Yield curve event tree construction for multi stage stochastic programming models

Kourosh Marjani Rasmussen* and Rolf Poulsen†

November 16, 2007

Abstract

Dynamic stochastic programming (DSP) provides an intuitive framework for modelling of financial portfolio choice problems where market frictions are present and dynamic re–balancing has a significant effect on initial decisions. The application of these models in practice, however, is limited by the quality and size of the event trees representing the underlying uncertainty. Most often the DSP literature assumes existence of "appropriate" event trees without defining and examining qualities that must be met (ex–ante) in such an event tree in order for the results of the DSP model to be reliable. Indeed defining a universal and tractable framework for fully "appropriate" event trees is in our opinion an impossible task. A problem specific approach to designing such event trees is the way ahead. In this paper we propose a number of desirable properties which should be present in an event tree of yield curves. Such trees may then be used to represent the underlying uncertainty in DSP models of fixed income risk and portfolio management.

1 Introduction

One of the main sources of uncertainty in analyzing risk and return properties of a portfolio of fixed income securities is the stochastic behavior in

*Informatics and Mathematical Modelling, Technical University of Denmark, Bldg. 305, DK–2800 Lyngby, Denmark. kmr@imm.dtu.dk

†Department of Applied Mathematics and Statistics, Institute for Mathematical Sciences, Universitetsparken 5, University of Copenhagen DK–2100, Denmark. rolf@math.ku.dk

the evolution of the shape of the term structure of the interest rates (yield curve). This uncertainty is sometimes referred to as shape risk, see for example Zenios (2007). Shape risk refers to the risk that interest rates with different maturities change in different ways as the time goes by. Figure 1 shows how the Danish yield curves have changed in the period 1995 to 2006.

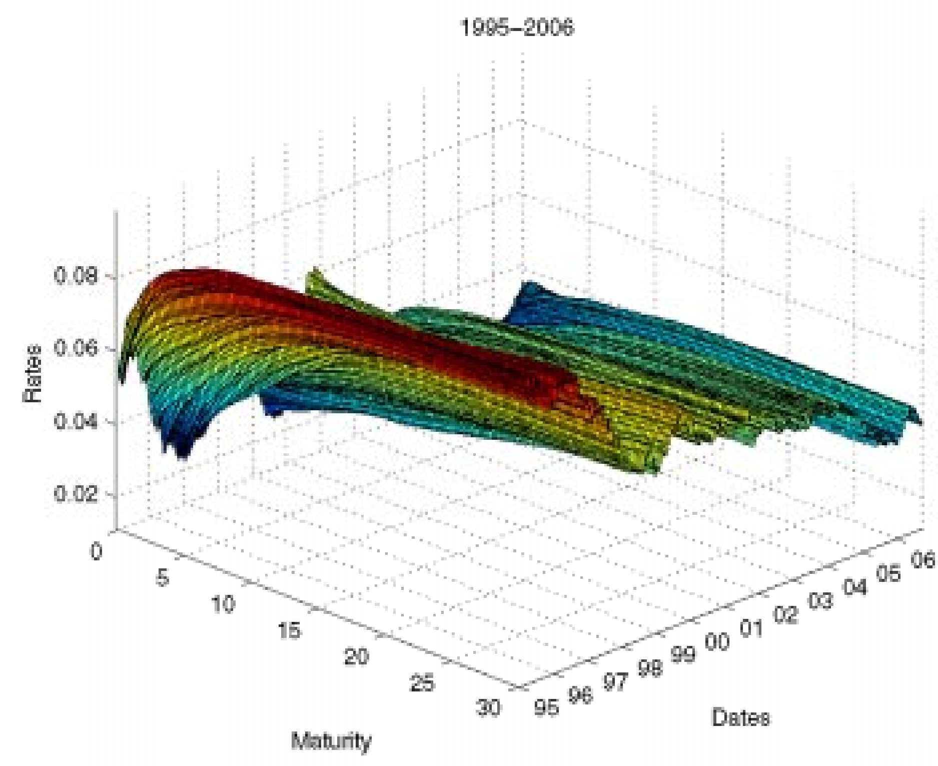

Figure 1: Historical data on Danish yield curves for the period 1995 to 2006.

We can see that the short rates have been more volatile than the long rates. We also observe that a simple parallel shift assumption does not hold; yield curves evolve in more complicated manners. Capturing the dynamics of yield curves in a multi period scenario tree is the purpose of this paper.

Dynamic stochastic programming (DSP) provides a flexible framework for portfolio and risk management problems. Trade frictions such as fixed costs, tax affects and limits on borrowing and short sale of assets can be incorporated in such models. Portfolio readjustments may as well be captured. This is in particular important for fixed income securities due to the usually long term perspectives of such investments. Finally no assumptions on the underlying uncertainty are required. This means that for example heavy tails which play an important role in extreme event considerations can be accounted for. But it also means that special care needs to be taken when it comes to modelling the underlying uncertainty. The event trees should be consistent with historical data as well as internally consistent with regards to the mechanisms governing the dynamics of the uncertain variables (see Ziemba 2001). Such consistency criteria include for example the no arbitrage conditions (see Klaassen 2002).

We suggest the following guidelines for generating an event tree of yield curves:

1. The distance between the underlying continuous interest rate process and the discretized event tree should be minimized.
2. The event tree should match the underlying continuous process both globally, i.e. for any given future period as well as locally, i.e. for any subtree of the event tree.
3. The actual levels of the generated scenarios should be realistic, for example the tree should not include any negative interest rates, or many extreme scenarios.
4. The volatilities of the interest rates of different maturity should be consistent with the implied volatilities of a market benchmark.
5. There should be no arbitrage opportunities in any of the subtrees of the event tree.
6. Types of changes in the shape of the yield curve in future nodes of the event tree should reflect those observed historically from an economical regime which is assumed similar to the one the event tree is built for.
7. The model should be mean reversive.
8. No volatility clumping; Volatility clumping refers to the case where a period of high volatility is followed by another period of high volatility. Volatility clumping is observable in the equity market, but empirical studies have shown that there is no volatility clumping for the interest rates.

There is a vast amount of literature on interest rate modelling. These models can in general be categorized as being discrete or continuous, normal or a log–normal, 1–factor or multi–factor and finally either more theoretically or more empirically inclined. What all such models have in common is the fact that they have been originally developed either for estimating current prices of interest rate sensitive assets, or for prediction purposes. None of the standard models therefore are designed in order to construct yield curve event trees fulfilling criteria 1 to 8 at the same time.

In this paper, we propose an overall framework for building a yield curve event tree and testing whether or not the consistency criteria are respected. The rest of this paper is organized as follows:

In section 2 we perform factor analysis (also known as principal component analysis) in order to identify the most significant factors in capturing yield curve variability. Then in section 3 we describe a simple 3–factor vector auto regressive model with lag 1 (VAR1) representing the underlying stochastic process. A non–linear discretization model of the stochastic process is then suggested in section 4. In section 5 we outline an approximative approach for solving the discretization model. In section 6 we argue why a simple 1–factor interest rate model such as the Vasicek model is not appropriate for stochastic programming applications and why the proposed 3–factor model provides more reliable solutions. Finally we conclude the paper in section 7.

2 Factor analysis of yield curves

Factor analysis is a statistical technique to detect the most important sources of variability among observed random variables. It may be used on historic time series of a multidimensional random variable to decide factors ordered after how much variability they explain. In linear algebraic terms it is an orthogonal linear transformation that transforms data to a new coordinate system in such a way that the greatest source of variance lies on the first factor, the second largest on the second factor and so on. It is used for reducing the dimensionality of a data set while keeping its characteristics. This is done by keeping only the main factors while ignoring the ones that only explain an insignificant proportion of the variance.

Litterman & Scheinkman (1991) and Knez, Litterman & Scheinkman (1994) use factor analysis to show that three factors explain – at a minimum – 96% of the variability on several American zero coupon yield curves in the period 1985 to 1988. Dahl (1994) shows similar results for the Danish data in the 1980's and Bertocchi, Giacometti & Zenios (2005) repeat the experiments for American and Italian data during 1990's again with similar results.

These findings are used by some practitioners to improve duration hedging (immunization) by factor based duration hedging (factor immunization). The main shortcoming of these hedging techniques is that they are myopic and do not consider the re–balancing effects in long term fixed income portfolio investments. Rather than using factor analysis for shape risk hedging, we use factor analysis as a means of finding a sufficient number of factors to be used as the underlying factors of uncertainty for the proposed interest rate model of this paper. We perform factor analysis on the Danish yield curves for the period 1995–2006. Like in earlier works we find that 3 factors are enough to capture almost all variability (99.99%) for the Danish yield curves. Figure 2 shows the factor loadings as a function of maturities in years.

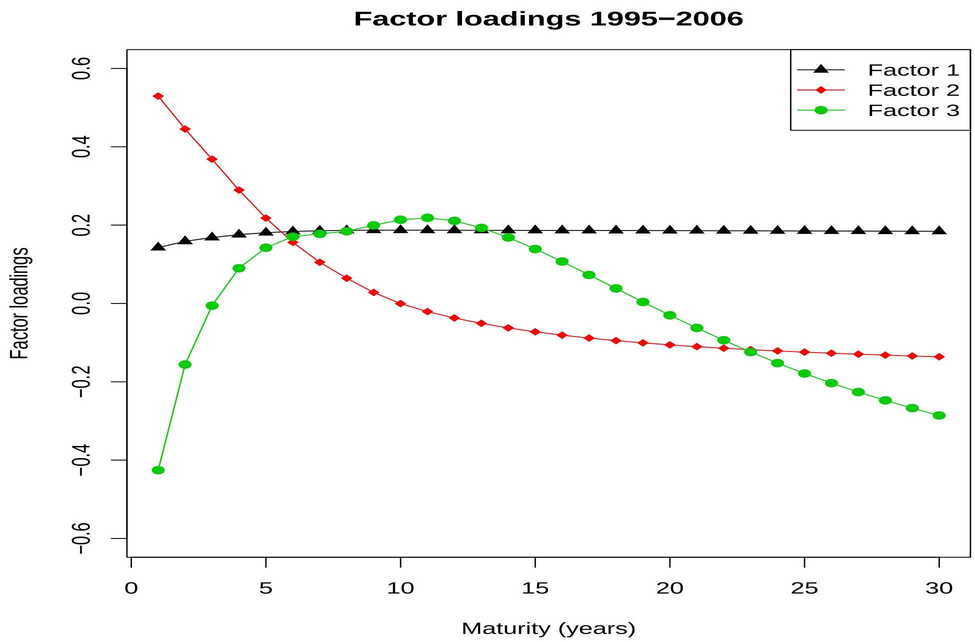

Figure 2: Factor loadings of the Danish yield curves for the period 1995 to 2006.

The first factor explains almost 95% of all variability. It can be interpreted as a slight change of slope for interest rates with maturities up to approximately 6 years together with a parallel shift for the rest of the curve. The second factor, explaining 4.7% of the variability, corresponds to a change of slope for the whole curve. However the slope change for the first 10 years is much more pronounced. Finally the third factor corresponds to a change of curvature in the yield curves. This factor explains only about 0.3% of the total variability.

From a statistical viewpoint we could suffice with level and slope as the main sources of variability. Nevertheless we do not reject the third factor, curvature, due to its economical appeal; changes of curvature are observed now and then, and a model not being able to represent those changes properly has a potential weakness of not capturing important movements in the interest rate market.

Inspired by the results found in this section we define the three factors which we want to use in our interest rate model as follows:

1. Level: An arbitrary rate such as the one year rate,Y_1, may be used as a proxy for level.
2. Slope: A good proxy for the slope would be $Y_{20} - Y_1$ where Y_{20} stands for the 20 year rate. This expression is an approximation of the average slope of the yield curve. The 20 year bond is chosen as the long rate here, since we observe in our historical data, that almost all yield curves flatten at about this maturity.
3. Curvature: The expression $Y_6 - (\omega Y_1 + (1 - \omega)Y_{20})$, with Y_6 as the 6 year rate, may be used as a proxy for the curvature. ω is the weight corresponding to the proportion of the distance in between the middle of long rates. It is chosen so that the curvature would be zero if the curve is a straight line, negative if the curve is convex and positive if the curve is concave.

In the rest of this paper we use level, slope and curvature defined as above as the factors of the interest rate model in question.

3 A vector autoregressive model of interest rates

A vector autoregressive model with lag 1 (VAR1) may be defined as:

$$x_{t+1} = \mu + A(x_t - \mu) + \epsilon_{t+1}$$

where x_t is an $n \times n$ matrix, μ is an $n \times 1$ vector and $\epsilon_{t+1} \sim \mathcal{N}_n(\overline{0}, \Omega)$ and Ω is an $n \times n$ matrix. In this formulation of the VAR1 model, μ is interpreted as the long term drift. A and μ are deterministic parameters which need to be calibrated based on historical data.

The conditional mean and covariance for the error term ϵ_{t+1} are given as:

$$E[\epsilon_{t+1}|x_t] = 0$$
$$E[\epsilon_{t+1}\epsilon_{t'+1}|x_t] = \Omega$$

Given the state of an uncertain variable at time x_t, the purpose of the model is to predict the state of the variable at time $t+1$, i.e. x_{t+1}. Based on the findings of the previous section we define the vector x_t as the proxies for level, slope and curvature $(l_t, s_t, c_t)^{\mathrm{T}}$ of the yield curves.

An example of the VAR1 model with 3 factors looks like:

$$
\begin{aligned}
l_{t+1} &= \mu_l + a_{ll}(l_t - \mu_l) + a_{ls}(s_t - \mu_s) + a_{lc}(c_t - \mu_c) + \epsilon_{l,t+1} \\
s_{t+1} &= \mu_s + a_{sl}(l_t - \mu_l) + a_{ss}(s_t - \mu_s) + a_{sc}(c_t - \mu_c) + \epsilon_{s,t+1} \\
c_{t+1} &= \mu_c + a_{cl}(l_t - \mu_l) + a_{cs}(s_t - \mu_s) + a_{cc}(c_t - \mu_c) + \epsilon_{c,t+1}
\end{aligned}
$$

To estimate the parameters of the VAR1 model (μ, A, Ω) we can use the parameter estimation for a general linear regression model of the form:

$$
y_i = \alpha + \beta x_i + \epsilon_i, \qquad \text{for all } i = 1, \cdots, n
$$

Or in matrix form:

$$
\begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} + \begin{bmatrix} \epsilon_1 \\ \vdots \\ \epsilon_n \end{bmatrix}
$$

This can be rewritten as:

$$
Y = \overline{X}\delta + \varepsilon
$$

The VAR1 model can be rewritten in this form. Now we may use standard least square estimators as follows:

$$
\hat{\delta} = (\overline{X}^{\mathrm{T}}\overline{X})^{-1}\overline{X}^{\mathrm{T}}Y
$$

which minimizes the sum of least squares in the expression $||Y - \overline{X}\delta||^2$.

The estimator for the residuals (ε) is given as:

$$
\begin{aligned}
res &= Y - \overline{X}\hat{\delta} \\
\hat{\Omega} &= res^{\mathrm{T}} res/(n-1)
\end{aligned}
$$

The estimator $\hat{\delta}$ is then decomposed into μ and A from the VAR1 model and the estimator $\hat{\Omega}$ can be directly used as the estimator for Ω in the VAR1 model.

The VAR1 model so far may only be used for one–period predictions (same interval length as in the historical time series). But it may easily be extended to predict k periods ahead:

$$x_{t+k} = \mu + A^k(x_t - \mu) + \epsilon_{t+k}$$

where $\epsilon_{t+k} \sim \mathcal{N}_n(\overline{0}, \sum_{i=1}^{k} A^{i-1}\Omega(A^{i-1})^{\mathrm{T}})$

The reasons for choosing a VAR1 model as the underlying model of interest rate uncertainty are the following:

1. One can choose any factors or any number of factors to describe the variability. This gives us maximum flexibility with respect to our observations from a factor analysis of interest rates.
2. Time step flexibility. Varying time steps can be easily implemented.
3. Mean reversion is built into the VAR1 model.

The VAR1 model is discrete in time but continuous in states, so in order to use the model as a scenario generator for stochastic programs we need to discretize it in states as well. This can be done using a moment matching model (See Høyland & Wallace (2001)). We propose a yield curve scenario discretization model in the next section.

4 Scenario generation and event tree construction

In DSP literature for fixed income securities often simple models of interest rates are used to represent the underlying interest rate uncertainty. In several applications lattice structures are either blown up into unique paths or sampled from so that to account for the path dependency of DSP problems. One immediate problem with such approaches is that the uncertainty space is not covered as efficiently as possible. This is due to the recombining structure of the original trees together with the fact that only a very coarse time step discretization is possible due to the curse of dimensionality when the recombining trees are blown up.

Others (Nielsen and Poulsen 2004, etc.) have used continuous interest rate models. Such models are either continuous both in time and state, or discrete in time and continuous in states. Discretizing in time is normally straight forward; it is a question of reformulating a differential equation into a difference equation. Discretizing in state, however, is often a more challenging issue. A number of nodes (in our case including yield curve information) have to be generated for each time point to give a discrete representation of the continuous distribution. There is no general consensus as to the best way of doing this discretization. In one stream of research the main focus is on generating discrete distributions which mimic the underlying continuous distribution as closely as possible. This is either done by sampling (see Shtilman and Zenios 1993) , or moment matching approaches (Høyland and Wallace 2001). In the other stream of research the aim is not necessarily to get the closest discrete representation of the continuous distribution, but rather finding a discrete representation which results in a closer approximation to the "true" optimal solution of the stochastic program in question. Here the "true" optimal solution refers to the solution we would get, if we were able to solve the stochastic program using the underlying continuous process directly. Indeed if we were able to do that, there would be no need to discretize the process in the first place, but it can be shown (See Pflug 2001) that in general if the discrete process has the smallest distance (using the transport metric) to the underlying continuous process, then the SP solutions found will be guaranteed to be within certain bounds of the "true" SP solutions. (See also Pflug 2001, Pflug and Hochreiter 2002, Pennanen 2004, Romisch and Heitsch 2003) . Although theoretically appealing, the guaranteed bounds are in many cases too large in order to have any practical interest, (See Wallace and Kaut 2003) . Comparison and further development of specialized models and solution algorithms for these two streams of scenario discretization approaches is the subject of future research.

An extensive comparative study of different yield curve scenario generation approaches is outside the scope of this paper. Instead we propose a yield curve scenario generation model which abides by the criteria 1 to 8 mentioned earlier in this paper. Note that the following model is single period. It can be extended to a multi–period model with some minor changes.

We define the following sets, parameters and variables:

Sets:
f: Set of factors (level, slope and curvature), f' is alias for f.
i: Set of zero coupon bonds (zcb's).
i': A subset of the set i corresponding to the zcb–rates which define the three factors. We have chosen i' to be the set of 1, 6 and 20 year zero coupon bonds.

j: Set of parameters of the Nelson Siegel function; 0 to 3.
t: Set of time points.
s: Set of scenarios.

Parameters:
$Mean_f$: The mean value for factor f. This value comes from the VAR1 model.
$Covar_{f,f'}$: The covariance matrix of the error term taken from the VAR1 model.
$Skewness_f$: Skewness of factor f. Assumed to be zero based on the normality assumption of the VAR1 model.
τ_i^t: Time to maturity for zcb i at time t.
PP_i^{parent}: Prices of the zero coupon bonds at the root, The prices are calculated using initial rates: $PP_i^{\text{parent}} = e^{-r_i \tau_i^{\text{parent}}}$.
ψ^{Const}: The martingale probability; assumed equal for all scenarios. It is found from the equation $PP_{i''}^{\text{parent}} = \sum_s \psi^{\text{Const}}$ where bond i'' matures exactly at the children nodes of the tree with a price of 1.

Variables:
$x_{f,s}$: A future estimate of factor f in scenario s given by the VAR1 model.
$E(x)_f$: The expected value of factor f over all scenarios.
$\sigma(x)_{f,f'}$: The covariance matrix of factors across all scenarios.
$E3(x)_f$: The skewness of factors across all scenarios.
$Y_{i',s}^{(VAR1)}$: The 3 yields comprising the 3 factors at scenario s.
$NSY_{i',s}$: The 3 yields comprising the 3 factors at scenario s as given by the Nelson Siegel function.
$\varphi_{s,j}$: Parameter j of the Nelson Siegel function at scenario s.
$R_{i,s}$: The entire yield curve given by the Nelson Siegel function at scenario s.
$CP_{i,s}$: Price of bond i at scenario s.

The overall objective of the optimization model is to match the moments of the underlying stochastic process (the VAR1 model) as closely as possible. At the same time the parameters of the Nelson Siegel function should be found so that the yields resulting from Nelson Siegel are as close as possible to those found by the VAR1 model. We need Nelson Siegel (or some other yield curve smoothing function) in order to get the rest of the yield curve, since the VAR1 model is based on 3 yields only.

The objective function is to minimize sums of least squares corresponding to the overall objective of the model:

$$
\begin{aligned}
\text{Minimize } & \sum_f (E(x)_f - Mean_f)^2 + \sum_f \sum_{f'} (\sigma(x)_{f,f'} - Covar_{f,f'})^2 + \\
& \sum_f (E3(x)_f - Skewness_f)^2 + \sum_s \sum_{i'} (Y_{i',s}^{(VAR1)} - NSY_{i',s})^2 \quad (1)
\end{aligned}
$$

The moments of the discrete scenarios as found by the optimization model are defined in Equations 2 to 4:

$$
\begin{aligned}
E(x)_f &= \sum_s p_s x_{f,s} && \text{for all } f && (2)\\
\sigma(x)_{f,f'} &= \sum_s p_s (x_{f,s} - E(x)_f)(x_{f',s'} - E(x)_{f'}) && \text{for all } f, f' && (3)\\
E3(x)_f &= \frac{\sum_s (x_{f,s} - E(x)_f)^3}{(\sum_s (x_{f,s} - E(x)_f)^2)^{3/2}} && \text{for all } f && (4)
\end{aligned}
$$

In Equation 5 the 3 yields corresponding to the 3 underlying maturities used in the VAR1 model are found by the Nelson Siegel model. Note that the final term of the objective function requires that $NSY_{i',s}$ should be as close as possible to the 3 yields found by the VAR1 model. So Equation 5 in interaction with the objective function calibrates the parameters of the Nelson Siegel function. These parameteres are used in Equation 6 to decide the entire yield curve at each scenario.

$$
\begin{aligned}
NSY_{i',s} &= \varphi_{s,0} + \varphi_{s,1} e^{-\varphi_{s,3}\tau_{i'}^{\text{parent}}} + \varphi_{s,2}\tau_{i'}^{\text{parent}} e^{-\varphi_{s,3}\tau_{i'}} && \text{for all } i', s && (5)\\
R_{i,s} &= \varphi_{s,0} + \varphi_{s,1} e^{-\varphi_{s,3}\tau_i} + \varphi_{s,2}\tau_i^{\text{parent}} e^{-\varphi_{s,3}\tau_i^{\text{parent}}} && \text{for all } i, s && (6)
\end{aligned}
$$

The VAR1 model is defined in terms of factors and not yields. Equations 7 to 9 find the yields corresponding to the factors estimated by the VAR1 model at each scenario.

$$
\begin{aligned}
Y_{1,s}^{(VAR1)} &= x_{1,s} && \text{for all } s && (7)\\
Y_{20,s}^{(VAR1)} &= x_{2,s} + Y_{1,s}^{(VAR1)} && \text{for all } s && (8)\\
Y_{6,s}^{(VAR1)} &= \frac{5}{19} Y_{20,s}^{(VAR1)} + \frac{14}{19} Y_{1,s}^{(VAR1)} + x_{3,s} && \text{for all } s && (9)
\end{aligned}
$$

The main reason to define the yield curve discretization process as an optimization model is that it enables us to add constraints which give the user a degree of control over the outcome. One such constraint may be forcing a lower bound on interest rates, for instance not allowing negative rates:

$$R_{i,s} \geq 0 \qquad \text{for all } i, s \tag{10}$$

Another condition may be not to allow arbitrage in the interest rates. In Equations 11 and 12 we introduce a more restrict condition than the no arbitrage condition, namely we require that martingale probabilities should be equal across all scenarios:

$$CP_{i,s}^{\text{child}} = e^{-R_{i,s}\tau_i^{\text{child}}} \qquad \text{for all } i, s \tag{11}$$

$$PP_i^{\text{parent}} = \sum_s \psi^{\text{Const}} CP_{i,s}^{\text{child}} \qquad \text{for all } i \tag{12}$$

The model 1 through 12 gives the user a great degree of flexibility over the outcome of the discretization process. Subjective expert opinion is integrated with objective econometrical and financial theory. The model, however, is non–linear, non–convex and as such has several local minima. Solving such a problem fall into the realm of global optimization. The general purpose global solvers are as of yet underdeveloped. Specialization of existing algorithms is therefore needed for solving this problem to optimality. This is outside the scope of the current paper. Instead we propose an approximative approach to find reasonable solutions in the next section.

5 An approximative solution approach

The approximation is in dividing the model into three parts and solving them in a serial manner instead of solving the entire problem in one go:

1. First we solve a model comprising of the objective function less the 4th term with constraints 2 to 4. This model results in discretized factors matching the first 3 moments of the underlying VAR1 model one period ahead. We also add constraints 7 through 10 to guarantee no negative rates.

2. Then we solve a second model where the objective function is made of the 4th term and the only constraint is Equation 5. Finding the parameters of the Nelson Siegel model we now simply use Equation 6 to find the entire yield curves for each scenario.

3. Finally we apply Equations 11 and 12 to remove arbitrage.

The two sub models are non–linear non–convex themselves but it is possible to find optimal solutions to these problems using standard non–linear solvers which is what we have done using GAMS/ConOpt[1].

Wasn't it due to the no–arbitrage conditions then solving the two models separately would corresponded to solving the entire problem. We therefore compare the yield scenarios before removing arbitrage with those after arbitrage removal, See Figures 3 to 6. The scenarios in the left are before the arbitrage removal part of the approximative algorithm has been applied. The scenarios in the right are after arbitrage removal. The smaller the change is between the left hand side and the right hand side scenarios the closer the results of the approximative approach will be to that of solving the entire problem.

The first 2 figures are from August 2005 when the initial term structure is rather steep (the stippled curve). In these cases we note that there is very little difference between the rates before and after arbitrage removal, meaning that the approximative approach generates near optimal solutions for the entire model. In the last 2 figures the starting point is May 2007 when the initial yield curve is essentially flat. In this case we note a considerable difference between the rates before and after removal of arbitrage. In both cases, however, the solutions found may be used as initial solutions for solving the entire problem.

We leave solving the entire problem as future work. Instead we replace the Nelson Siegel function with an affine function developed for our 3–factor VAR1 model of interest rates (See Poulsen 2007) . It is known from interest rate theory that Nelson Siegel does not produce arbitrage free curves in any continuous model. Given that, there is little hope that the discretized models will be arbitrage free regardless of the number of scenarios generated. The affine function is, however, constructed arbitrage free in the continuous setting. So the hope is that by adding scenarios we will satisfy the no–arbitrage

[1]GAMS/CONOPT is a non linear problem (NLP) solver available for use with General Algebraic Modeling System (GAMS). See http://www.gams.com/solvers/solvers.htm

condition in the discrete scenarios as well. The graphs in the bottom of Figures 3 to 6 are the result of an affine smoothing of the yield curves. Again the yield curve scenarios before and after removing of arbitrage are considered.

In the rest of this work we use the scenario trees based on the affine model. In the next section we will compare interest rate scenarios generated by our VAR1 model with the well known 1–factor Vasicek model.

6 Vasicek versus VAR1 for event tree construction

A central theme in this paper is to convince the reader that simple 1–factor interest rate models do not capture the dynamics of historic rates as indicated by a factor analysis of historic interest rates. Even though that does not necessarily have an influence on how well such models are in pricing fixed income securities here and now, that does have an impact on estimates of prices of assets in future nodes. That is why using simple models of interest rate as the underlying source of uncertainty in a stochastic program might result in misleading solutions to the asset allocation and risk management problems that are formulated based on such interest rate scenario trees. How wrong the solutions of such stochastic programs will be is problem dependent and need to be studied for individual applications. In this section we show how we can get a graphical feel of how well an interest rate scenario tree mirrors what we expect interest rates to behave based on the criteria mentioned in the introductory part of this work.

Figures 7 to 9 show interest rate trees for 1, 6 and 20 year maturities starting on the 1th of May 2007 and running over 5 years once using the 1–factor Vasicek model as the underlying source of uncertainty and twice using our VAR1 model. The only difference between the VAR1 representations is the manner in which discretization takes place. We use our approximative discretization approach described in the last section iteratively to the future nodes of the tree to produce these multi period tree structures.

It is obvious from the figures that the trees using the Vasicek model have almost no volatility for the long rates. Looking at historic yield curves in Figure 10 this seems very unrealistic. On the other hand the VAR1 trees branched in a 4–4–4–4 fashion seem to produce overly large volatilities for all maturities. This is better seen in Figure 11 where we only consider the yield curves 5 years from May 2007. The initial yield curve is presented using a solid line. Note, however, that in the Vasicek model the initial yield curve is not the observed curve but reproduced by the model. By only looking at these graphs there is little room for suspicion left as for the insufficiency of a 1–factor Vasicek model in capturing future dynamics of interest rate, in particular the long rates.

Obviously we do not wish for our model of choice to reproduce historical yield curves exactly. That said, it is desired that the model captures characteristics seen in historic data. Our VAR1 model with a 16–4–2–2 discretization seems to produce a good approximation to the real world data from 1995–2006 as seen in Figure 12. Whether or not this is a good historical period which characteristics to mimic is a subjective question, but it is a subjective question at a high level of abstraction; we do not choose how the yield curves should exactly look like, but we make a decision as to which historic period we believe gives rise to a good approximation of future yield curve scenarios.

7 Conclusions

We have set up a number of qualitative conditions with which a yield curve scenario generation method should comply. We have shown that the 1–factor Vasicek model, even though suitable for option pricing, is unable to capture future dynamics of interest rate, which disqualifies this model as a source of uncertainty for stochastic programs. We have tailored a 3–factor VAR1 model using the 3 factors, level, slope and curvature, describing over 99% of variability in historical interest rates and we have introduced a discretzation scheme on top of that. We have presented graphs which give the user a feel of whether or not the scenarios generated are representative of what is observed in historical data as well as what is prescribed by econometrical and interest rate theory. Our VAR1 model with a 16–4–2–2 discretization gives rise to a reasonable representation of uncertainty over a 5–year period with a modest number of scenarios, 256. The three major types of yield curve shifts are present in representative quantities and the volatility of the last 10 years historic data is captured properly. There is also reversion towards the long term drifts. No negative rates or extremely low rates are observed. There are, however, some gaps in between the extreme scenarios and the main bulk of scenarios in the high end of the scale in particular for long rates. The gap can be closed if we generate more scenarios for example 32–4–4–4, but this results in 2048 scenarios which is probably about the highest number of scenarios most realistic linear stochastic programming applications can handle on a standard pc. Given that the stochastic programming problems we have in mind have 0–1 constraints we find the trees of approximately 200–300 scenarios more appealing. Whether or not this leads to serious solution deficiencies as compared to using 2000-3000 scenarios is subject of future work. We need special purpose algorithms and/or parallel routines to perform the comparison. Super computers may as well provide sufficient computing power for these tests. Our preliminary trials on LP–relaxed version of our optimization problems at hand show, however, that the first stage solution structures stabilize already at about 200–300 scenarios despite the gaps in between the high extreme scenarios and the main bulk of scenarios. Another idea that we leave to future work is trying another moment matching approach where the first four moments (kurtosis being the fourth) are matched simultaneously at each period conditioned on the root, and that only the first 2 or 3 moments are matched for the sub–trees in between the periods. Likewise applying the ideas of Pflug (2001) and Hochreiter and Pflug (2006) on optimal discretization to our problem remain as future work.

References

Bertocchi, M., Giacometti, R. & Zenios, S. A. (2005), 'Risk Factor Analysis and Portfolio Immunization in the Corporate Bond Market.', *European Journal of Operational Research* **161(2)**, 348–363.

Dahl, H. (1994), 'A Flexible Approach to Interest–Rate Risk Management.', *Financial Optimization.* pp. 189–209.

Hochreiter, R. & Pflug, G. C. (2002), Scenario tree generation as a multidimensional facility location problem. AURORA TR2002-33.

Høyland, K. & Wallace, S. W. (2001), 'Generating Scenario Trees for Multistage Decision Problems.', *Management Science* **47**, 295–307.

Kaut, M. & Wallace, S. W. (2003), Evaluation of scenario generation methods for stochastic programming. Working paper, University of Molde.

Klaassen, P. (2002), 'Comment on "Generating Scenario Trees for Multistage Decision Problems".', *Management Science.* **48**, 1512–1516.

Knez, P. J., Litterman, R. & Scheinkman, J. (1994), 'Explorations Into Factors Explaining Money Market Returns.', *Journal of Finance* **49(5)**, 1861–1882.

Litterman, R. & Scheinkman, J. (1991), 'Common Factors Affecting Bond Returns.', *The journal of fixed income.* **June**, 54–61.

Nielsen, S. S. & Poulsen, R. (2004), 'A Two–Factor, Stochastic Programming Model of Danish Mortgage–Backed Securities', *Journal of Economic Dynamics and Control* **28**, 1267–1289.

Pennanen, T. (2004), Epi–convergent discretizations of multistage stochastic programs via integration quadratures. Working paper, Helsinki school of economics.

Pflug, G. C. (2001), 'Scenario tree generation for multiperiod financial optimization by optimal discretization.', *Math. Program., Ser. B* **89**, 251–271.

Poulsen, R. (2007), A Gaussian Factor Model. Working paper.

Roemisch, W. & Heitsch., H. (2003), 'Scenario reduction algorithms in stochastic programming.', *Computational Optimization and Applications.* **24**, 187–206.

Shtilman, M. S. & Zenios, S. A. (1993), 'Constructing optimal samples from a binomial lattice.', *J. Inf. Optim.* **14**, 1–23.

Zenios, S. A. (2007), *Practical Financial Optimization: Decision making for financial engineers.*, Blackwell, Cambridge, UK.

Ziemba, W. T. (2001), *Handbook of Asset Liability Management using Stochastic Programming.*

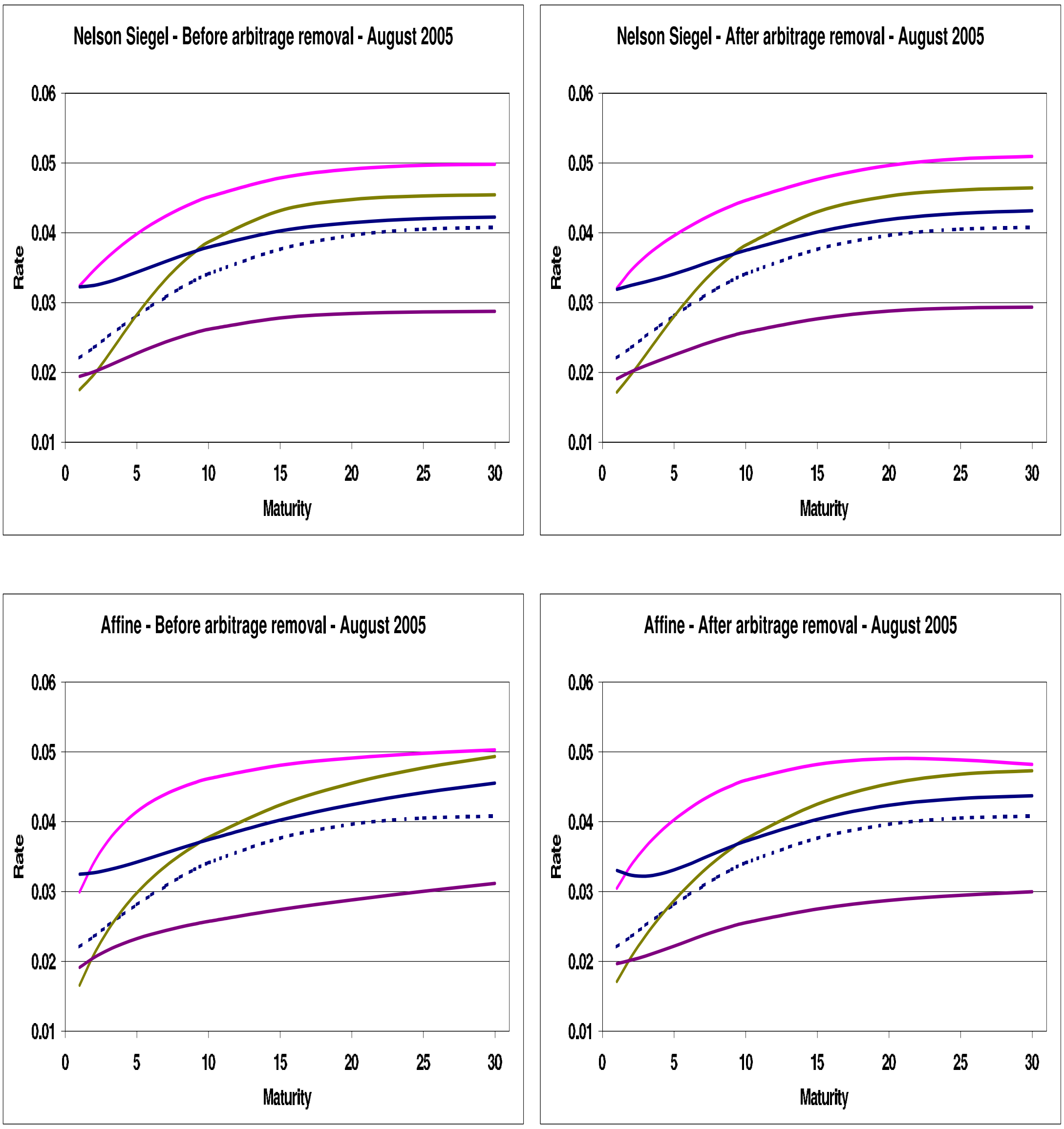

Figure 3: Each graph includes the observed yield curve on the 1th of August 2005 (the stippled curve). Four yield curve scenarios one year ahead are included as well. In the top figures the Nelson Siegel method is used to smooth the curves. In the bottom figures an affine function is used. Figures to the left are before removing arbitrage from the yield curves and figures to the right are after removal of arbitrage.

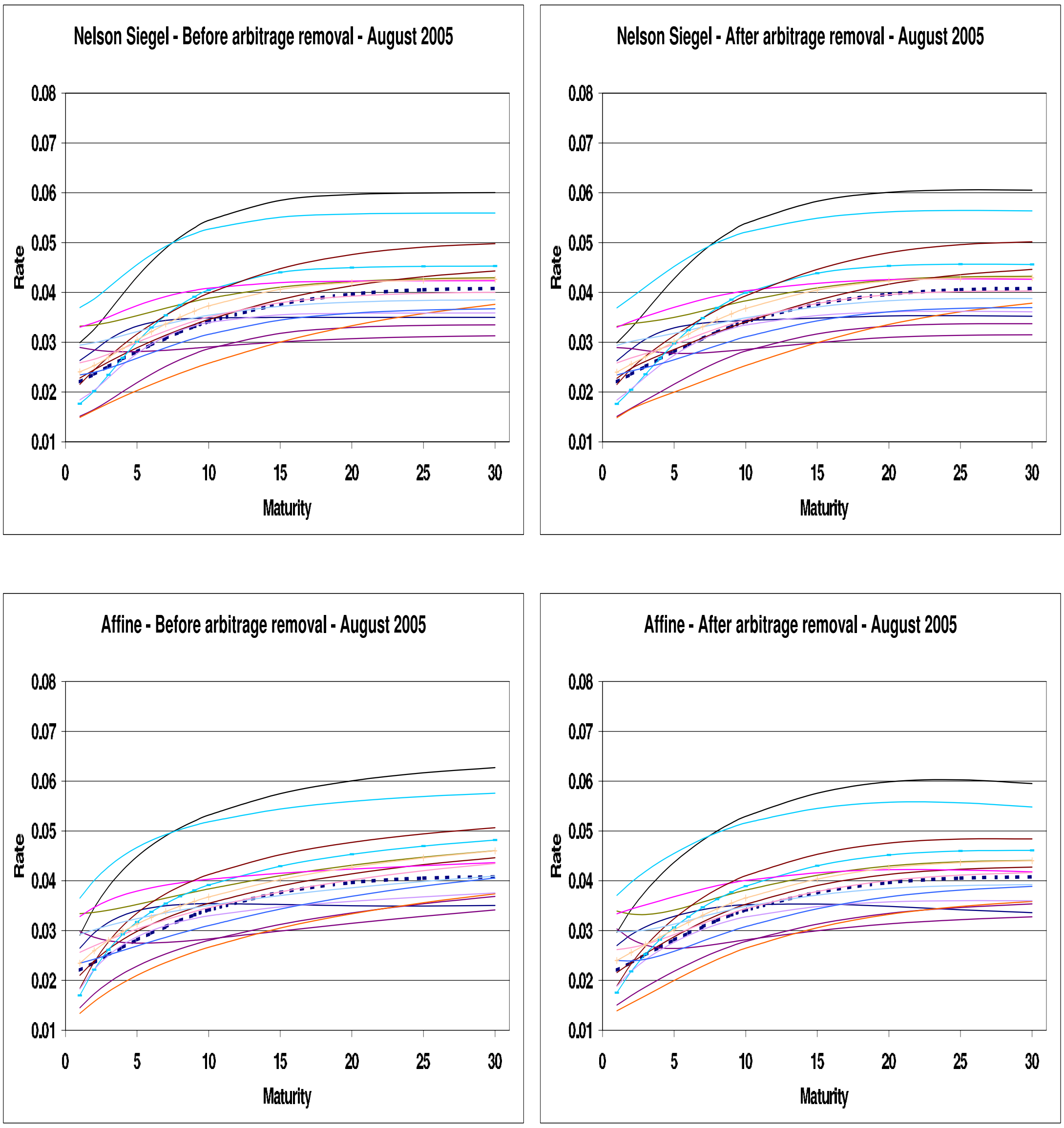

Figure 4: Each graph includes the observed yield curve on the 1th of August 2005 (the stippled curve). 16 yield curve scenarios one year ahead are included as well. In the top figures the Nelson Siegel method is used to smooth the curves. In the bottom figures an affine function is used. Figures to the left are before removing arbitrage from the yield curves and figures to the right are after removal of arbitrage.

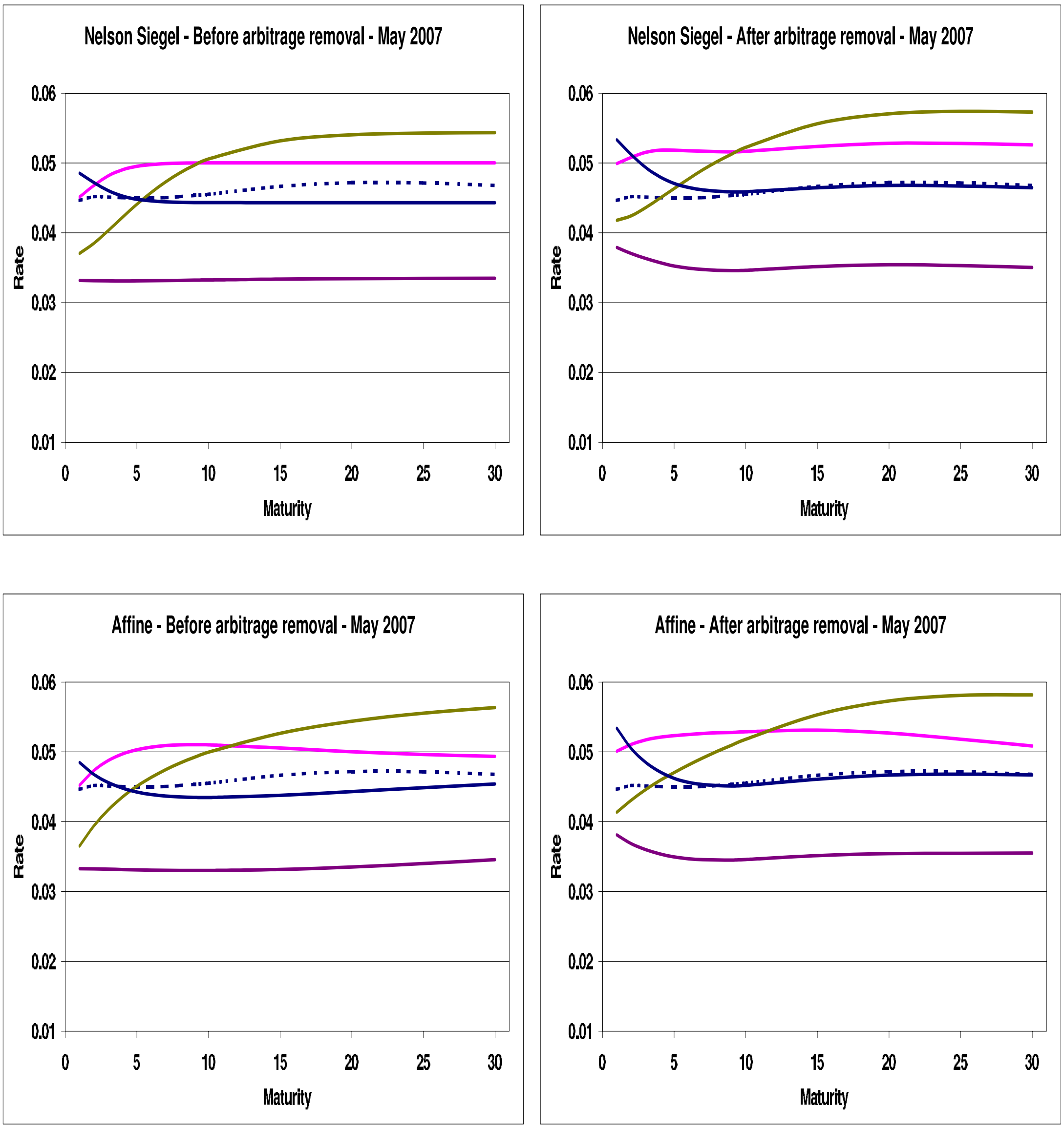

Figure 5: Each graph includes the observed yield curve on the 1th of May 2007 (the stippled curve). Four yield curve scenarios one year ahead are included as well. In the top figures the Nelson Siegel method is used to smooth the curves. In the bottom figures an affine function is used. Figures to the left are before removing arbitrage from the yield curves and figures to the right are after removal of arbitrage.

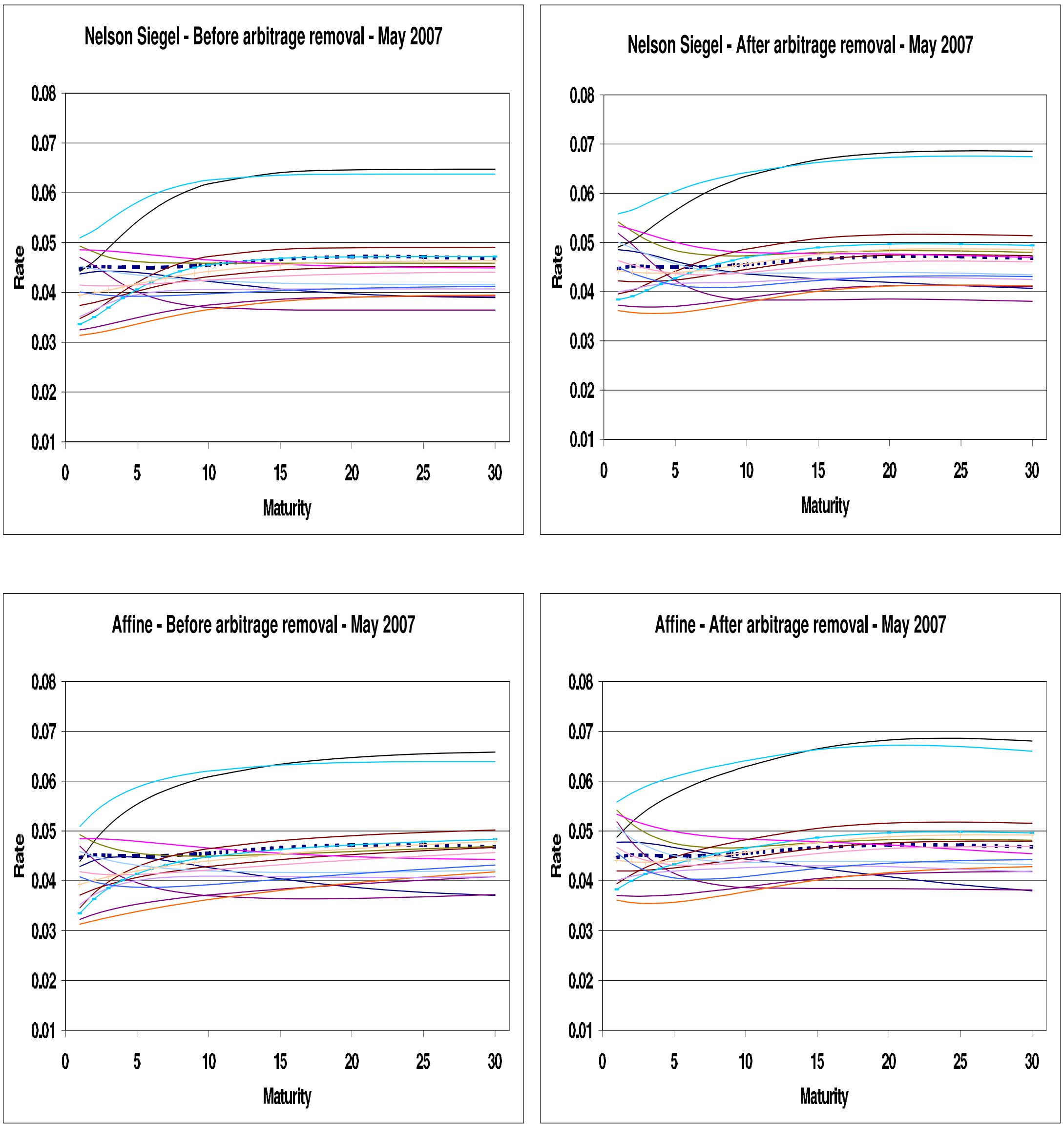

Figure 6: Each graph includes the observed yield curve on the 1th of May 2007 (the stippled curve). 16 yield curve scenarios one year ahead are included as well. In the top figures the Nelson Siegel method is used to smooth the curves. In the bottom figures an affine function is used. Figures to the left are before removing arbitrage from the yield curves and figures to the right are after removal of arbitrage.

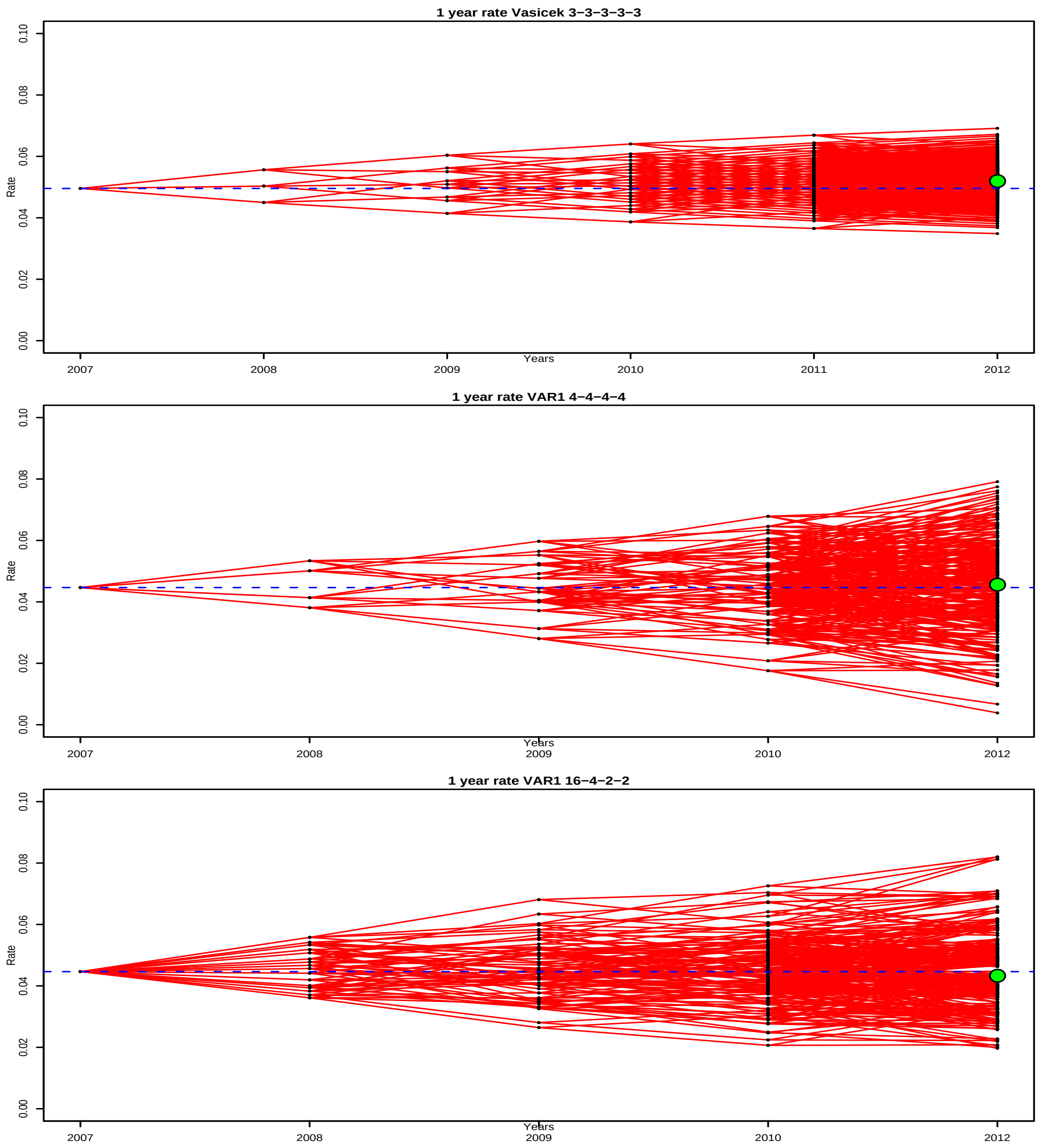

Figure 7: Scenario trees for 1–year rates over 5 years as produced by a 1–factor Vasicek model with a 3–3–3–3–3 discretization (top), our 3–factor VAR1 model with a 4–4–4–4 discretization (middle) and our 3–factor VAR1 model with a 16–4–2–2 discretization (down). The green circle shows the average level of scenarios. Note that there is a jump from year 3 to year 5 in the VAR1 trees.

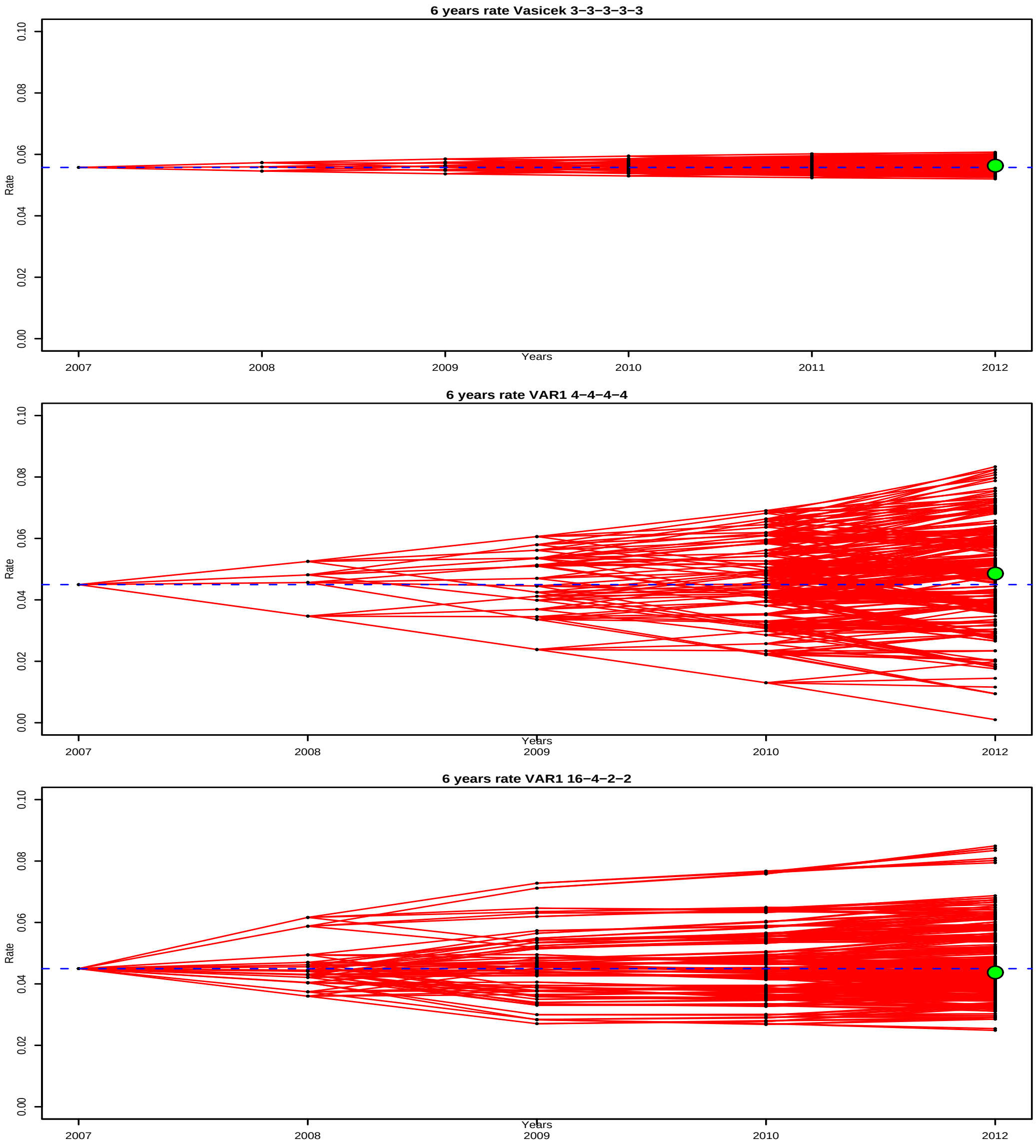

Figure 8: Scenario trees for 6–year rates over 5 years as produced by a 1–factor Vasicek model with a 3–3–3–3–3 discretization (top), our 3–factor VAR1 model with a 4–4–4–4 discretization (middle) and our 3–factor VAR1 model with a 16–4–2–2 discretization (down). The green circle shows the average level of scenarios. Note that there is a jump from year 3 to year 5 in the VAR1 trees.

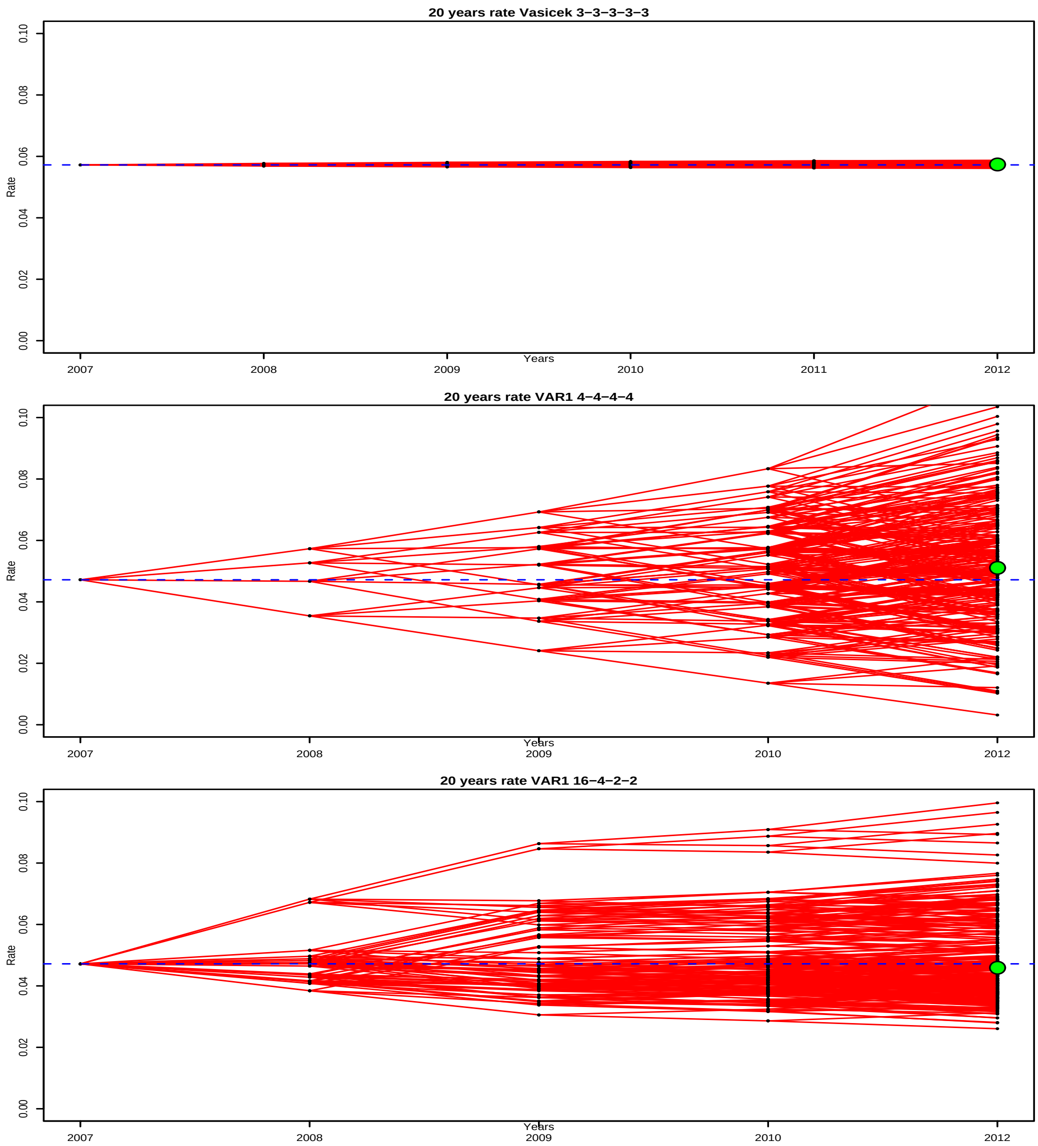

Figure 9: Scenario trees for 20–year rates over 5 years as produced by a 1–factor Vasicek model with a 3–3–3–3–3 discretization (top), our 3–factor VAR1 model with a 4–4–4–4 discretization (middle) and our 3–factor VAR1 model with a 16–4–2–2 discretization (down). The green circle shows the average level of scenarios. Note that there is a jump from year 3 to year 5 in the VAR1 trees.

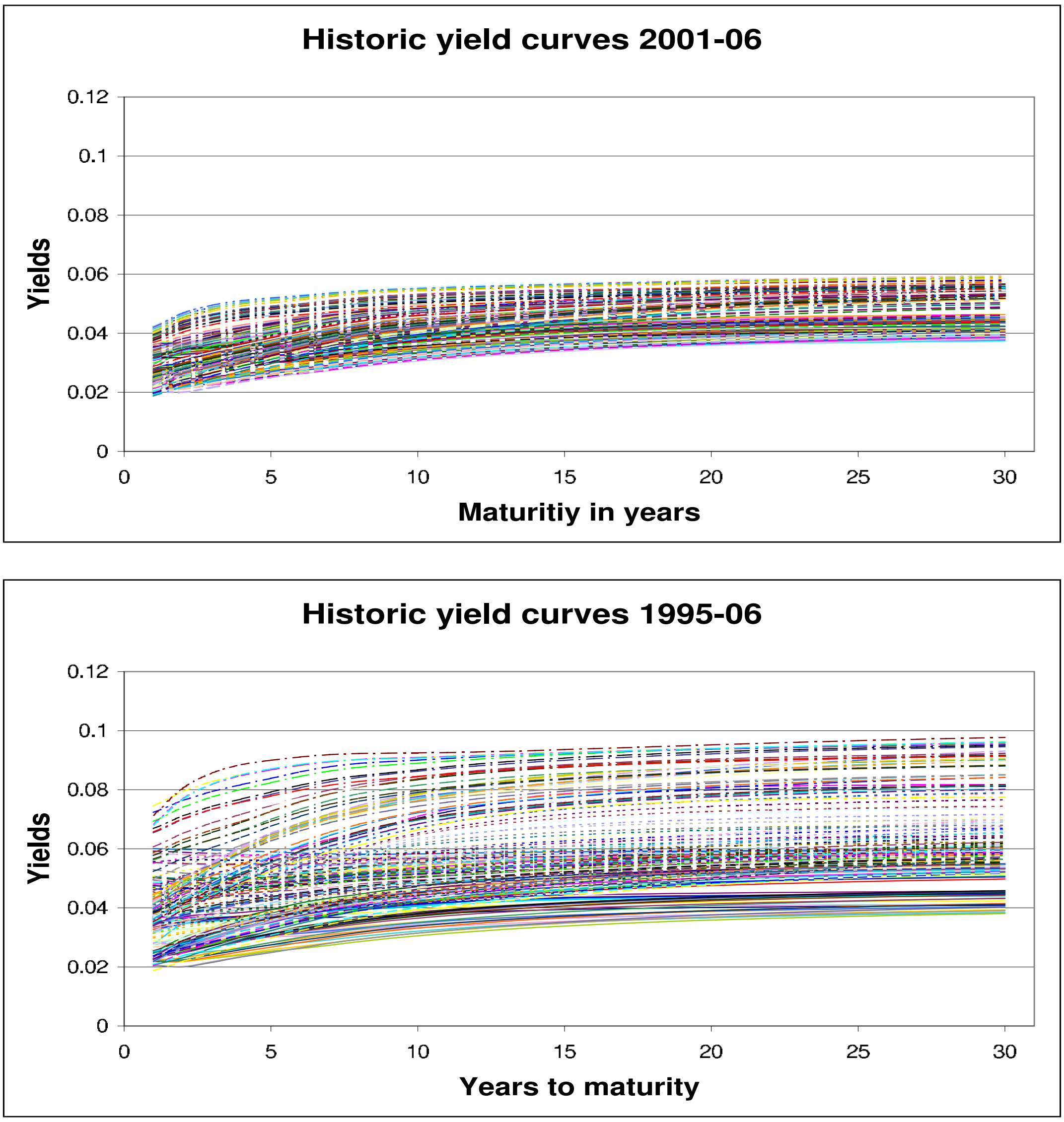

Figure 10: Historic yield curves from 2001 to 2006 (top) and from 1995 to 2006 (down).

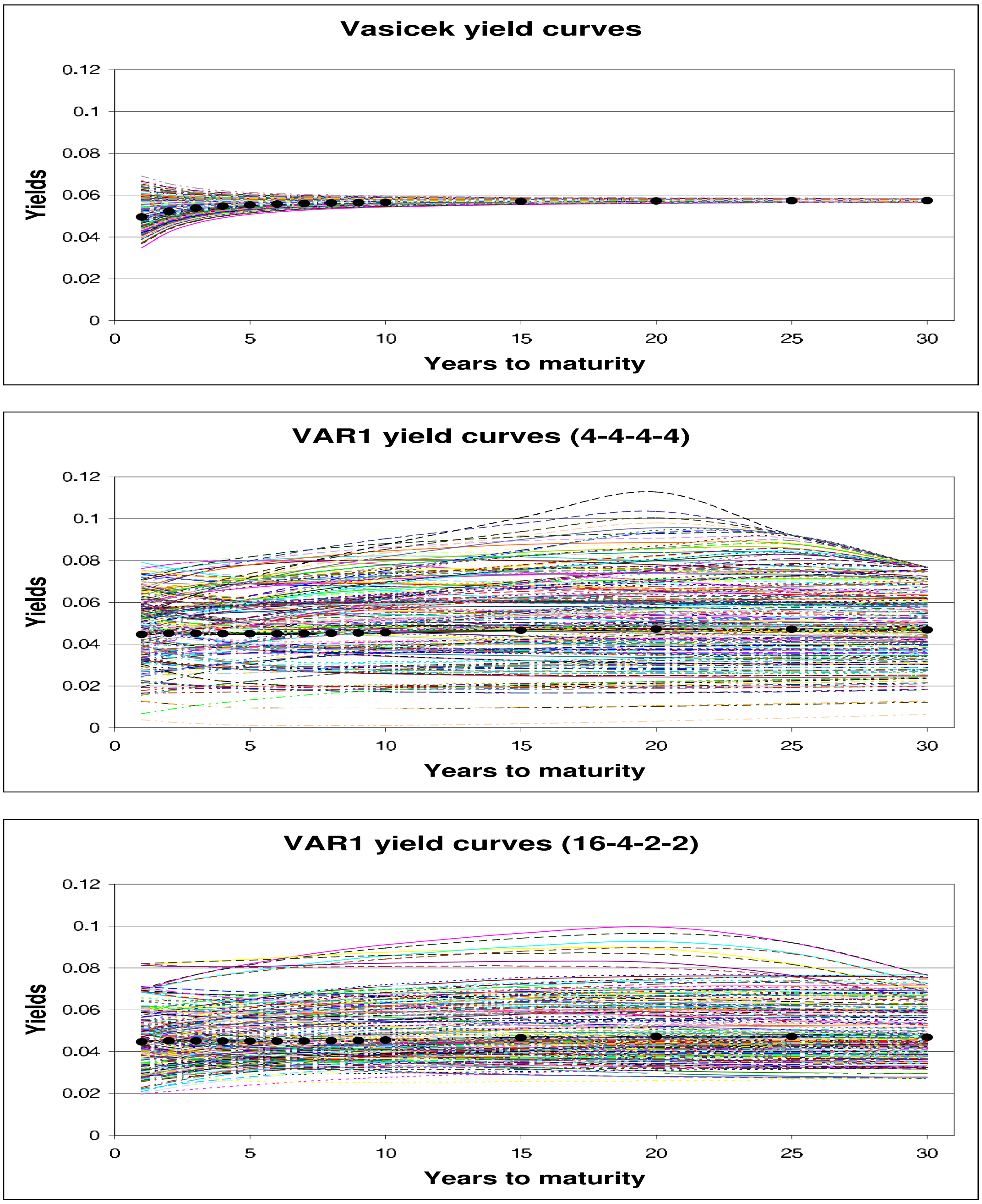

Figure 11: Yield curves generated 5 years from now (May 2007) using the 1–factor Vasicek model with a 3–3–3–3–3 discretization (top), our VAR1 model with a 4–4–4–4 discretization (middle) and our VAR1 model with a 16–4–2–2 discretization (down). The initial yield curve is also presented using solid lines for comparison.

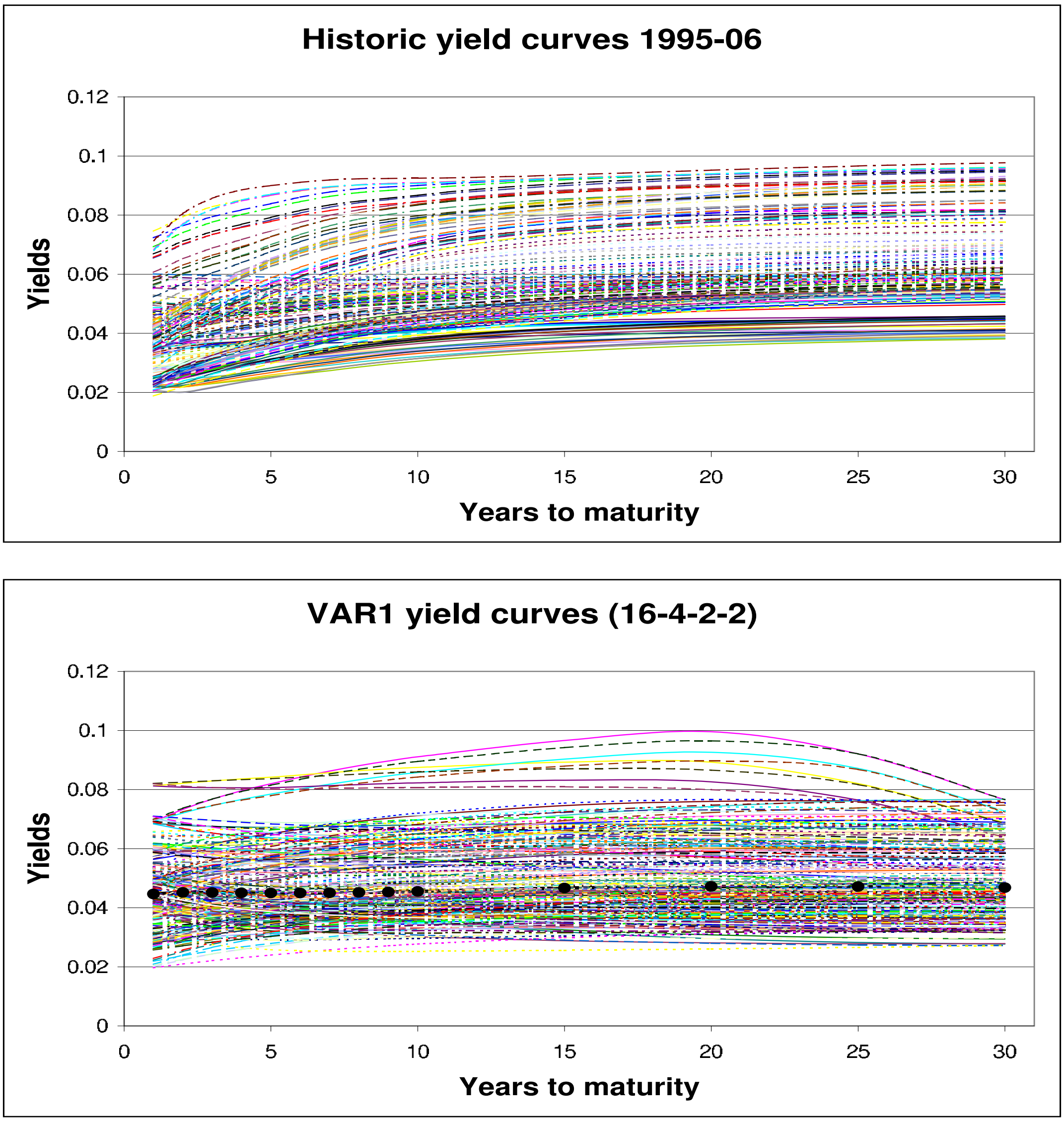

Figure 12: Comparison of the historic yield curves from 1995 to 2006 (top) with Yield curves generated 5 years from now (May 2007) using our VAR1 model with a 16–4–2–2 discretization (down).