into account that OAS is a function interest rate changes[26]

As far as we know VaR for Danish MBBs has only been considered by Jacobsen (1996), his method is however completely different than the methodology we employ in this paper.

Let me for that reason briefly mention a few stylized facts about Jacobsens approach:

The main idea in Jacobsen is to construct a delta equivalent hedge portfolio of zero-coupon bonds - that is a 1 order approximation. For that purpose he selects 17 key-rates[27] and using the triangular method of Ho (1992) it is possibly to construct 17 different shifts to the initial yield-curve - where these shifts by definition are constructed in such a manner that the sum of the shifts equals an additive shift to the yield-curve. This delta equivalent hedge portfolio is now being used to calculate VaR measures using the parametric approach suggested in RiskMetric[TM] for linear securities - that is the delta equivalent cash-flow (DECF) is being treated as a straight bond.

This approach is very simple and the computational burden is acceptable - "just" 17 price calculations - but there are some undesirable features in the methodology:

- Firstly, the DECF is not constant, it depends on all the information that is

---

[26] It lies outside the boundary of this paper to come into details about our MBB pricing model, and associated models, like our OAS-Model.

[27] The reason being that this is the number of maturities that RiskMetric[TM] operates with.

        related to the pricing of MBBs (this is also pointed out by Jacobsen). This means that it is only useful for calculating VaR over short periods and/or limited yield-curve changes

- Secondly, it does not produce a probability distribution - which especially for non-linear securities is of importance. Ofcourse we can construct a probability distribution using the DECF approach - but because the DECF is not constant - this probability distribution will not be of limited use

It could of course be argued that in the scenario simulation procedure advocated here we need to perform 315 price calculations in order to derive the probability distribution - which however still could be quit time-consuming for large portfolios of MBBs - at least relatively to the DECF approach from Jacobsen.

8. Backtesting of our Model Setup

In this section we will perform a backtest of out model setup for the period 2. January 2008 - 2. January 2009.

We will test the abilities of our model setup to capture the "true" risk for the following 2 bonds.

Table 6: **Backtest Portfolio (Date: 2. January 2008)**

sin	Instrument-Name	Clean Price	Accrued IR	Dirty Price
DK0009769895	6 NYK MBB 2041	98,800	0,099	98,899
DK0009918138	7 Stat Ink 2024	127,312	1,109	128,421

In figure 6 below we have shown the price pattern for the 2 above bonds over our backtest period.

Figure 6: **Price evolution for the period 2. January 2008 - 2. January 2009**
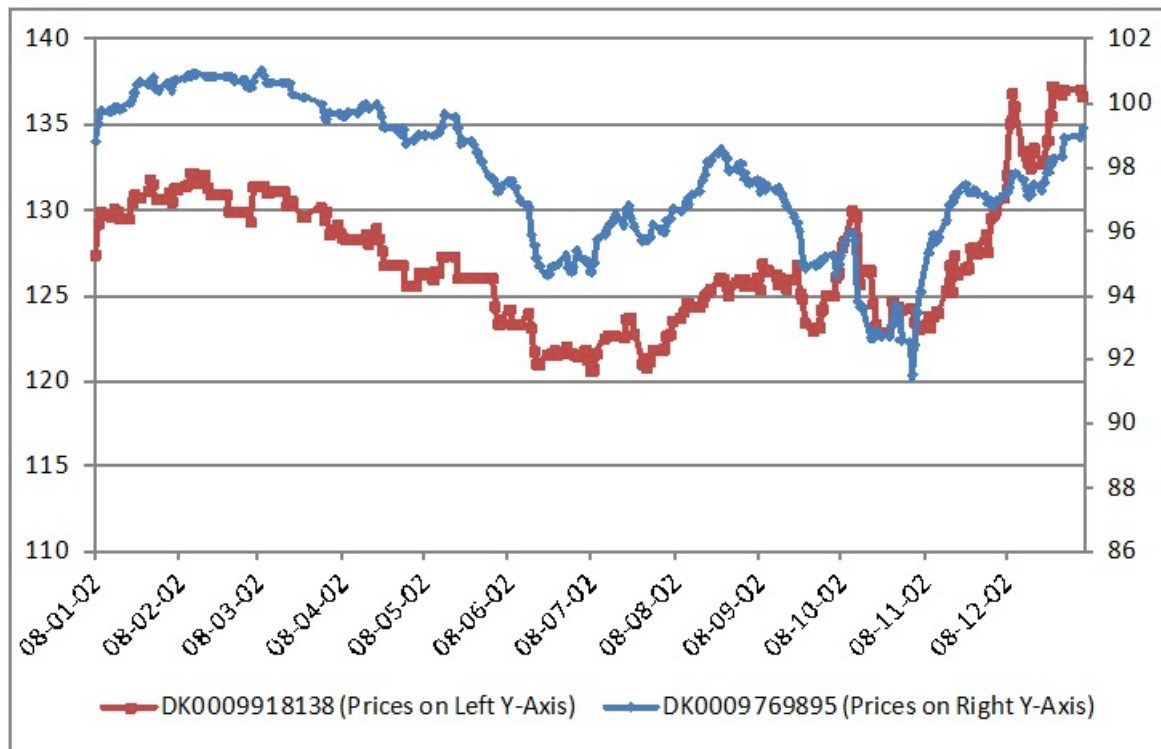
We have performed the following calculations:

Starting with the 2. January 2008 we have every 10-days for the whole year 2008 done the following:

- Re-Estimated the PCA model, assuming 4-factors are to be selected and rotated the factor loadings using the Varimax-Method
- Re-Estimated our Garch(1,1) model and forecastet the volatility structure for the next 10-days
- Generate our 315 Yield-Curves using the JZ Approach with states(7,5,3,3) (as described in the previous sections)
- Given the simuleted 31 Yield-Curves we will now price our 2 Bonds 10-day in the future.
  - For the Government Bond that is straightforward as the value is only a function of the yield-curve
  - For our Mortgage Backed Bond, things are somewhat more complicated. The procedure is as follows:
    - We will first estimate the volatility parameters for our Log-Normal Hull-White model using at-the-money Caps, and assume the estimated volatility is constant over the next 10-days
    - Using the estimated volatility parameters and each of the 315 (simulated) yield-curves as inputs, we will now price the MBB 10-day in the future, using our one-factor Log-Normal Hull-White model combined with our

Prepayment-Model
- Knowing the 315 simulated values (prices) of our 2-Bonds 10-day in the future, will allows us to calculate the Profit-Loss distribution, as we know the value of the Bonds at the beginning of the 10-day period (=today)
- From the Profit-Loss Distributions determine VaR and ETL

These calculated risk-measures we wish to compare with the real-changes in the value of these bonds - especially will it be interesting to see how the model has performed during the Financial Crisis - notable autumn of 2008.

These data are shown in the following 2 figures.

Figure 7:   **Profit-Loss, VaR and ETL (DK0009918138)**
                        (2. January 2008 - 2. January 2009)
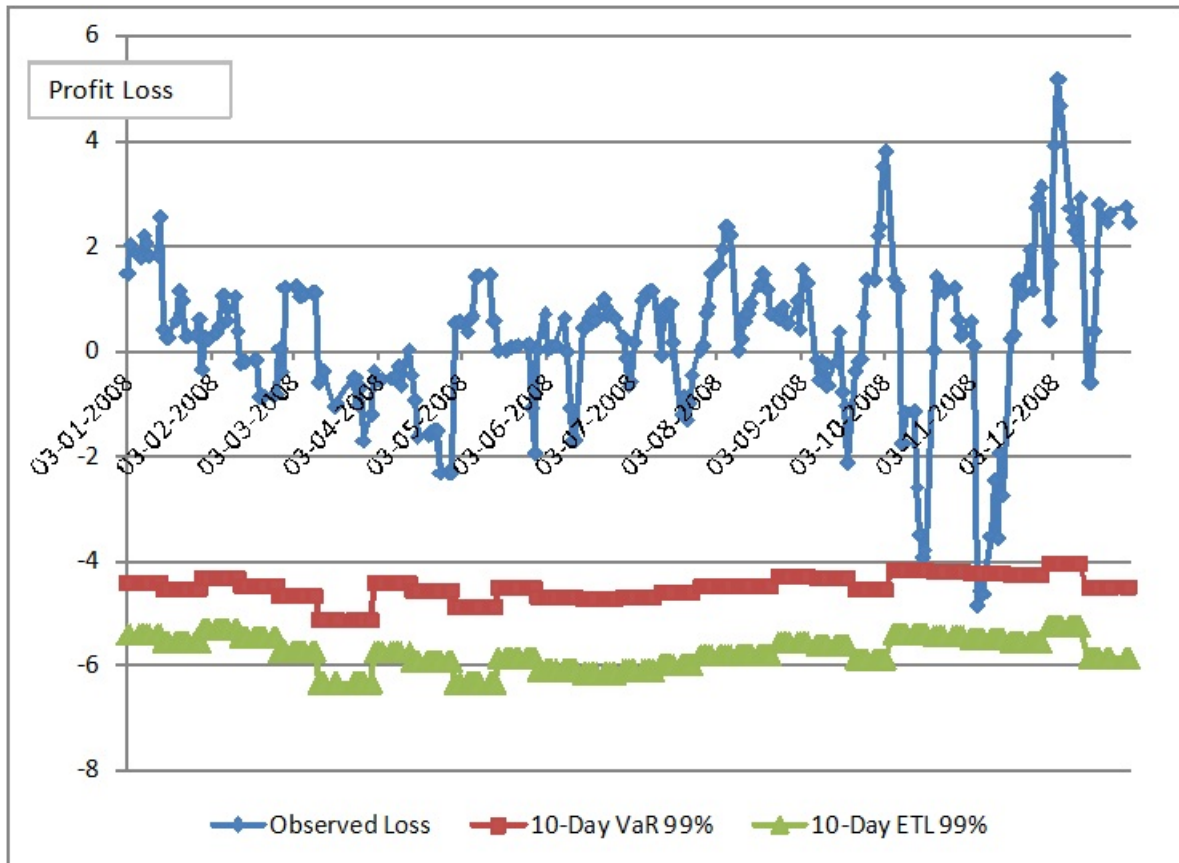
Figure 8: **Profit-Loss, VaR and ETL (DK0009769895)**
(2. January 2008 - 2. January 2009)
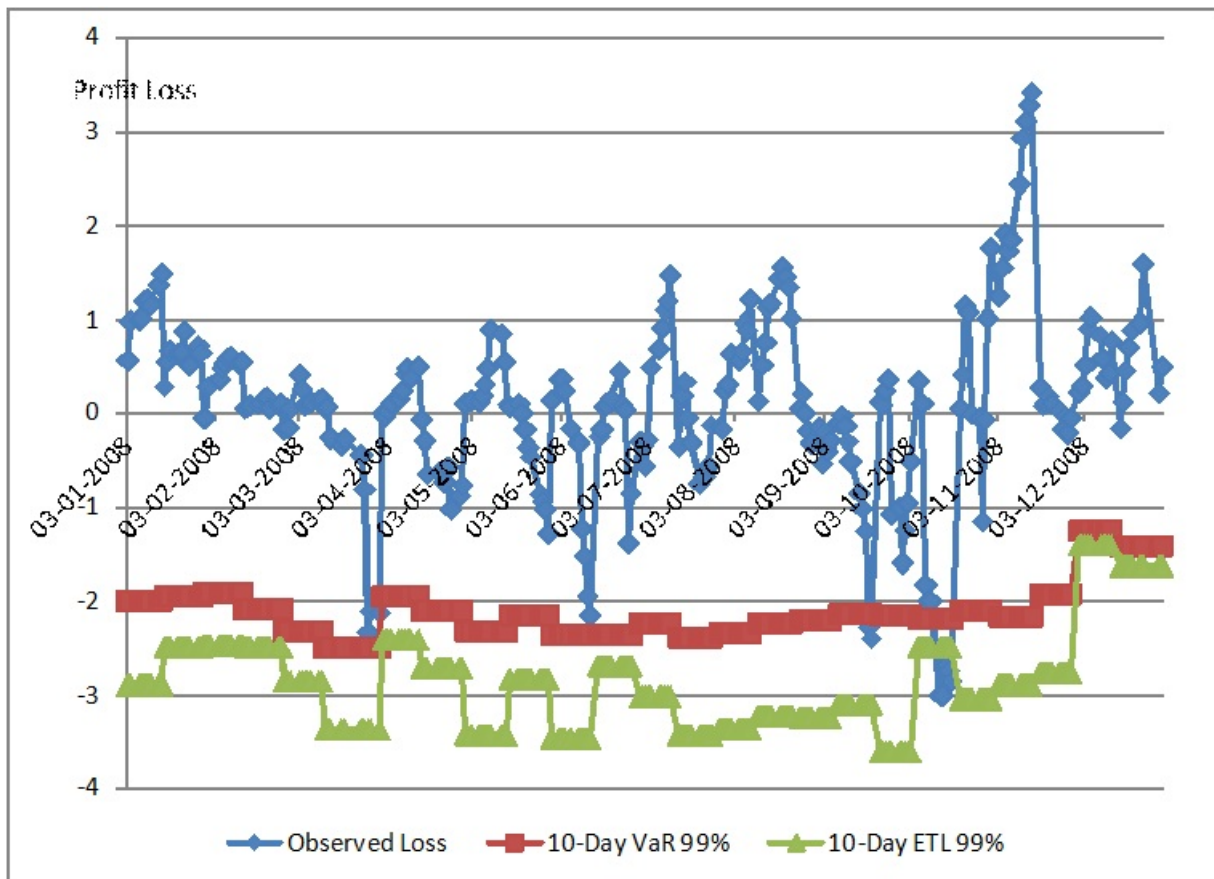
From the above 2 figures we can draw the following conclusion:

- Number of breeches of our VaR number for the Government-Bond is 3 - expected breeches is approximately 2.5
    - The 3 breeches occur on the following 3 dates:
        - 5. November, 6. November and 7. November
    - The highest breech is less then 15% worse than anticipated by our VaR number
- There are no breeches of our ETL number for the Government-Bond
- Number of breeches of our VaR number for the Mortgage-Bond is 7 - expected breeches is approximately 2.5
    - The breeches occur on the following 7 dates:
        - 18. September, 19. September, 13-17 of October
    - The highest breech is approximately 37% worse than anticipated by our VaR number
- There are 5 breeches of our ETL number for the Mortgage-Bond
    - The breeches occur on the following 5 dates:
        - 13-17 of October
    - The highest breech is approximately 20% worse than anticipated by our ETL number

If we consider the results for our Government-Bond then we obtain very interesting results, one might even say that it is incredible that our model setup is so robust that under such extreme price movements we were witness to during the autumn of 2008.

The result obtained here is a strong indication that the model setup as defined in this paper is of practical interest due to its ability to capture the risk under extreme market conditions.

When we turn our attention to the results for the Mortgage-Bond, the results are not nearly as strong as for the Government-Bond. However, considering the fact that there are a lot more in play in modelling a Mortgage-Bond, it is not really that surprising - where the most notable problem during the financial crises was the following 2 facts:

- The market-marker arrangements for Mortgage-Bonds was suspended on several occasions during this period, which means that no real trading was done
- The was a huge changes in OAS-Spread - which no OAS-Model was able to capture, this can be seen in figure 9 below

Figure 9: Evolution in OAS-Spread for DK0009769895
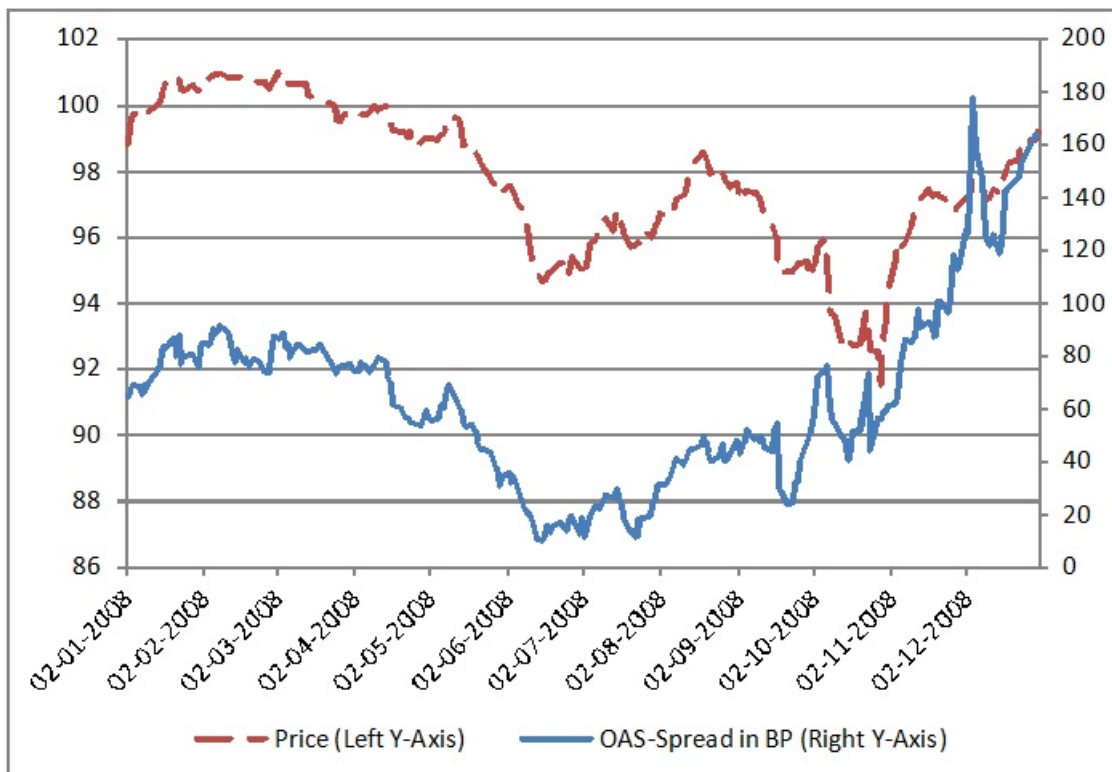
If one takes that into account, we believe the results we obtain in our model setup also is very promising in connection with Danish MBB's.

8. Conclusion

The first result in this paper was the construction of a general model for the variation in the term structure of interest rates - that is we defined a model for the shift function. In this connection we showed - using the Heath, Jarrow and Morton (1991) framework - that the shift function could be understood as a volatility structure - more precisely the spot rate volatility structure.

The class of shift functions considered in this paper was of the linear type, with independence between the individual factors; the model was therefore comparable to the Ross (1976) APT model.

Using PCA we showed that it took a 4-factor model to explain the variation in the term structure of interest rates over the period from the 2. January 2002 to the 2. January 2012. These 4 factors can be called a Slope factor, a Short-Curvature Factor, a Short Factor and a Curvature Factor.

Because of the relationship between the volatility structure in the Heath, Jarrow and Morton framework and the shift function implied from our empirical analysis of the evolution in the yield-curve we concluded that PCA could be used to determine the volatility structure in the Heath, Jarrow and Morton framework - a non-parametric approach.

We then extended this model setup by adding stochastic volatility. This we did by assuming that the stochastic volatility was driven by a GARCH(1,1) model. Due to the fact that our stochastic volatility process was independent of our yield-curve factor-model, means that the extended model belonged to the class of USV (unspanned stochastic volatility) models.

In the last part of the paper we turned our attention to Risk model. Our approach to the calculation of VaR/ETL was a scenario simulation based methodology with relied on the framework of Jamshidian and Zhu (1997). This scenario simulation procedure builds on the factor loadings derived from a PCA of the same kind we used in our analysis of the empirical dynamics in the yield-curve.

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

From our analyses of both straight bonds and Danish MBBs using the scenario simulation procedure we conclude the following:

- The scenario simulation procedure is a computational effective alternative to Monte Carlo simulation
- The scenario simulation procedure is capable of producing reasonable good approximations of the probability distributions with a limited number of states

- There is much better control over the extreme values in the scenario simulation than in Monte Carlo simulation
- The scenario simulation procedure is more efficient than the parametric approach for "linear" securities and it rivals the parametric approach for "linear" securities because of the speed of calculation
- We suggested using 7 x 5 x 3 x 3 states for the scenario simulation. Because we "only" need to perform 315 re-evaluations of the portfolio the scenario simulation procedure is feasible for large portfolios consisting of highly complex non-linear securities
- One last important feature is that we in the spirit of Barone-Adesi, Bourgoin and Giannopoulos (1997) extended the scenario simulation procedure to include stochastic volatility - in our case we showed that a Garch(1,1) was appropriate

We also backtested our Risk-Model setup for the year 2008 on 2 bonds - a Government-Bond and a Mortgage-Bond. The overall conclusion for this was:

- Our Risk-Model setup was able to capture the extreme movements influenced by the financial crisis for our Government-Bond case - notable the autumn of 2008
- Our Risk-Model setup did not manage to completely capture the extreme price movements for our Mortgage-Bond. Due to the fact that the reason for that was mostly (if not entirely) because of the enormous changes on level of OAS, we found the model indeed very promising - both of term of efficient from a calculation point of view, to its ability to capture the Risk of Mortgage-Bonds

More work is of course necessary - both with respect to backtesting the model and with respect to the determination of the input to the scenario simulation procedure.

In connection with the inputs to the scenario simulation method we especially need to address the following:

- How to forecast the volatility?
  - In this paper we argued that GARCH is a logical method to utilize
- How to select the expected yield-curve at a given time-horizon?
  - In this paper we suggested using the initial yield-curve as the expected yield-curve
- Would it be possible to make our OAS-Model even more robust?
- More examination on the appropriate number of states that is used in the scenario simulation method

Literature

Adams and Deventer (1994) "Fitting Yield Curves and Forward Rates with Maximum Smoothness", Journal of Fixed Income vol. 4, no. 1, page 52-62

Alexander (1996) (editor) "Risk Management and Analysis", John Wiley and Sons 1996

Amin and Morton (1993) "Implied Volatility Functions in Arbitrage Free Term Structure Models", Working Paper University of Michigan, May 1993

Barrett, Heuson and Gosnell (1992) "Yield Curve Shifts: An Empirical Solution to a Theoretical Dilemma", Working Paper, University of Miami, October 1992

Baron (1989) "Time Variation in the Modes of Fluctuation of the Treasury Yield Curve", Bankers Trust, October 1989

Barone-Adesi, Bourgoin and Giannopoulos (1997) "A Probabilistic Approach to Worst Case Scenarios", working paper University of Westminster, August 1997

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

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

Bierwag, Kaufmann and Toevs (1983) "Innovation in Bond Portfolio Management", JAI Press Inc.

Bierwag (1987a) "Duration Analysis Managing Interest Rate Risk", Balling Publishing Company 1987

Bierwag (1987b) "Bond Returns, Discrete Stochastic Processes, and Duration", Journal of Financial Research 10, pp. 191-209

Bierwag, Kaufmann and Latta (1987) "Bond Portfolio Immunization: Tests of Maturity, One- and 2-factor Duration Matching Strategies", Financial Review 22, pp. 203-219

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

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

Boudoukh. Richardson and Whitelaw (1997) "Investigation of a Class of Volatility Estimators", Journal of Derivatives vol. 4, no. 3, page 63-71

Casassus, Collin-Dufresne and Goldstein (2005) "Unspanned stochastic volatility and fixed income derivatives pricing", Journal of Banking & Finance 29 (2005) 2723–2749

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

Chambers, Carleton and Waldman (1984) "A New Approach to Estimation of the Term Structure of Interest Rates", Journal of Financial and Quantitative Analysis, Vol. 19, pp. 233-252

Chambers and Carleton (1988) "A Generalized Approach to Duration", Research in Finance, Vol. 7, pp. 163-181

Chambers, Carleton and Mcenally (1988) "Immunizing Default-Free Bond Portfolios with a Duration Vector", Journal of Financial and Quantitative Analysis, pp. 89-104

Christiansen (1998) "Value at Risk Using the Factor-ARCH Model", working paper D 98-6, Department of Finance, The Aarhus School of Business

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

Cox and Rubinstein (1985) "Options Markets", Englewood Cliffs: Prentice Hall 1985

Dahl (1989) "Variations in the term structure of interest rates and controlling interest-rate risks" [in Danish], Finans Invest, February 1989

Elton, Gruber and Nabar (1988) "Bond Returns, Immunization and The Return Generating Process", Studies in Banking and Finance 5, pp. 125-154

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

Engle and Ng (1991) "Measuring and Testing the Impact of News on Volatility", working paper University of California

Engle, Ng and Rothschild (1990) "Asset Pricing with a Factor-ARCH covariance structure", Journal of Econometrics 45, page 213-237

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

Garbade (1989) "Polynomial Representations of The Yield Curve and its Modes of Fluctuation", Bankers Trust, July 1989

Giannopoulos (1995) "The Theory and Practice of ARCH Models", working paper University of Westminster, 27 November 1995

Glosten, Jagannathan and Runkle (1991) "Relationship Between the Expected Value and the Volatility of the Nominal Excess Return on Stocks", working paper Northwestern University

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

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

Hald (1979) "Updating formulas in Quasi-Newton methods" (in Danish), Department for Numerical Analysis, report no. NI-79-03

Harman (1967) "Modern Factor Analysis", University of Chicago Press, 1966

Harvey, Ruiz and Shepard (1994) "Multivariate Stochastic Variance Models", Review of Economic Studies, 61, page 247-264

Heath, Jarrow and Morton (1990) "Contingent Claim Valuation with a Random Evolution of Interest Rates", The Review of Futures Markets 9, pp. 54-76

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

Ho (1992) "Key Rate Durations: Measures of Interest Rate Risks", Journal of Fixed Income, September, page 29-44

Hull and White (1994) "Numerical Procedures for Implementing Term Structure Models", Working paper University of Toronto, January 1994

Ingersoll, Skelton and Weil (1978) "Duration forty years later", Journal of Financial and Quantitative Analysis, November, pp. 627-652

Jacobsen (1996) "Value-at-Risk for Danish Mortgage-Backed-Bonds" (in Danish), Finans/Invest 1/96, page 16-21

Jamshidian and Zhu (1997) "Scenario Simulation: Theory and methodology", Finance and Stochastic vol. 1, no. 1, page 43-67

J. P. Morgan (1995) "RiskMetric[TM] - Technical Document, third edition", New York 26 May

J. P. Morgan (1997) "RiskMetric[TM] - Monitor, fourth quarter 1997", New York 15 December

Karner, Kelstrup and Schelde (1998) "A Guide to the Danish Bond and Money Market", February 1998, Handelsbanken Markets, Reference Library no. 10

Kelstrup, Madsen and Rom-Poulsen (1999) "A Guide to the Danish Mortgage Bond Market", Primo 1999, Handelsbanken Markets, Reference Library no. 18

LeBaron (1994) "Chaos and Nonlinear forecastability in Economic and Finance", working paper University of Wisconson, February 1994

Litterman and Scheinkman (1988) "Common factors affecting bond Returns", Goldmann, Sachs & Co., Financial Strategies Group, September 1988

Longstaff and Schwartz (1991) "Interest-rate volatility and the term structure: A two-factor general equilibrium model", Working Paper, Ohio State University, November 1990

Lord and Pelsser (2007) "Level-Slope-Curvature - Fact or Artefact?", Applied Mathematical Finance, Vol. 14, No. 2, 2007

Madsen (1997) "Determination of the implicit credit spread on Danish Mortgage-Backed Securities (MBS)", Risk Magazine December 1997

Madsen (1998) "The Modelling of Debtor Behaviour - Mortgage Backed Securities I", February 1998, Handelsbanken Markets, Reference Library no. 13

Madsen (1995) "Danish Mortgage Modeling in FinE", working paper 17. June 2005, http://www.fineanalytics.com/images/stories/pdf/fine_mbb_modelling.pdf

Mandelbrot (1997) "Fractals and Scaling in Finance - Discontinuity, Concentration, Risk", Springer-Verlag New York Berlin Heidelberg

McDonald (1996) "Probability Distributions for Financial Models", in Maddala and Rao, eds., Handbook of Statistics vol. 14, page 427-461

Nelson (1992) "Filtering and Forecasting with Miss-specified ARCH Models: Getting the Right Variance with the Wrong Model", Journal of Econometrics, vol. 52, page 61-90

Nelson and Siegel (1987) "Parsimonious modelling of Yield Curves", Journal of Business, October, pp. 473-489

Pedersen (1996) "Forecasting of interest rate volatility (3) - an empirical comparison of selected methods" (in Danish), Finans/Invest 6/96, page 22-27

Prisman and Shores (1988) "Duration Measures for Specific Term Structure Estimations and Applications to Bond Portfolio Immunization", Journal of Banking and Finance 12, pp. 493-504

Reitano (1992) "Non-Parallel Yield Curve Shifts and Immunization", Journal of Portfolio Management, spring, pp. 36-43

Repplinger (2008) "Pricing of Bond Options - Unspanned Stochastic Volatility and Random Field Models", Springer-Verlag, 615 Lecture Notes in Economics and Mathematical Systems

Risk Books - editor Jarrow (1998) "Volatility - New Estimation Technique for Pricing Derivatives", Risk Publications

Ross (1976) "The Arbitrage Theory of Capital Asset Pricing", Journal of Economic Theory 13, pp. 341-360

Rubenstein (1984) "A simple Formula for the Expected Rate of Return of an Option over a Finite Holding Period", Journal of Finance, Vol. 39, pp. 1503-1509.

Sentana (1991) "Quadratic ARCH Models: A Potential Re-Interpretation of ARCH Models", working paper London School of Economics

Shea (1985) "Term Structure Estimation with Exponential Splines", Journal of Finance vol. 40, page 319-325

Steeley (1990) "Modelling the Dynamics of the Term Structure of Interest Rates", The Economic and Social Review, Vol. 21, no. 4, pp. 337-361

Svensson (1994) "Estimating and Interpreting Forward Interest Rates: Sweden 1992-1994, IMF Working Paper, WP/94/114 (September 1994), 1-49

Tanggaard (1991) "A model for bond portfolio choice" [in Danish], Working Paper, the Århus Business School, January 1991

Tanggaard (1997) "Nonparametric smoothing of yield curves". Review of Quantitative Finance and Accounting 9, page 251-267

Trolle and Schwartz (2008) "Unspanned stochastic volatility and the pricing of commodity derivatives", The Review of Financial Studies / v 22 n 11 2009

Trzcinka (1986) "On the Number of Factors in the Arbitrage Pricing Model", Journal of Finance, Vol. XLI, pp. 347-368

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

Webber and James (2000) "Interest Rate Modelling" Wiley, 2000

Wilson (1994) "Plugging the Gap", Risk Magazine, vol. 7, no. 10, October

Zakoian (1991) "Threshold Heteroskedastic Model", working paper, INSEE, Paris

Appendix A

In principle, the literature contains three different ways in which it has been proposed to extend the traditional 1-factor duration model so that not only additive shifts in the term structure of interest rates are considered. In this connection, we have decided not to include the literature on stochastic term structure models.

The first class of models consists of the so-called multi-factor duration models, where they can be formulated as follows if we consider a 2-factor duration model:

$$PR_i = a_i + b_{i1}F_1 + b_{i2}F_2 + e_i \tag{30}$$

Where PR_i is the return of the i'th bond, a_i is the risk-free return, b_{i1} and b_{i2} could be defined as the modified duration of a short bond and a long bond, respectively. In addition, F_1 and F_2 are defined as factors that relate to the first and second risk factors, which means that ε_i is the i'th bond residual. An alternative formulation could be to let b_{i2} be defined as a spread duration. With respect to duration models, the situation is that these factor values (F_1 and F_2) each represent the relation between changes in these basic rates (or basis spreads) and the term structure of interest rates itself; see, for instance, Ingersoll (1983).

Models that can be classified as being of this type have been defined by Elton, Gruber and Naber (1988), Reitano (1992), Bierwag (1987b) and Bierwag, Kaufmann and Latta (1987), and moreover Bierwag, Kaufman and Toevs (1983).

The other type of model is based on the functional form used in the estimation of the term structure of interest rates. The first model introduced in this class was Chambers, Carleton and McEnally (1988)[28]. The fundamental condition underlying these models is the fact that the functional form of the term structure of interest rates has been chosen in such a way that the unknown parameters that determine the term structure estimation are independent. One obvious functional form of the term structure that complies with this requirement is polynomial models, as follows:

$$P(t,T) = e^{-\sum_{j=0}^{n} a_j \tau^{j+1}} \tag{31}$$

Where j is the polynomial degree, n is the maximum degree, a_j is the j'th coefficient and τ is the term to maturity, i.e. $T - t$[29]. From this can be clearly seen that there is independence between the individual unknown parameters, each single a_j. If a second-degree polynomial is considered, then the impacts on the term structure of interest rates can be seen to be defined by

---

[28] In addition, reference can be made to Chambers and Carleton (1988), Prisman and Shores (1988), Steeley (1990) and Barret, Heuson and Gosnell (1992).

[29] This formulation can be seen to be identical to Chambers, Carleton and Waldmanns (1984) model, and was in fact tested on the Danish market by Tanggard and Jacobsen in a number of working papers in the late 1980's.

an additive factor, a term structure slope factor and a term structure curvature factor.

The risks of interest rate contingent claims as a function of changes in these coefficients can then to be measured, as these parameters uniquely determine the term structure of interest rates. At this point, it should be mentioned that Steeley (1990) used a cubic-spline model as his starting point, whereas Barret, Heuson and Gosnell (1992) use Nelson and Siegel's (1987) model. The fact that Nelson and Siegel's model also fulfils the condition regarding independence between the individual unknown parameters is not entirely the case, only to the extent that the τ-parameter is determined explicitly.

However, note 2 from Chambers, Carleton and McEnally (1988) demonstrated, in connection with the duration vector model, that the underlying principles of this model are only dependent on the fact that changes in the term structure of interest rates can be described by a polynomial of the n'th degree.

Where this leads us on to the third class of models, namely models that take the term structure of interest rates for given and explicitly specify some sort of functional form to describe the dynamics. At this point Garbade (1989) can be mentioned, which assumes that impacts on the term structure of interest rates are defined by a polynomium of the n'th degree. If this is formulated on the basis of the shift function, then this model can be expressed as follows:

$$S(t,T) \;=\; \sum_{j=0}^{n} a_j \tau^{j+1} \tag{32}$$

Or to put it differently, the spot rate volatility structure can be described by a polynomial of the n'th order. Where it is this last formulation of the shift function that is closest to the model constructed in section 3 in the text. On the basis of formula 3, it can be seen that the traditional 1-factor duration model is achieved for $n = 0$, i.e. the impacts on the term structure of interest rates are of the order a_0.

Appendix B

Some Statistics for the Relative Log-Returns

Table 1:   **(period: 2 January 2002 - 2 January 2012)**[30]

Maturity	0,083	1	2	5	10	20	30
Centered Mean	-0,04752	-0,04909	-0,05439	-0,04504	-0,03327	-0,03019	-0,03272
Centered Variance	24,39939	12,14947	3,42270	3,09164	1,43235	1,47645	1,75291
Centered Skewness	0,06600	-0,10737	0,41372	-0,53219	-0,25574	-0,38302	-0,77399
Centered Kurtosis	23,86390	34,53680	22,97702	26,91020	7,49355	13,17215	12,65991
Auto-Correlation 1	-0,39468	-0,39681	-0,05862	-0,14179	-0,00966	-0,02977	-0,00274
Box-Ljung 1	387,56906	391,74695	8,55047	50,02209	0,23223	2,20483	0,01865
Chi-Square 1	3,84315	3,84315	3,84315	3,84315	3,84315	3,84315	3,84315
Auto-Correlation 5	-0,01558	-0,06142	0,01129	0,07892	0,04331	0,01012	0,01490
Box-Ljung 5	399,11255	423,48431	29,10601	98,69074	17,54937	15,51930	8,62700
Chi-Square 5	11,07050	11,07050	11,07050	11,07050	11,07050	11,07050	11,07050
Auto-Correlation 10	0,01408	0,00389	-0,04445	-0,01647	0,01907	0,01418	-0,01201
Box-Ljung 10	417,93811	456,36793	55,14677	125,32822	27,87941	19,98903	14,43744
Chi-Square 10	18,30705	18,30705	18,30705	18,30705	18,30705	18,30705	18,30705
Auto-Correlation 20	-0,03383	0,02147	0,01308	-0,00758	-0,02019	-0,03006	-0,00193
Box-Ljung 20	429,77804	495,70763	91,63903	195,42823	48,20522	32,54401	31,63084
Chi-Square 20	31,41044	31,41044	31,41044	31,41044	31,41044	31,41044	31,41044
Auto-Correlation 50	-0,02594	0,01039	-0,01013	-0,01163	-0,00842	0,01712	0,01810
Box-Ljung 50	575,22287	616,44750	157,82031	255,73414	86,35404	83,97849	92,86378
Chi-Square 50	67,50061	67,50061	67,50061	67,50061	67,50061	67,50061	67,50061

In general we observe that there is (significant) evidence of GARCH effect especially in the short end of the yield-curve - for short maturities.

In the next section we will briefly introduce the GARCH technique and in that connection specify how we have chosen to forecast the volatility in this paper.

B.1   GARCH Models, A Survey[31]

In theory there is no problem in finding useful estimations for the volatility, since volatility estimates are in principle a reflection of how the variable in question is expected to fluctuate. The direction of movements in the variable is in fact of no importance. The important thing, however, is how the variable fluctuates around its average value.

In practice, the volatility estimation is made complicated by a number of circumstances.

---

[30] Box-Ljung x is the Box-Ljung test statistics for the squared Auto-Correlation x - where x represents the number of lags.
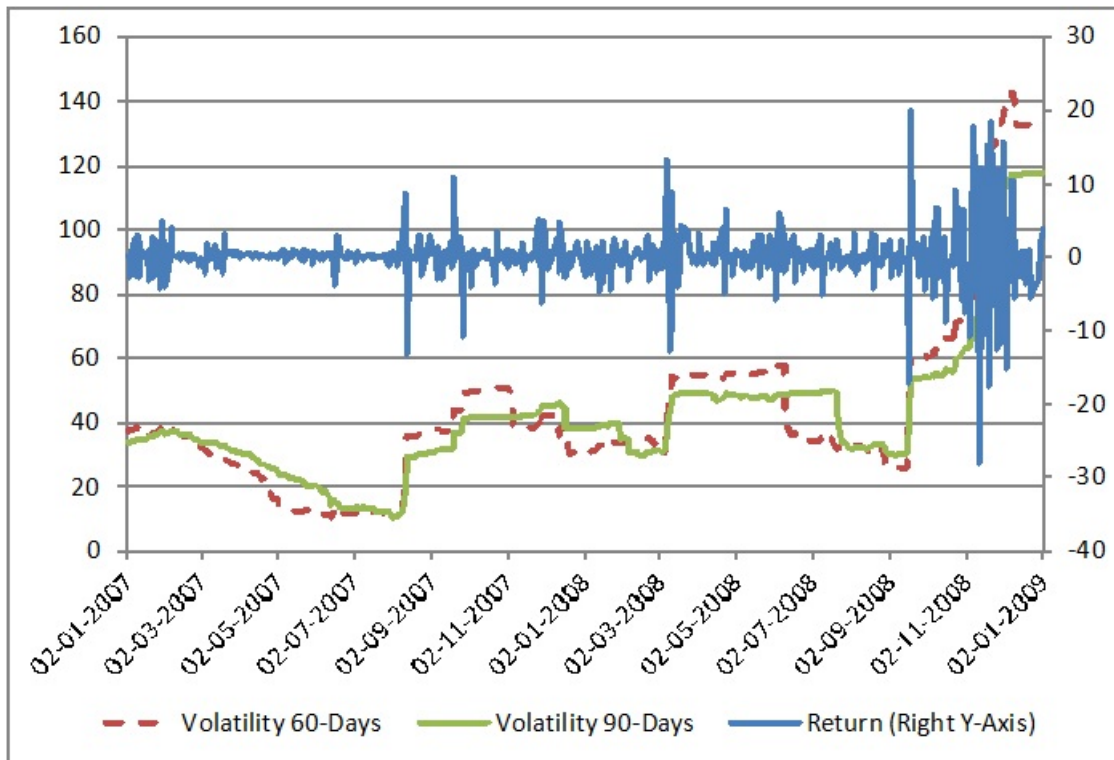
[31] We will not here try to reference the huge amount of litterature on the GARCH framework but instead refer to "Volatility - New Estimation Technique for Pricing Derivatives", Risk Books 1998 and RiskMetrics[TM] - Technical Document (1995,1997). Some particular references will however be given in the section whenever appropriate. On Danish data see in particular Pedersen (1996).

Probably the most significant problem is the instability of the volatility over short time-periods, i.e short estimation periods. On the other hand there is a certain degree of stability when considering longer estimation periods. This means of course that if the volatility changes to a relatively high degree, there is no reason to believe that volatility estimates achieved on the basis of previously observed interest rate movements will reflect future actual interest rate movements.

This has the implication that the assumption of constant volatility over the maturity spectrum is clearly not consistent with the observation that the volatilities which we have observed in the market are not constant.

This is illustrated in figure 1 below for the 1-year zero coupon rate[32]:

Figure 1   Historical Volatility (2. January 2007 - 2. January 2009)

The maximum likelihood estimator is used to calculate the volatility, thus:

$$\sigma = \sqrt{\frac{1}{n-1}\sum_{t=1}^{n} y_t^2}$$
$$for \qquad\qquad (33)$$
$$y_t = \ln\frac{r_{t+1}}{r_t}$$

where r_t is the interest rate at time t, n is the number of observations, σ is the volatility pr. day and y_t represents the return.

From the discussion above we have that if the volatility estimator obtained from formula 33 is to be useful for forecasting purpose, then:

- The interest rate volatility has to be constant
- The interest rates have to be lognormally distributed and for this reason the returns be normally distributed

More precisely we postulate the following model for the returns:

$$y_t = \varepsilon_t \qquad \varepsilon_t \approx N(0,\sigma^2) \qquad\qquad (34)$$

That is the returns are treated as a time series of independent, normally distributed stochastic variables with a constant variance.

Clearly this assumption is not valid - as seen from figure 1. What we instead see is that large changes tend to be followed by large changes - of either sign - and small changes tend to be followed by small changes[33] - of either sign. This is often referred to as the clustering effect.

In general the following observation has been reported in the literature with respect to returns:

- Return distributions have fat tails and higher peak around the mean, that is we can observe excess kurtosis (normal distribution = 3)
- Returns are often negatively skewed (normal distribution = 0)
- Squared returns often have significant autocorrelation

For our data (see table 1) there is evidence of excess kurtosis, positive skewness and autocorrelation in the squared returns across the wbole maturity spectrum - though more pronounced for shorter maturities.

These findings indicates that a normal distribution assumption might not be appropriate when

---

[33] This was first observed by Mandelbrot 1963 - see Mandelbrot (1997).

modelling returns.

The devotion to find more precise estimation techniques has inspired reseach in finance. Some of the most notable approaches can be categorized as follows:

- Introducing other probability distributions, for an overview, see McDonald (1996)
- GARCH[34]-type models, see Bollerslev (1986)
- Stochastic volatility models, see for example Harvey, Ruiz and Shepard (1994)
- Application of chaotic dynamics, see LeBaron (1994)
- MDE (Multivariate Density Estimation), see Boudoukh, Richardson and Whitelaw (1997)
- Jump Diffusion models, see for example J.P. Morgan (1997)

In the academic literature GARCH-models have been the most popular - which is due to the evidence that time series realizations of returns often exhibit time dependent volatility[35]. Because of that we will restrict myself to discussing GARCH-models.

Since the introduction of GARCH in the literature a number of new models in this framework have been developed. To name a few we could mention IGARCH, EGARCH and AGARCH.

Most financial studies conclude that a GARCH(1,1) is adequate, ie:

$$
\begin{aligned}
y_t &= \varepsilon_t \qquad \varepsilon_t \approx N(0,h_t) \\
h_t &= \omega + \alpha\varepsilon_{t-1}^2 + \beta h_{t-1}
\end{aligned}
\tag{35}
$$

In order for a general GARCH process to be stationary we must have that the sum of the roots lies inside the unit circle, more precisely we have the following stationarity condition for the GARCH(1,1) model in formula 35; $1 > \alpha + \beta$.

Using the law of iterated expectation the unconditional variance for a GARCH(1,1) can be expressed as:

$$
E_t[\varepsilon_t^2] = E[\varepsilon_t^2|F_{t-1}] = E[h_t] = \omega + [\alpha + \beta]E[\varepsilon_{t-1}^2]
\tag{36}
$$

which can be recognized as a linear difference equation for the sequence of variances.

Assuming the process began infinitely far in the past with a finite initial variance - then the sequence of variances will converge to a constant:

---

[34] GARCH mean Generalized Autoregressive Conditional Heteroskedasticity - and it represents an extension of the ARCH-model from Engle (1982).

[35] It is of course possible to model time dependent volatility directly using stochastic volatility models and the MDE-approach - but as of now limited investigation is available. It is nevertheless an interesting alternative, which however is left for future research.

$$E_t[\varepsilon_t^2] \;=\; h \;=\; \frac{\omega}{1-\alpha+\beta} \tag{37}$$

which implies $\;1 > \alpha+\beta\;$ in order for h to be finite.

From this we have that even though the conditional distribution of the error is normal then the unconditional distribution is non-normal - which is an very attractive feature of GARCH models.

An important property arises by inspecting equation 37, namely: shocks to the volatility decay at a speed that is measured by $\alpha+\beta$. Furthermore the closer to one (1) $\alpha+\beta$ is, the higher is the persistence of shocks to the current volatility.

For $\alpha+\beta=1$ the shocks to volatility will persist for ever and in this case the unconditional variance is not determined. A process with such a property is known as an IGARCH - Integrated GARCH. It is of special interest to focus a little on the IGARCH since as will become apparent in a moment it is identical to the EVMA model[36] (apart from a constant) that is used in RiskMetric[TM].

The EVMA (Exponentially Weighted Moving Average) places more weight on recent observation - which will have the effect of diminishning the "ghost-features" which are apparent in figure 1. This exponential weighting is done by using a smoothing constant λ - where we have that the larger the value of λ the more weight is placed on past observations.

The infinite EVMA model can be expressed:

$$h_t \;=\; (1-\lambda)\sum_{i=1}^{\infty}\lambda^{i-1}\varepsilon_{t-i}^2 \tag{38}$$

To show the equivalence between the IGARCH and the EVMA model represented by formula 38 we first recall that the IGARCH(1,1) can be written as:

$$h_t \;=\; \omega + (1-\lambda)\varepsilon_{t-1}^2 + \lambda h_{t-1} \tag{39}$$

A repeated substitution in formula 39 yields:

$$
\begin{aligned}
h_t &= \omega + (1-\lambda)\varepsilon_{t-1}^2 + \lambda[\omega + (1-\lambda)\varepsilon_{t-2}^2 \\
&\quad + \lambda[\omega + (1-\lambda)\varepsilon_{t-3}^2 + \lambda[..... \\
&= \frac{\omega}{1-\lambda} + (1-\lambda)[\varepsilon_{t-1}^2 + \lambda\varepsilon_{t-2}^2 + \lambda^2\varepsilon_{t-3}^2 + ...]
\end{aligned}
\tag{40}
$$

---

[36] It is not equal to the EVMA in general but the IGARCH is equal to a particular EVMA model - which happen to be the one used in RiskMetric[TM].

If we compare formula 40 with formula 38 it follows that they are identical apart from a constant.

One very powerful result with respect to GARCH is that GARCH models are not so sensitive (as ARCH-models) to misspecification - because as Nelson (1992) shows, then even if the conditional variance in a linear GARCH model has been misspecified - then the parameters of the model are still consistent. Furthermore the GARCH model is not sensitive to the distribution assumption - that is the parameters are still consistent if there is evidence of non-normality in the squared normalized residuals[37], ie: $\hat{\varepsilon}_t^2 = \dfrac{\varepsilon_t^2}{h_t} \approx N(0,1)$. In this connection it

however worth mentioning that if we decide to estimate a GARCH model under a normal distribution assumption - then we (normally) have to do a (< 99%) fractile adjustment of the original data - this fractile adjustment is however not necessary if we assume for example a t-distribution.

Before returning to the data in table 1, let me briefly mention a few of the non-linear GARCH models and their implications[38]:

Engle and Ng (1991) propose the AGARCH[39] (asymmetric GARCH) which can be expressed as:

$$
\begin{aligned}
y_t &= \varepsilon_t \qquad \varepsilon_t \approx N(0,h_t) \\
h_t &= \omega + \alpha[\varepsilon_{t-1} + \gamma]^2 + \beta h_{t-1}
\end{aligned}
\tag{41}
$$

which has similar properties as the EGARCH from Nelson (1990) - but from an implementaion point of view much simpler. The AGARCH has similar properties as the GARCH model but unlike the GARCH model which explores the magnitude of one-period lag-errors, the AGARCH model allows for the past error to have an asymmetric effect on the variance. For example if γ is negative then the conditional variance will be higher when ε_{t-1} is negative than when it is positive. For $\gamma = 0$ the AGARCH model degenerates to the GARCH model. Because of the γ-parameter the AGARCH model can - like the EGARCH model - capture the leverage effect which has been observed in the stock market. The unconditional variance for the AGARCH model has the following form:

---

[37] An investigation of the autocorrelation in the squared normalized residuals may also reveal model failure. Pagan and Schwert (1990) furthermore proposed to regress the squared normalized residuals against a constant and h_t, ie: $\hat{\varepsilon}_t^2 = a + bh_t + error$. If the forecast is unbiased, a = 0 and b = 1. In addition if a obtained by performing the regression - this will indicate that the model has a high forecasting power (for the variance).

[38] This section relies mostly on Giannopoulos (1995, section 4.2).

[39] The QGARCH from Sentana (1991) is similar to the AGARCH model except for the sign of γ.

$$E_t[\varepsilon_t^2] = h = \frac{\omega + \gamma^2\alpha}{1 - \alpha + \beta} \tag{42}$$

which indicates that the stationarity condition for the AGARCH model is identical with the condition for the GARCH model.

Lastly let me list a few other asymmetric GARCH specifications:

- The Threshold GARCH from Glosten, Jagannathan and Runkle (1991) and Zakoian (1991), ie:

$$h_t = \omega + \alpha\varepsilon_{t-1}^2 + \beta h_{t-1} + \gamma I_t\varepsilon_{t-1}^2 \tag{43}$$

where I_t is an indicator function defined as: $I_t = \begin{pmatrix} \varepsilon_t < 0 \;\Rightarrow\; 1 \\ \varepsilon_t \geq 0 \;\Rightarrow\; 0 \end{pmatrix}$. The reason why this mode

referred to as a threshold model is because when γ is positive then negative values of ε_{t-1} will have an additive impact on the conditional variance. As in the AGARCH model then negative errors have a greater impact on the variance if $\gamma < 0$.

- The non-linear AGARCH from Engle and Ng (1991), ie:

$$h_t = \omega + \alpha[\varepsilon_{t-1} + \gamma\sqrt{h_{t-1}}]^2 + \beta h_{t-1} \tag{44}$$

- The VGARCH model from Engle and Ng (1991), ie:

$$h_t = \omega + \alpha\left[\frac{\varepsilon_{t-1}}{\sqrt{h_{t-1}}} + \gamma\right]^2 + \beta h_{t-1} \tag{45}$$

B.2   Modelling the Volatility structure

We will now return to the data in table 1.

For all the date series we have estimated the GARCH(1,1) model and the IGARCH(1) model.

The volatility structure as at 2 January 2012 for each of the estimation approaches compared to the 60-and 90 day rolling window technique, is shown below in figure 2:
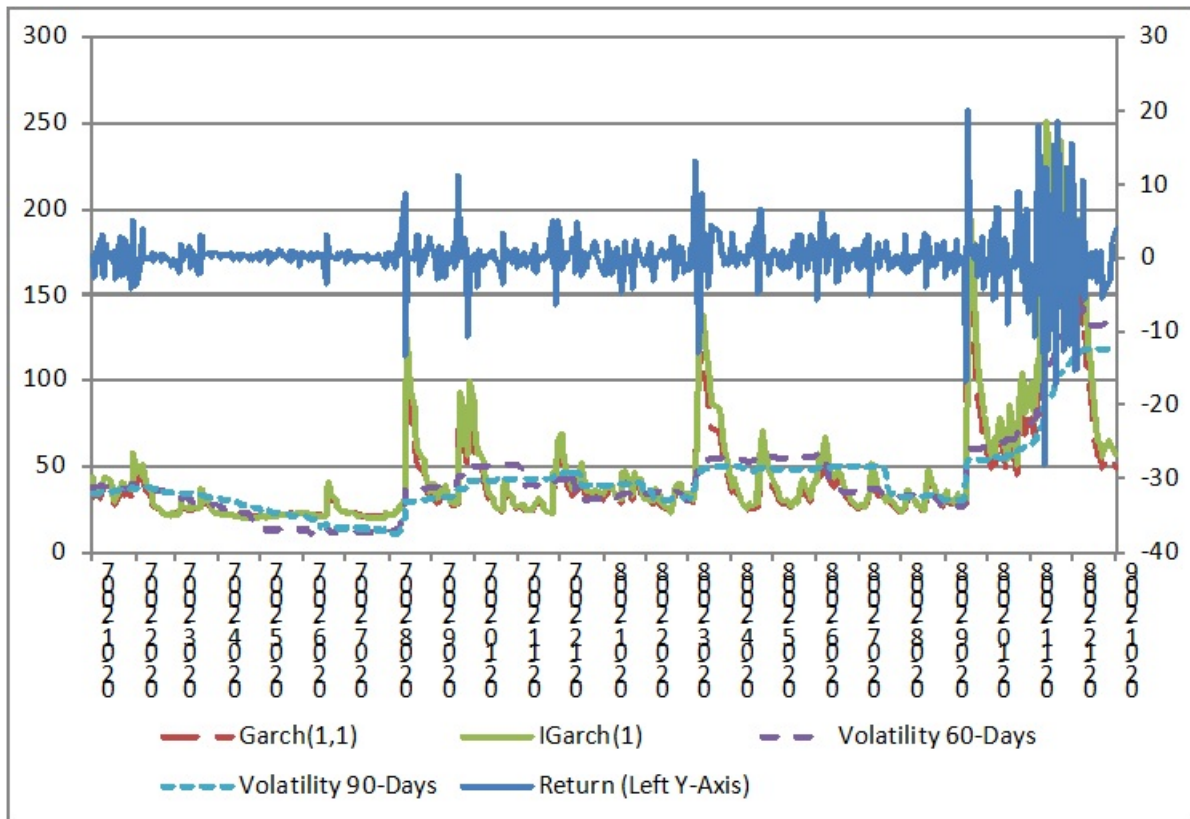
Figure 2   **Comparison of Volatility Structures**
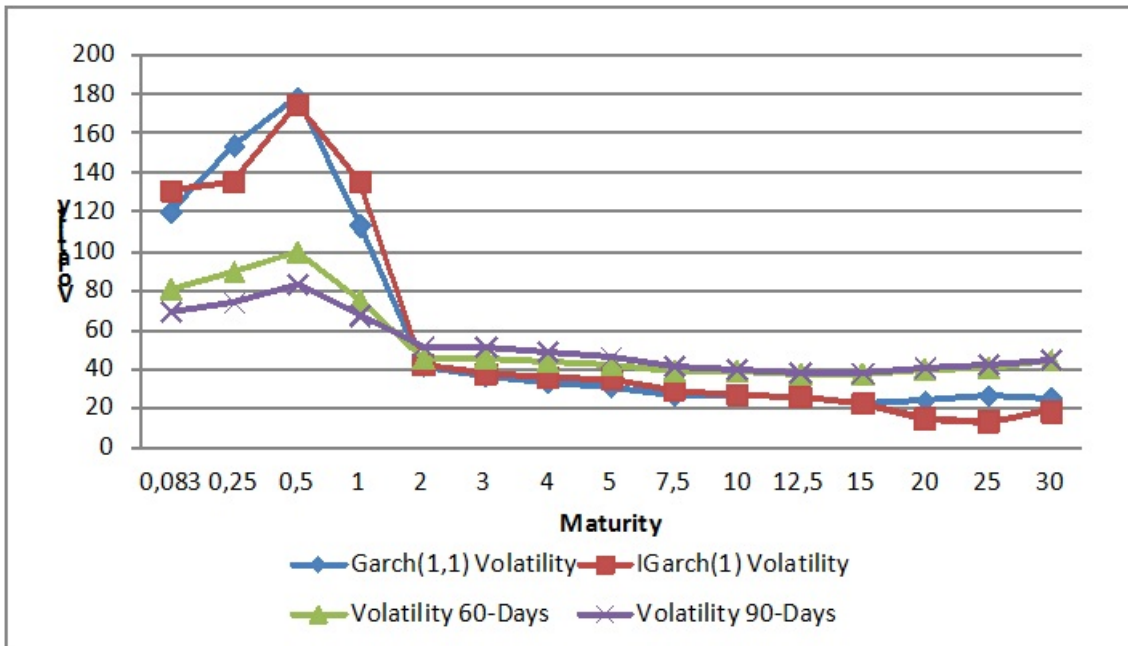
As seen from the figure the differences in volatility are more pronounced for short maturities.

To illustrate this we have in figure 3 shown the volatility pattern for the 1-year rate over the period 2. January 2007 - 2 January 2009:

Figure 3   Evolution in the Volatility for the 1-year Interest Rate

In connection with the estimation results we can report the following[40]:

- The persistence in volatility is comparable to what is usually reported in the literature. Furthermore, our estimate of β is largest for the 1-month rate (0.92) and in general is declining with maturity - however it never falls below 0.81
- All the time-series are stationary - $1 \gg \alpha + \beta$
- In most of the cases R^2 is higher in the GARCH model than in the IGARCH model - allthough little difference is observed. In general we have that R^2 is between 11-27% - with the highest R^2 for shorter maturities
- For both the GARCH model and IGARCH model we observe that we can accept the hypothesis of no autocorrelation in the squared normalized residuals
- In general we have that the normalized residuals exhibit lesser excess kurtosis and less skewness - that is we are close to the normal distribution (standardized normal distribution)

At the end of this Appendix we have shown a detailed description of our estimation results for both the GARCH model and the IGARCH model for the 1-year interest rate.

---

[40] For reasons of space we have omitted the tables here - except for the 1-year rate - but they can be obtained from the author.

From our estimation results we have decided to use the GARCH model instead of the IGARCH model, because of the following two observations:

- In most of the cases R^2 is higher in the GARCH model than in the IGARCH model
- In all cases Akaike's Information Criterion (AIC) was in favor of the GARCH model

Construction of a term structure of GARCH forecast for any time horizon can now be derived from the estimated model. This is in general straightforward[41], and is performed iteratively using the appropriate variance specifications - though taking into account that for the j-step ahead forecast for j > 1 the squared return is to be set equal to the variance for the j-1 step - this is what we do in our calculations.

Remark:

As mentioned by Alexander (1996) insufficient GARCH effect in data may lead to convergence problems in the optimization procedure. This is true if the GARCH model is being estimated using a variant of the BHHH-algoritm[42] - which seems to be the preferable optimization procedure suggested in the literature, see for example Greene (1993) and Bollerslev (1986). We howerver did not encounter any convergence problems, probably for the following reasons:

- There is significant evidence of GARCH effect in data
- We used the BFGS[43] as our main optimization procedure, which is much more robust than the BHHH

Lastly it is worth mentioning that the BFGS method was initialized with starting values obtained from the downhill-simplex procedure.

---

[41] Except for the EGARCH-model, see Alexander (1996).

[42] See Berndt, Hall, Hall and Hausman (1974).

[43] See Hald (1979).

Appendix C

In this appendix we will briefly explain how we have designed our yield-curve interpolation.

Why this is of importance can be formulated as:

- From the swap market it is possible to build a yield-curve using a bootstrapping procedure - which ensures that all swaps and any included forward-and spot-rates are priced perfectly. However, between observable data points, some yield-curve smoothing techniques are neccesary
- If we use Monte Carlo simulation procedure (or any other sampling technique) to generate interest rate paths from a subset of maturity points - then a smoothing technique is needed to generate a continuous yield-curve at every time-step

These are probably the two most important reasons why a flexible, robust and appropriate smoothing procedure has to be designed[44].

We have here selected the maximum smoothness approach of Adams and Deventer (1994). Other methods exists, such as for example spline-methods - but the problems outlined in Shea (1985) still remain. Shea has the following comments to forward-rates obtained by polynomial splines - namely that they are unstable, fluctuate widely and often drift off to very large positive or even negative values.

The yield-curve can either be formulated in terms of prices, spot-rates or forward-rates, ie:

$$
P(0,T) = e^{-R(0,T)T}
$$

$$
P(0,T) = e^{-\int_0^T f(0,s)ds}
$$

$$
\text{where}
$$

$$
f(0,T) = R(0,T) + T\frac{\partial R(0,T)}{\partial T}
$$

(46)

where P(0,T), R(0,T) and f(0,T) are respectively the bond-price, the spot-rate and the forward-rate.

The idea of Adams and Deventer is to determine the maximum smoothness term structure within all possible functional forms.

If the maximum smoothness criteria is defined as the forward rate curve on an interval (0,T) that minimizes the functional:

---

[44] Of course smoothing techniques can also be used directly to obtain the yield-curve from prices of coupon bonds - this issue will however not be addressed here - see instead Tanggaard (1997).

$$\min \int_{0}^{T} \left[\frac{\partial^2 f(0,s)}{\partial s^2} \right]^2 ds \qquad (47)$$

then Adams and Deventer show that the forward-rate model that satisfies the maximum smoothness criteria from formula 47 while fitting the observed prices is a fourth-order spline of the following special kind:

$$\begin{gathered} f(0,t) = a_i + b_i t + c_i t^4 \\ \text{for } t_{i-1} < t \le t_i \\ \text{and} \\ i = 1,2,3,....,m + 1 \end{gathered} \qquad (48)$$

where m is the number of observed bond-prices.

The smoothness criteria and the requirement that we want to price the m bonds without measurement error gives rise to the following $3m + 3$ system of equations for the coefficients a_i, b_i and c_i, for all i[45]:

$$\begin{gathered} a_i + b_i t_i + c_i t_i^4 = a_{i+1} + b_{i+1} t_i \quad \text{for } i = 1,2,3,...,m \\ b_i + 4c_i t_i^2 = b_{i+1} + 4c_{i+1} t_i^3 \quad \text{for } i= 1,2,3,...,m \\ a_i(t_i - t_{i-1}) + \frac{1}{2} b_i(t_i^2 - t_{i-1}^2) + \frac{1}{5} c_i(t_i^5 - t_{i-1}^5) = -\ln\left(\frac{P_i}{P_{i-1}} \right) \quad \text{for } i = 1,2,3,...,m \\ c_{m+1} = 0 \\ b_{m+1} = 0 \\ a_1 = r_1 \end{gathered} \qquad (49)$$

It should be mentioned that the last two criteria have been selected for practical reasons - they are not associated with the smoothness criteria or the zero (0) measurement requirement for the observable bond-prices.

The last criteria ensures that the shortest observable spot-rate is equal to the shortest forward-rate - a logical feature. The second last criteria makes restrictions on the asymptotic behaviour of the yield-curve - more precisely it ensures that the slope of the forward-rate curve is zero (0) at the endpoint of the last interval.

As a final remark we might note that the algebraic linear system leads to a banded symmetric diagonally dominant and positive definite linear system which can be easily solved using

---

[45] See Adams and Deventer (1994).

special algorithms, see Golub and Van Loan (1993, chapter 5).

Let us finally illustrate the method with a simple example:

For that purpose we assume we know the prices for the following maturity dates:

Period	Interest Rate	Bond Price
3-Month	3,75	99,067
1-Year	4,25	95,839
2-Year	4,35	91,668
5-Year	4,56	79,612
7.5-Year	4,20	72,979
10-Year	4,75	62,189
15-Year	4,31	52,388
30-Year	5,65	18,360

In figure 1 below we have shown the spot-rate curve and the forward-rate curve using the Adams and Deventer procedure.
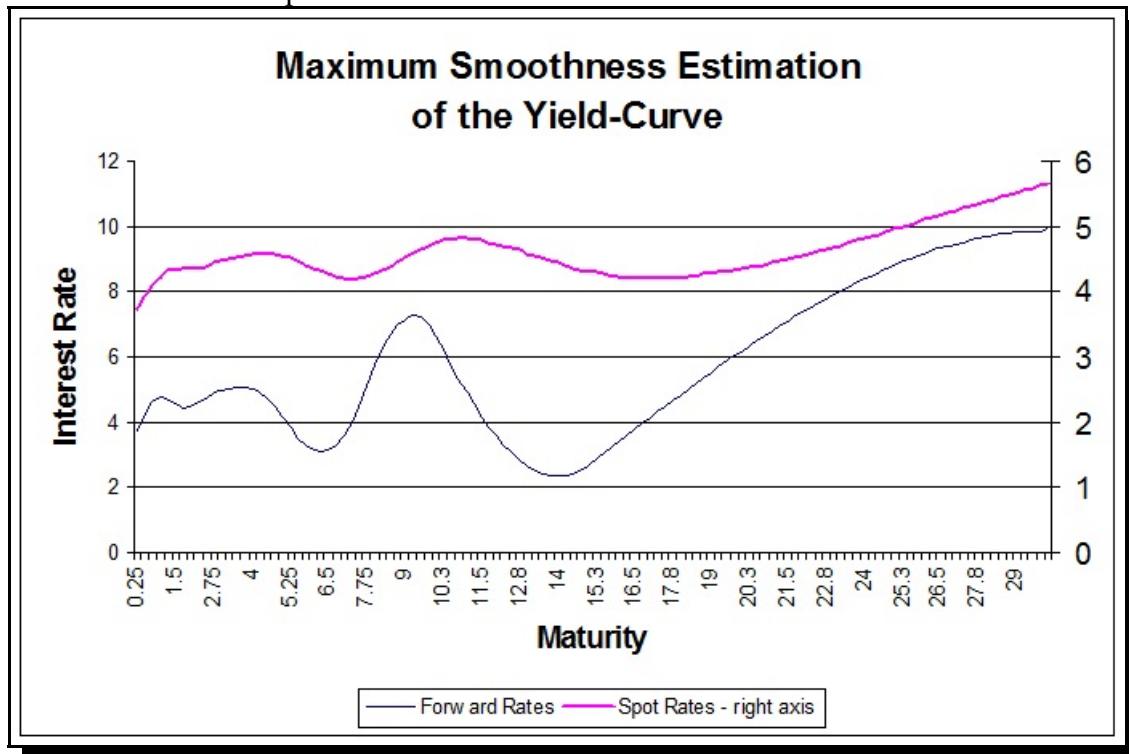

Figure 1

The data might be a bit extreme but it serves to illustrate the maximum smoothness procedure.