Generating Yield Curve Stress-Scenarios

Arthur Charpentier, Christophe Villa

Generating Yield Curve Stress-Scenarios

Arthur CHARPENTIER   Christophe VILLA

March 23, 2010

Abstract: Several authors have proposed to combine movements in principal components to generate scenarios of "large" historical changes in term structures, i.e. stress-scenarios. This approach, however, has at least two shortcommings. This paper answers at these two problems and proposes a general two-steps procedure. The first step relies on fitting the discount bond yields and the second step relies on estimating statistically independent variables. Using the distribution of independent components identified, we combine their movements to produce stress-scenarios by specifying separate "shocks" in each of the directions given by the three independent components. We apply our methodology to the U.S. term structure of interest rates over the last three decades.

Keywords: Term structure; Yield curve; IRR; Stress-Testing; Scenario analysis; Principal Component Analysis; Independant Component Analysis.

I Introduction

While, stress testing is mostly used in managing market risk, a recent survey of financial firms by the Committee on the Global Financial Sytem (2005) found that stress tests based on movements in interest rates remain the main dominant type of stress test. The Basel Committee on Banking Supervision (BCBS, 2004) has issued several guidelines regarding interest rate risk (IRR) management that is defined as the change in a bank's portfolio value due to interest rate fluctuations both in the trading book and in the banking book. They pointed out that IRR management system should be *stressed* by examining uncommon, although not implausible, scenarios. As a wide range of shifts in the shape of the yield curve are observed in practice, the quality and the effectiveness of IRR management depends on the ability to generate relevant yield curve stress-scenarios. Indeed, in order to estimate their IRR exposure, banks should use multiple stress-scenarios, including potential effects in "large" changes in the relationships among interest rates as well as "large" changes in the general level of interest rates. Measuring IRR may seem to require specifying a very large number of perturbations of interest rates. In fact, interest rates movements are highly correlated contemporaneously and numerous studies have found that shifts or changes in the shape of the yield curve are attributable to three unobservable factors, which are often called *level*, *slope*, and *curvature*. The interpretation of these factors describes how the yield curve shifts or changes shape in response to a "shock" on a factor. These labels have turned out to be extremely useful in thinking about the driving forces of the yield curve until today and have important macroeconomic and monetary policy underpinnings (see Diebold, Piazzesi, and Rudebusch, 2005). The "effective dimensionality" of IRR is therefore considerably less than the number of large number of nominal bonds that are trading and held in a typical portfolio.

Principal Component Analysis (PCA) is often proposed as a tractable method for extracting such factors from yield curves (see Litterman and Scheinkman, 1991). Several authors have proposed to combine movements in principal components to produce scenarios as a method to: 1) separate computationally intensive fixed-income portfolio revaluations from the simulation step in Value-at-Risk by Monte Carlo, i.e. Scenario simulation (see Frye, 1996 and Jamshidian and Zhu, 1997); 2) generate scenarios of "large" historical changes in term structures, i.e. stress-scenarios (see Loretan, 1997 and Rodrigues, 1997). In fact, the approach involves creating a separate scenario for each possible combination of changes in the princpal components, say N_k for $k = 1, 2, 3$. Thus, with 3 principal components there are $N_1 \times N_2 \times N_3$ possible scenarios. The extreme outcomes for each principal component could be selected using either observed values in the tails of the empirical distribution or multiples of the standard deviation (with an assumption of elliptical distributions). Each actual scenario is derived by multiplying the corresponding principal component value by the yield sensitivities. This approach, however, has at least two shortcommings. Firstly, as pointed out by Diebold and Li (2006), PCA have unappealing features, including: (1) they cannot be used to produce yields at maturities other than those observed in the data, (2) they do not guarantee a smooth yield curve and forward curve,

(3) they do not guarantee positive forward rates at all horizons, and (4) they do not guarantee that the discount function starts at 1 and approaches 0 as maturity approaches infinity. Secondly, Fung and Hsieh (1996) have shown that during periods of large interest rate moves, the change in the shape of the yield curve is usually correlated to the level of interest rate itself. This fact means that specifying separate "shocks" in each of the directions given by the retained principal components is not appropriate to genetare stress scenarios. Ironically, this is also the key scenario of concern from the risk management perspective. This paper answers at these two problems.

In this paper we propose a general two-step procedure which is computationally feasible and can account for the dependence of interest rates at all available maturities. In the first step, we apply the term-structure model introduced by Nelson and Siegel (1987) which is popular among market and central bank practitioners and since it has been recently re-interpreted by Diebold and Li (2005) as a linear three-factor model of *level slope* and *curvature*. The second step relies on estimating three statistically independent components, as linear combination of level, slope and curvature factors. A popular method for solving the above problem is Independent Component Analysis (ICA), see Hyvärinen, Karhunen and Oja (2001). Using the distribution of independent components identified, we combine movements in ICs to produce stress-scenarios by specifying separate "shocks" in each of the directions given by the three independent components. We apply our methodology to the U.S. term structure of interest rates over the last three decades. The data used in this empirical analysis are the zero-coupon bond yields from January 1972 to December 2002. This sample period spans several major recessions and major expansions according to the NBER, and covers the terms of four Federal Reserve chairmen, namely Burns, Miller, Volcker, and Greenspan. The remainder of the paper proceeds as follows. Section I describes the procedures. Section II presents an empirical analysis based on the U.S. term structure of interest rates over the last three decades. Section III offers a summary and concluding comments.

II General Framework

A Parsimonious model of bond yields

Numerous studies have found that shifts or changes in the shape of the yield curve are attributable to three unobservable factors, which are often called *level*, *slope*, and *curvature* . The interpretation of these factors describes how the yield curve shifts or changes shape in response to a "shock" on a factor. These labels have turned out to be extremely useful in thinking about the driving forces of the yield curve until today and have important macroeconomic and monetary policy underpinnings (see Diebold, Piazzesi, and Rudebusch, 2005). The "effective dimensionality" of IRR is therefore considerably less than the number of large number of nominal bonds that are trading and held in a typical portfolio.

While various models have been developed and estimated to characterize the movement of these unobservable factors and the associated factor loadings that relate yields to different maturities to those factors, we claim in this paper that a particular interesting approach is the Diebold and Li (2006) reformulation of a fitted Nelson-Siegel curve (1987). A little background is required to understand what follows. Nelson and Siegel (1987) proposed the parsimonious yield curve model,

$$y_t(\tau) = b_{1t} + b_{2t}\left(\frac{1-e^{-\lambda_t\tau}}{\lambda_t\tau}\right) - b_{3t}e^{-\lambda_t\tau}$$

where $y_t(\tau)$ denotes the continuously-compounded zero-coupon nominal yield at maturity τ, and b_{1t}, b_{2t}, b_{3t} and λ_t are (time-varying) parameters. The Nelson-Siegel model can generate a variety of yield curve shapes including upward sloping, downward sloping, humped, and inversely humped. Recently, this model has been re-interpreted by Diebold and Li (2005) as a modern linear three-factor model. The corresponding yield curve is

$$y_t(\tau) = \beta_{1t} + \beta_{2t}\left(\frac{1-e^{-\lambda_t\tau}}{\lambda_t\tau}\right) + \beta_{3t}\left(\frac{1-e^{-\lambda_t\tau}}{\lambda_t\tau} - e^{-\lambda_t\tau}\right).$$

The advantage of this representation is that we can easily give economic interpretations to the parameters β_{1t}, β_{2t} and β_{3t}. In particular, we can interpret them as a *level factor* and two *shape factors*: a *slope factor*, and a *curvature factor*, respectively. To see this, note that the loading on β_{1t} is 1, a constant that doesn't depend on the maturity. Thus β_{1t} affects yields at different maturities equally and hence can be regarded as a *level factor*. The loading associated with β_{2t} is $\left(\frac{1-e^{-\lambda_t\tau}}{\lambda_t\tau}\right)$, which starts at 1 but decays monotonically to 0. Thus β_{2t} affects primarily short-term yields and hence changes the slope of the yield curve. Finally, factor β_{3t} has loading

$\left(\frac{1-e^{-\lambda_t \tau}}{\lambda_t \tau} - e^{-\lambda_t \tau}\right)$, which starts at 0, increases, and then decays. Thus β_{3t} has largest impact on medium-term yields and hence moves the curvature of the yield curve.

Rather than estimating the three factors by nonlinear least squares, Diebold and Li (2006) fix the value of $\lambda_t = \lambda$ (with maturities measured in months) and estimate the model for each period using ordinary least squares. They argue that this not only greatly simplifies the estimation, but likely results in more trustworthy estimates of the level, slope and curvature factors. More precisely, they set $\lambda = 0.0609$ precisely the value where the loading on the curvature factor reaches it maximum on the assumption that the curvature of the yield curve reaches its maximum at 30 months. In short, we can express the yield curve at any point of time as a linear combination of the level, slope and curvature factors, the dynamics of which drive the dynamics of the entire yield curve. Diebold, Rudebusch, and Aruoba (2005) examine the correlations between Nelson-Siegel yield factors and macroeconomic variables. While, they find that the level factor is highly correlated with inflation, and the slope factor is highly correlated with real activity, the curvature factor appears unrelated to any of the main macroeconomic variables.

B Independent Component Analysis

First of all, let us present an example of the problem of finding independent component in the data. A popular example is the so called cocktail party problem: Imagine a room full of people discussing with each other. A few microphones, located at different positions in the room, collect the sounds of mixed human voices and possible external noises. An outsider listening to the mixtures of sounds recorded by the microphone cannot decipher what was actually discussed in the room. The task now is to decompose the mixtures of sounds back into their original form, that is, human voices and external noises. These original sounds are called the latent sources, as they are "hidden" from the outsider listener. The task is often referred to as source separation. The computational methods discussed later in this paper are aimed at solving problems similar to this one.

More generally, assume that we observe n linear mixtures $x_1, ..., x_n$ of n independent components $s_1, ..., s_n$

$$x_j = a_{j1}s_1 + a_{j2}s_2 + ... + a_{jn}s_n, \text{ for all } j$$

Using a vector-matrix notation we note the ICA model as:

$$x = As$$

Denoting a_j the j^{th} columns of matrix A, the model becomes

$$x = \sum_{j=1}^{n} a_j s_j$$

The ICA model is a generative model, i.e., it describes how the observed data are generated by mixing the components s_i. The independent components are latent variables, i.e., not directly observable. The mixing matrix A is also unknown. We observe only the random vector x, and we must estimate both A and s. This must be done under as general assumptions as possible. ICA is a special case of blind source separation. Blind means that we know very little, if anything, on the mixing matrix, and make little assumptions on the source signals.

Like in standard factor models, ICA have two major ambiguities. The first ambiguity is that we cannot determine the variances of the independent components. The reason is that any scalar multiplier in one of the independent components could always be cancelled by dividing the corresponding column a_i of A by the same scalar. Whitening (sphering) of the independent components, i.e., choose all variances equal to one: $E\left(s_i^2\right) = 1$ can solve this ambiguity but the ambiguity of the sign remains, as we can multiply any independent component by -1 without affecting the model. The second ambiguity is that we cannot determine the order of the independent components. In fact we can freely change the order of the terms in the sum in $x = \sum_{j=1}^{n} a_j s_j$, and call any of the independent components the first one. A permutation matrix P and its inverse can be substituted in the model to give

$$x = AP^{-1}Ps = A's', \quad s' = Ps, \quad A' = AP^{-1}$$

There are two schools of thought with respect to what actually is the aim in estimating the ICs in the data. A first point of view is to regard ICA as a method (like factor analysis) of presenting the data in a more comprehensible way by revealing the hidden structure in the data and often reducing the dimensionality of the representation. According to this school of thought, data is tranformed as a combination of a few latent factors that are statistically as independent as possible. Second, one may regard the data being generated by a combination of some existing but unknown independent source signals and the task is to estimate them. This viewpoint is chosen in the so called Blind Source Separation framework — there are some sources which have been mixed, and the mixing process is completely unknown to us (hence the word "blind"). This paper mostly concentrates on the last viewpoint of ICA. There are several approaches to estimating the independent components, resulting in different algorithms. An interesting approach to ICA estimation are tensorial methods. Tensors are generalizations of linear operators — in particular, cumulant tensors are generalizations of the covariance matrix. Minimizing the higher order cumulants approximately amounts to higher order decorrelation, and can thus be used to solve the ICA model. The most well-known among these are the JADE (Joint Approximate Diagonalization of Eigenmatrices) algorithm, see Cardoso and Souloumiac (1993).

III JADE Algorithm

Cardoso (1999) has shown how higher-order correlations can be efficiently exploited to reveal independant component. These ICA algorithms use higher order two statistical information for separating the signals. Note that uncorrelatedness alone is not enough to separate the desired components. To keep the following exposition simple it is restricted to symmetric distributions. For any $n \times n$ matrix M, we define the associated cumulant matrix $T_x(M)$

as the $n \times n$ matrix defined component-wise by

$$[T_x(M)]_{ij} = \sum_{k,l} Cum(x_i, x_j, x_k, x_l) M_{kl}$$

where the subscript ij means the $(i,j)-$th element of a matrix T is a linear operator, and thus has n^2 eigenvalues that correspond to eigenmatrices. $Cum(x_1, x_2, x_3, x_4)$ are fourth-order cumulants defined by

$$Cum(x_1, x_2, x_3, x_4) = E[\bar{x}_1\bar{x}_2\bar{x}_3\bar{x}_4] - E[\bar{x}_1\bar{x}_2] E[\bar{x}_3\bar{x}_4] - E[\bar{x}_1\bar{x}_3] E[\bar{x}_2\bar{x}_4] - E[\bar{x}_1\bar{x}_4] E[\bar{x}_2\bar{x}_3]$$

where $\bar{x}_i = x_i - E[x_i]$. Recall that for symmetric distributions odd-order cummulants are zero and that second order cumulants are $Cum(x_1, x_2) = E[\bar{x}_1\bar{x}_2]$.

The variance and the kurtosis of a real random variable x are defined as

$$\begin{aligned} \sigma^2(x) &= Cum(x, x) = E\left[\bar{x}^2\right] \\ k(x) &= Cum(x, x, x, x) = E\left[\bar{x}^4\right] - 3E^2\left[\bar{x}^2\right] \end{aligned}$$

that is, they are the second and fourth-order autocumulants. A cumulant involving at least two different variables is called a cross-cumulant.

Under $x = As$ which also reads $x_i = \sum_p a_{ip} s_p$ where a_{ij} denotes the $(ij)-$th entry of matrix A the cumulants of order 4 transform can be written as:

$$Cum\left(x_i, x_j, x_k, x_l\right) = \sum_{pqrs} a_{ip}a_{jq}a_{kr}a_{ls}Cum\left(s_p, s_q, s_r, s_s\right)$$

Using the assumption of independence of s by which $Cum\left(s_p, s_q, s_r, s_s\right) = k\left(s_p\right)\delta\left(p, q, r, s\right)$, we readily obtain the simple algebraic structure of the cumulants of $x = As$:

$$Cum\left(x_i, x_j, x_k, x_l\right) = \sum_{u=1}^{n} k\left(s_u\right) a_{iu}a_{ju}a_{ku}a_{lu}$$

The structure of a cumulant matrix in the ICA model is easely deduced from this last equation that:

$$\begin{aligned} T_x\left(M\right) &= A\Lambda\left(M\right)A' \\ \Lambda\left(M\right) &= diag\left(k\left(s_1\right)a_1'Ma_1, ..., k\left(s_n\right)a_n'Ma_n\right) \end{aligned}$$

where a_i denotes the $i-$th column of A. In this factorization the kurtosis enter only in the diagonal matrix. Solving for the eigenvectors of such eigenmatrices, the ICA model can be estimated. This is typically a joint diagonalization of several matrices but the most difficulty is that A is not an orthogonal matrix. The following section outlines the JADE algorithm (Cardoso and Souloumiac, 1993). The approach for the JADE algorithm is the following two-stage procedure : Whitening and Rotation.

JADE like some other ICA algorithms require a preliminary whitening of the data x. This means that before the application of the algorithm (and after centering), we transform the observed vector x linearly so that we obtain a new vector $\tilde{x}$ which is white, ($\tilde{x} = Qx$) i.e. its components are uncorrelated and their variances equal unity. In other words $E[\tilde{x}\tilde{x}'] = I$. The whitening transformation is always possible. The whitening transform Q can be determined by taking the inverse square root of the covariance matrix via an eigenvalue decomposition of the covariance matrix or by a PCA. Denoting Γ the orthogonal matrix of eigenvectors of $E[xx']$ and Λ is the diagonal matrix of its eigenvalues, one has $E[xx'] = \Gamma\Lambda\Gamma'$. Whitening can now be performed by

$$\tilde{x} = \Lambda^{-\frac{1}{2}}\Gamma' x$$

that is $Q = \Lambda^{-\frac{1}{2}}\Gamma'$.

It is easy to check that $E[\tilde{x}\tilde{x}'] = I$ by taking the expectation of $\tilde{x}\tilde{x}' = \Lambda^{-\frac{1}{2}}\Gamma' x \left[\Lambda^{-\frac{1}{2}}\Gamma' x\right]' = \Lambda^{-\frac{1}{2}}\Gamma' x x' \Gamma \Lambda^{-\frac{1}{2}}$. Since $x = As$ and after whitening $\tilde{x} = Qx$, one has

$$\tilde{x} = \tilde{A}s$$

where $\tilde{A} = QA$. It can be easily shown that $\tilde{A}$ is an orthogonal matrix. Indeed

$$E\left[\tilde{x}\tilde{x}'\right] = I = \tilde{A}E\left[ss'\right]\tilde{A}' = \tilde{A}\tilde{A}'$$

Recall that we assumed that the independent components s_i have unit variance. Whitening reduced the problem of finding an arbitrary matrix in model $x = As$ to the simpler problem of finding an orthogonal matrix $\tilde{A}$. Once it is found, $\tilde{A}$ is used to solve the independent components from the observed by

$$\hat{s} = \tilde{A}^{-1}\tilde{x} = \tilde{A}'\tilde{x}$$

A couple of remarks merites to be done. First, whitening alone does not solve the separation problem. This is because whitening is only defined up to an additional rotation: if Q_1 is a whitening matrix, then $Q_2 = UQ_1$ is also a whitening matrix if and only if U is an orthogonal matrix. Therefore, we have to find the correct whitening matrix that equally separates the independent components. This is done by first finding any whitening matrix Q, and later determining the appropriate orthogonal transformation from a suitable non-quadratic criterion. Second, whitening reduces the number of parameters to be estimated. Instead of having to estimate the n^2 parameters that are the elements of the original matrix A, we only need to estimate the new, orthogonal mixing matrix $\tilde{A}$. An orthogonal matrix contains $\frac{n(n-1)}{2}$ degrees of freedom. In larger dimensions, an orthogonal matrix contains only about half of the number of parameters of an arbitrary matrix. Thus one can say that whitening solves half of the problem of ICA. Because whitening is a very simple and standard procedure, much simpler than any ICA algorithms, it is a good idea to reduce the complexity of the problem this way.

For any $n \times n$ matrix M, we can define the associated cumulant matrix

$T_{\tilde{x}}(M)$ defined component-wise by

$$[T_{\tilde{x}}(M)]_{ij} = \sum_{k,l} Cum(\tilde{x}_i, \tilde{x}_j, \tilde{x}_k, \tilde{x}_l) M_{kl}$$

where the subscript ij means the $(i,j)-$th element of a matrix T. We have shown that whitening yields to the model $\tilde{x} = \tilde{A}s$ with $\tilde{A}$ orthonormal. This model is still a model of independant components. From the above section the structure of the corresponding cumulant matrix of $\tilde{x}$ can be written as:

$$\begin{aligned} T_{\tilde{x}}(M) &= \tilde{A}\tilde{\Lambda}(M)\tilde{A}' \\ \tilde{\Lambda}(M) &= diag\left(k(s_1)\tilde{a}_1' M\tilde{a}_1, ..., k(s_n)\tilde{a}_n' M\tilde{a}_n\right) \end{aligned}$$

where $\tilde{a}_i$ denotes the $i-$th column of $\tilde{A}$ and for any $n \times n$ matrix M.

Let $\Pi = \{M_1, M_2, ..., M_p\}$ be a set of p matrices of size $n \times n$ and denote by $T_{\tilde{x}}(M_i)$ $1 \leq i \leq p$ the associated cumulant matrices of the whitened data $\tilde{x} = \tilde{A}s$. Again for all i we have $T_{\tilde{x}}(M_i) = \tilde{A}\tilde{\Lambda}(M_i)\tilde{A}'$ with $\tilde{\Lambda}(M_i)$ a diagonal matrix. As a measure of nondiagonality of a matrix H, define Off(H) as the sum of the squares of the non diagonal elements: Off$(H) = \sum_{i \neq j}(H_{ij})^2$. We have in particular $\mathrm{Off}\left(\tilde{A}' T_{\tilde{x}}(M_i)\tilde{A}\right) = \mathrm{Off}\left(\tilde{\Lambda}(M_i)\right) = 0$ since $T_{\tilde{x}}(M_i) = \tilde{A}\tilde{\Lambda}(M_i)\tilde{A}'$ and $\tilde{A}\tilde{A}' = I$.

For any matrix set Π and any orthonormal matrix V, the Jade algorithm optimize an orthogonal contrast

$$\arg\min \sum_{i} Off\left(V' T_{\tilde{x}}\left(M_i\right) V\right)$$

This criterion measures how close to diagonality an orthonormal matrix V can simultaneously bring the cumulants matrices generated by Π. With JADE algorithm, this joint diagonalizer is found by a jacobi technique.

IV Summary

We propose here a sequential procedure for yield curve stress-scenarios. Our procedure follows the generic strategy that is divided in three steps:

1. We express the yield curve at any point of time, t, (each month) as a linear combination of the *level*, *slope* and *curvature* factors (β_{1t}, β_{2t} and β_{3t}) and estimate the model for each period using ordinary least squares. At that stage the yield curve is summarized by 3 coefficients. We then assume a h months holding period and we calculate $\Delta\beta_{it} = \beta_{i(t+h)} - \beta_{it},\ i = 1, 2, 3$. This is due to the linearity of the model: factors affecting bond yield changes are the changes of the estimated bond yield factors.

2. We employ an the JADE algorithm to estimate three independent component $s_j,\ j = 1, 2, 3$ as a linear combination of $\Delta\beta_i$, $i = 1, 2, 3$.

3. Once these independent components have been computed, we use their distribution and combine movements in each independent component

to produce stress-scenarios by specifying separate "shocks" in each of the directions given by the three independent components. Yield curve scenarios are then computed by multiplying the value of each independent component to first, factor sensivities and second, yield sensitivities.

V Empirics

A The data

We use end-of-month price quotes (bid-ask average) for U.S. Treasuries, from January 1972 through December 2002, taken from the CRSP government bonds files (372 months). CRSP filters the data, eliminating bonds with option features (callable and flower bonds), and bonds with special liquidity problems (notes and bonds with less than one year to maturity, and bills with less than one month to maturity), and then converts the filtered bond prices to unsmoothed Fama-Bliss (1987) forward rates. Then, using programs and CRSP data kindly supplied by Rob Bliss, we convert the unsmoothed Fama-Bliss forward rates into unsmoothed Fama-Bliss zero yields. At each month, we consider a set of fixed maturities: the maturities spanned include 3 months to 12 months at one month increment, then at three months increment to the 5 year maturity, and finally at six months increment to the 10 year maturity, where a month is defined as 30.4375 days. These cover the range of bond maturities, and they also reflect the different trading volumes at different maturities. In particular, at the short end where bonds concentrate, we include more fixed maturities, and at the long end where there are fewer bonds, we include fewer fixed maturities. We did not use the one month or the two month yield to represent the short end, because they are more likely to be influenced by liquidity needs (see Duffee, 1996). According to the NBER, this sample period contains several major recessions and major expansions.[1] Several major historical and economic events occurred during our period of analysis (e.g. the Vietnam war, the oil price shocks, the monetary experiment, the 1987 crash, the Gulf war), among which some strongly impacted U.S. interest rates. Moreover, in this sample there have been four different Federal Reserve chairmen (see Thornton, 1996, for more details): Arthur F. Burns (February 1970 - January 1978), G. William Miller (March 1978 - August 1979), Paul A. Volcker (August 1979 - August 1987), and finally Alan Greenspan (August 1987 - present). The various yield as well as the yield curve level, slope and curvature defined above, will play a prominent role in the sequel. Hence we focus on them now in some details. In Fig. 1 we provide a three-dimensional plot of our yield curve data. The large amount of temporal variation in the yield curve is visually important. Moreover we consider a one year holding period and we plot in the same vein as for yield curves a three-dimensional plot of changes in our yield curve data in Fig. 2. Again a large amount of temporal variation in the yield change is visually important.

< Insert Figures 1 and 2>

[1] The NBER peaks are 1969:12, 1973:11, 1980:01, 1981:07, 1990:07, and 2001:03, and the NBER troughs are 1970:11, 1975:03, 1980:07, 1982:11, and 1991:03.

B Generating standardised stress-scenarios

In the first step, as in Fama and Bliss (1987), we use a bootstrap method to infer zero bond yields from available bill, note, and bond prices. In the second step, we treat the factor loadings in the above equation as regressors and we lambda equal to 0.0609 to calculate the regressor values for each zero bond. In the third step, we run a cross-sectional regression of the zero yields on the calculated regressor values. The regression coefficients are the estimated factor values. We do this in each month to get the time series of three factors. As discussed above, we fit the yield curve using the three-factor model,

$$y_t(\tau) = \beta_{1t} + \beta_{2t}\left(\frac{1-e^{-\lambda\tau}}{\lambda\tau}\right) + \beta_{3t}\left(\frac{1-e^{-\lambda\tau}}{\lambda\tau} - e^{-\lambda\tau}\right).$$

and estimate it through OLS. The residual plot in Fig. 3 and Mean Squared Errors plots in Fig. 4 indicate a good fit. In Fig. 5 we plot the implied average fitted yield curve (resp. yield curve change) against the average actual yield curve (resp. yield curve change). Both the bootstrapped zero yields and the three-factor fitted yield curves are included. The two agree quite closely.

< Insert Figures 3, 4, 5>

In Fig. 6 (resp. Fig. 7) we plot $\left\{\hat{\beta}_{1t}, \hat{\beta}_{2t}, \hat{\beta}_{3t}\right\}$ (resp. $\left\{\Delta\hat{\beta}_{1t}, \Delta\hat{\beta}_{2t}, \Delta\hat{\beta}_{3t}\right\}$). Scatter plots of the changes of Level, Slope and Curvature Factors is presented in Fig. 8 What we can say is the deviation from normality in the data.

< Insert Figures 6, 7, 8>

In selecting the standardised intrest rate shock the following guiding principle for IRR exposures in G-10 currencies is proposed by the BCBS (2004, Annex 3, p. 36): banks should consider either a parallel rate change of ± 200 basis points or the changes implied by the 1^{st} and 99^{th} percentiles of historically observed interest rate changes over at least five years (using a one year holding period). A one year holding period was selected by the BCBS both for practical purposes and in recognition that within a one-year period most institutions have the ability to restructure or hedge their positions to mitigate further losses in economic value should rates appear exceptionally volatile. We use a similar strategy in our generation of stress-scenarios.

How does one specify scenarios? Consider first the case where a single independant component is estimated. Since an independent component is a one-to-one transformation of the observed data, it is possible to "reverse" the calculations and to compute the values of each of the series that correspond to given values of this independent component. Next, since the independent component is a random variable we may pick tailevent quantiles of the empirical distribution of the independent component to generate corresponding tail events of the observable series. When more than one independent component is added, one may proceed by specifying separate "shocks" in each of the directions given by the retained independent components, in analogy to the case of a single one. Alternatively, one may choose to form arbitrary linear combinations of the estimated independent components to generate "combined" shocks. For each of the three types of independent stress-scenarios, the following quantiles of the resulting distributions are reported: 0.5%, 1%, 5%, 10%, 90% 95%, 99%, and 99.5%. By measuring the exposure to shocks of increasing severity-from 10% to 0.5%, and from 90% to 99.5%-it may be possible to determine if there is "curvature" in the exposure, i.e., if there is gamma risk that could lead to systemic breakdowns if these exposures are hedged by dynamic trading strategies. Note that the quantiles of the shock distributions should not be interpreted as meaning that any of these particular scenarios will occur with the specified probabilities; "real world" shocks are combinations of the shocks in the directions of the various IC-shocks. The results are listed in Figure 9. They show that shocks in the direction of the first independent component seems lead to a "shift" in all rates. The second scenario is a "tilt" of the yield curve, and the third serves to increase or decrease curvature. However, it should be remembered that they are "pure factor shocks," and that "actual" shocks are combinations of the "pure" shocks. In Figure 10 we present simultaneous shock to the three independent components. The following quantiles of the resulting distributions are reported: 1% and 99% which lead to $2^3 = 8$ scenarios.

< Insert Figures 9 and 10>

VI Conclusion

In support of the revised Basel Capital Accord, which was released in June 2004 and is to be fully implemented by year-end 2007, the Basel Committee on Banking Supervision (BCBS, 2004) has issued several guidelines regarding interest rate risk (IRR) management that is defined as the change in a bank's portfolio value due to interest rate fluctuations both in the trading book and in the banking book. A wide range of shifts in the shape of the yield curve are observed in practice. Therefore the quality and the effectiveness of IRR management depends on the ability to identify relevant yield curve shocks. Taking on IRR is a key part of what banks do; but taking on excessive IRR could threaten a bank's earnings and its capital base, raising concerns for bank supervisors. For instance, in order to facilitate supervisory monitoring of IRR exposures across institutions, banks should try to use a standardized interest rate shock to provide the results of their internal measurement systems, expressed in terms of changes to economic value. These rate shocks should in principle be determined by banks but based on the recommended criteria. For example, the following guiding principle for IRR exposures in G-10 currencies is proposed: banks should consider either a parallel rate change of ± 200 basis points or the changes implied by the 1^{st} and 99^{th} percentiles of historically observed interest rate changes over at least five years (using a one year holding period both for practical purposes and in recognition that within a one-year period most institutions have the ability to restructure or hedge their positions to mitigate further losses in economic value should rates appear exceptionally volatile).

However recognizing the relative simplicity of this parallel rate shock, supervisors will continue to expect institutions to examine multiple shocks that include yield curve twists, inversions, and other relevant shocks in evaluating the appropriate level of their IRR exposures. This paper has described a methodology for summarising historical movements in interest rates across government term structures. The methodology could be used to construct scenarios of large historical changes in term structures. These scenarios could be used both for computing the exposure of a financial firm's trading book to term structure movements as well as for stress testing purposes. The outputs from the scenarios could also be aggregated over market-making firms to analyse the impact of large term structure shifts on aggregate profit & loss statements.

References

[1] Anderson, N and Sleath, J, 2001, New estimates of UK yield curves, Bank of England Working Paper no. 126.

[2] Ang, A. and Piazzesi, M. (2003), "A No-Arbitrage Vector Autoregression of Term Structure Dynamics with Macroeconomic and Latent Variables," Journal of Monetary Economics, 50, 745-787.

[3] Back, A. and Weigend, A. 1998. A first application of independent component analysis to extracting structure from stock returns, International Journal of Neural Systems 8: 473–484.

[4] Balanda, K. P., and MacGillivray, H. L. (1988), "Kurtosis: A Critical Review," The American Statistician, 42, 111–119.

[5] Barnett, V., and Lewis, T. (1978) Outliers in Statistical Data, New York: Wiley.

[6] Beim, D.O., 1992. Term structure and the non-cash value in bonds, Working Paper 92-40. Graduate School of Business, Columbia University.

[7] Basel Committee on Banking Supervision. 2004. "Principles for the Management and Supervision of Interest Rate Risk." http://www.bis.org/publ/bcbs108.pdf

[8] Board of Governors of the Federal Reserve System. 2003. "Risk Management and Valuation of Mortgage Servicing Assets Aris-

ing from Mortgage Banking Activities." Supervisory Letter SR03-4. http://www.federalreserve.gov/boarddocs/SRLETTERS/2003/sr0304.htm

[9] Bomfim A. N., 2003, Monetary Policy and the Yield Curve, Working Paper, Federal Reserve Board

[10] Bliss, R. R. (1997) : Movements in the term structure of interest rates, Economic Review, Fourth Quarter, pp 16-33.

[11] Bogner, R. E. (1992). Blind separation of sources, Technical Report 4559, Defence Research Agency, Malvern.

[12] Cardoso, J. (1989). Source separation using higher order moments, International Conference on Acoustics, Speech and Signal Processing, pp. 21092112.

[13] Cardoso, J. and Souloumiac, A. (1993). Blind beamforming for non-Gaussian signals, IEE Proc. F. 140(6): 771774.

[14] Chaumeton, Lucie, Gregory Connor and Ross Curds (1996): "A Global Stock and Bond Model", Financial Analysts Journal, vol. 52, no. 6, November/December, pp. 65-74.

[15] Comon, P. (1994). Independent component analysis a new concept?, Signal Processing 36(3): 287314.

[16] Derivatives Policy Group (1995): Framework for Voluntary Oversight. To be accessed at //risk.ifci.ch/137790.htm

[17] Dewachter, H. and Lyrio, M. (2003), "Macroeconomic Factors in the Term Structure of Interest Rates," Manuscript, Catholic University of Leuven and Erasmus University Rotterdam.

[18] Diebold, F.X., Rudebusch, G.D. and Aruoba, S.B. (2003), "The Macroeconomy and the Yield Curve: A Non-Structural Analysis," Working Paper, University of Pennsylvania.

[19] Diebold, F. X. and C. Li (2006). Forecasting the term structure of government bond yields. Journal of Econometrics 130, 337–364.

[20] Duffee, G., 1996, Idiosyncratic Variation of Treasury Bill Yields, *Journal of Finance*, 51, 527-551.

[21] Estrella, Arturo, and Gikas A. Hardouvelis. 1991, The Term Structure as a Predictor of Real Economic Activity. *The Journal of Finance* XLVI, No. 2, 555-576.

[22] Evans, M. D. D. 2003. Real Risk, Inflation Risk, and The Term Structure, *The Economic Journal*, pages 345{389.

[23] Evans, C.L. and Marshall, D. (2001), "Economic Determinants of the Nominal Treasury Yield Curve," Manuscript, Federal Reserve Bank of Chicago.

[24] Frye, Jon (1996): "Principals of Risk: Finding Value-at-Risk through Factor-based Interest Rate Scenarios, manuscript, Chicago: Nationsbanc-CRT (June).

[25] Ferguson, R. and S. Raymar, 1998, A comparison analysis of several popular term structure estimation methods, *Journal of Fixed Income*, March, 17-32

[26] Fama E. F., Bliss R. R., 1987, The Information in Long-Maturity Forward Rates, *American Economic Review* 77, 680-692.

[27] Friedman, J. H., and Tukey, J. W., 1974, "A Projection Pursuit Algorithm for Exploratory Data Analysis," IEEE Transactions on Computers, C-23, 881–889.

[28] Fung, W., and Hsieh, D. A., 1996, Event Risk and Risk Management of Global Yield Curve Exposure, Journal of Fixed Income, 1996, Vol 6, No 2, 37-48.

[29] Harvey, Campbell R. The Real Term Structure and Consumption Growth. *Journal of Financial Economics* 22, 1988, 305-333.

[30] Hördahl, P., Tristani, O. and Vestin, D., 2002, "A Joint Econometric Model of Macroeconomic and Term Structure Dynamics," Manuscript, European Central Bank.

[31] Huber, P., 1985, "Projection Pursuit (with discussion)", The Annals of Statistics, 13,435-525.

[32] Hyvärinen A., Karhunen J., and Oja E., 2001, Independent Component Analysis. Wiley Interscience.

[33] Ioannides, M., 2003, A comparison of yield curve estimation techniques using UK data, *Journal of Banking and Finance* 27, 1-26

[34] Jutten C. and Herault J., 1991, Blind separation of sources, part I: An adaptive algorithm based on neuromimetic architecture. Signal Processing, 24:1–10, 1991.

[35] Jamshidian, Farshid, and Yu Zhu, 1997, Scenario Simulation Model: Theory and Methodology. Finance and Stochastics, vol. 1, no. 1.

[36] Jones, M.C. and Sibson, R., 1987, "What is Projection Pursuit (with discussion)?", Journal of the Royal Statistical Society A, 150, 1-36.

[37] Kambhu, John and Anthony Rodrigues (1997): ”Stress Tests and Portfolio Composition”, Mimeo, Federal Reserve Bank of New York.

[38] Knez, Peter, Robert Litterman and José Scheinkman (1994): ”Explorations into Factors Explaining Money Market Returns”, The Journal of Finance, vol. XLIX, no. 5, December, pp. 1861-1882.

[39] Litterman, Robert and Thomas Iben (1991): ”Corporate Bond Valuation and the Term Structure of Credit Spreads”, The Journal of Portfolio Management, Spring, pp. 52-64.

[40] Litterman, Robert and José Scheinkman (1991): ”Common Factors Affecting Bond Returns” The Journal of Fixed Income, June, pp. 54-61.

[41] Linton, O., E. Mammen, J. Nielsen, and C. Tanggaard, 2003, Estimating Yield Curves by Kernel Smoothing Methods, forthcomming *Journal of Econometrics.*

[42] Lopez, J.A. 2003. "Disclosure as a Supervisory Tool: Pillar 3 of Basel II." FRBSF Economic Letter 2003-22 (August 1). http://www.frbsf.org/publications/economics/letter/2003/el2003-22.html

[43] Loretan M., 1997, Generating market risk scenarios using principal components analysis: Methodological and practical considerations. In The Measurement of Aggregate Market Risk, CGFS Publications No. 7, pages 23–60. Bank for International Settlements. Available at http://www.bis.org/publ/ecsc07.htm.

[44] Mankiw G., Miron J., 1986, The Changing Behavior of the Term Structure of Interest Rates, *Quarterly Journal of Economics* 101, 211-228.

[45] Nelson C. R., Siegel A. F., 1987, Parsimonious Modeling of Yield Curves, *Journal of Business* 60, 473-489.

[46] Shea, G.S., 1984. Pitfalls in smoothing interest rate term structure data: equilibrium models and spline approximations. *Journal of Financial and Quantitative Analysis* 19 3 , 253–269.

[47] Siegel, A.F. and Nelson, C.R. 1988, Long-term Behavior of Yield Curves, *Journal of Financial and Quantitative Analysis*, 23, 105-110.

[48] Svensson, L.E. 1994. Estimating and Interpreting Forward Interest Rates: Sweden 1992-1994. Centre for Economic Policy Research, Discussion Paper 1051.

[49] Thornton D. L., 1996, Discount Rate Policies of Five Federal Reserve Chairmen, Working Paper, Federal Reserve Bank of St. Louis (http://research.stlouisfed.org/wp/1996/96-001.pdf).

[50] Wu, T. (2001), "Monetary Policy and the Slope Factor in Empirical Term Structure Estimations," Working Paper 02-07, Federal Reserve Bank of San Francisco.

[51] Wu, T. (2003), "What Makes the Yield Curve Move?" Economic Letter, Federal Reserve Bank of San Francisco, Number 2003-15, June 6, 2003.

VII Annexes

A Independence versus orthogonality

Consider two scalar-valued random variables y_1 and y_2. The variables y_1 and y_2 are said to be independent if information on the value of y_1 does not give any information on the value of y_2, and vice versa. Let us denote by $p(y_1, y_2)$ the joint probability density function (pdf) of y_1 and y_2. Let us further denote by $p_1(y_1)$ the marginal pdf of y_1, i.e. the pdf of y_1 when it is considered alone:

$$p_1(y_1) = \int p(y_1, y_2)\, dy_2$$

and similarly for y_2.Then we define that y_1 and y_2 are independent if and only if the joint pdf is factorizable in the following way:

$$p(y_1, y_2) = p_1(y_1)\, p_2(y_2)$$

The definition can be used to derive a most important property of independent random variables. Given two functions, h_1 and h_2, we always have

$$E[h_1(y_1)\, h_2(y_2)] = E[h_1(y_1)]\, E[h_2(y_2)]$$

A weaker form of independence is uncorrelatedness. Two random variables y_1 and y_2 are said to be uncorrelated, if their covariance is zero:

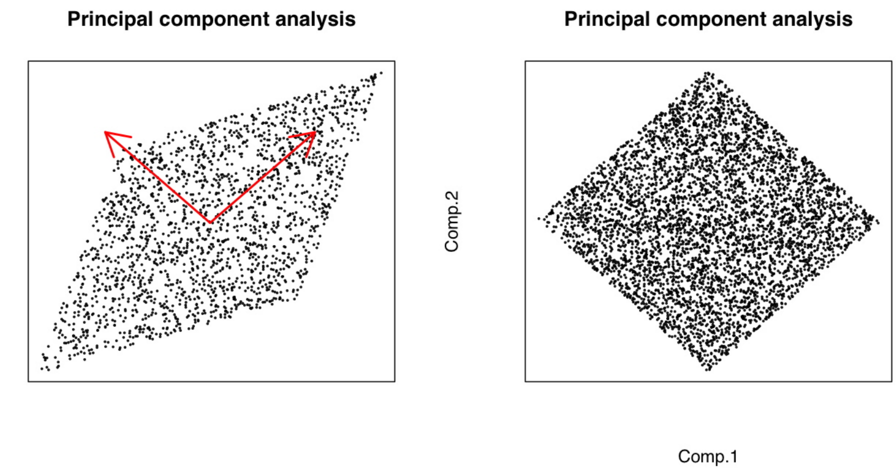
$$E[y_1 y_2] - E[y_1]\, E[y_2] = 0$$

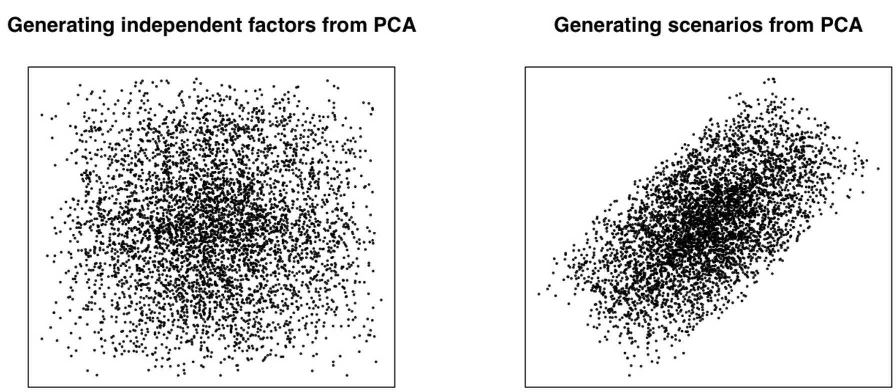
If the variables are independent, they are uncorrelated, which follows directly, taking $h_i(y_i) = y_i$.On the other hand, uncorrelatedness does not imply independence. Assume that (y_1, y_2) are discrete valued and follow such a distribution that the pair are with probability 1/4 equal to any of the following values: $(0,1), (0,-1), (1,0), (-1,0)$. Then y_1 and y_2 are uncorrelated, as can be simply calculated. On the other hand,

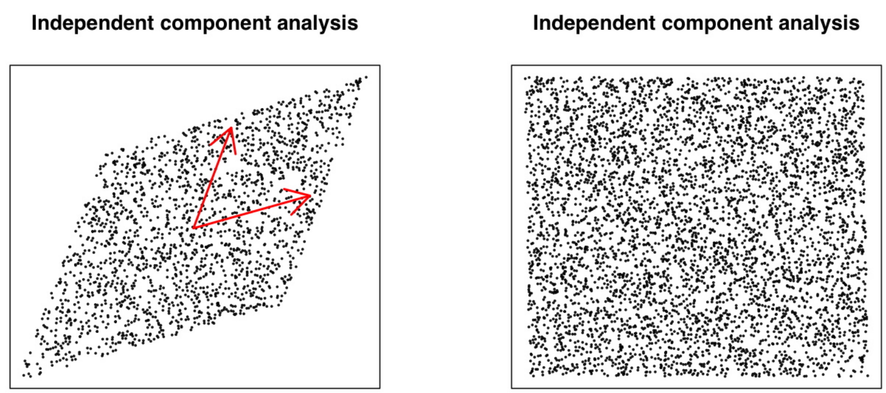
$$E\left[y_1^2 y_2^2\right] = 0 \neq \frac{1}{4} = E\left[y_1^2\right] E\left[y_2^2\right]$$

so the variables cannot be independent. Since independence implies uncorrelatedness, many ICA methods constrain the estimation procedure so that it always gives uncorrelated estimates of the independent components. This reduces the number of free parameters, and simplifies the problem.

Consider the artificial dataset, where the two components are non-independent. The goal of PCA is to detect the strongest uncorrelated factors in a mixture of components (the two arrays on the Figure representing the directions of the factors). If the two factors are uncorrelated (on the right of the Figure), they are obviously non-independent. Next Figure shows random generations of the two factors, where the two factors are assumed to be independent. Observe on the right that assuming independent yield underestimation of tail events: there is undoublty less weight in corners of the polygon. On the other hand, with ICA, we are able to detect independent factors. Using those factors, generated independently, we are able to generate scenarios identical with initial observations.


Principal component analysis
Principal component analysis
Comp.2
Comp.1


Generating independent factors from PCA
Generating scenarios from PCA


Independent component analysis
Independent component analysis

B Figures

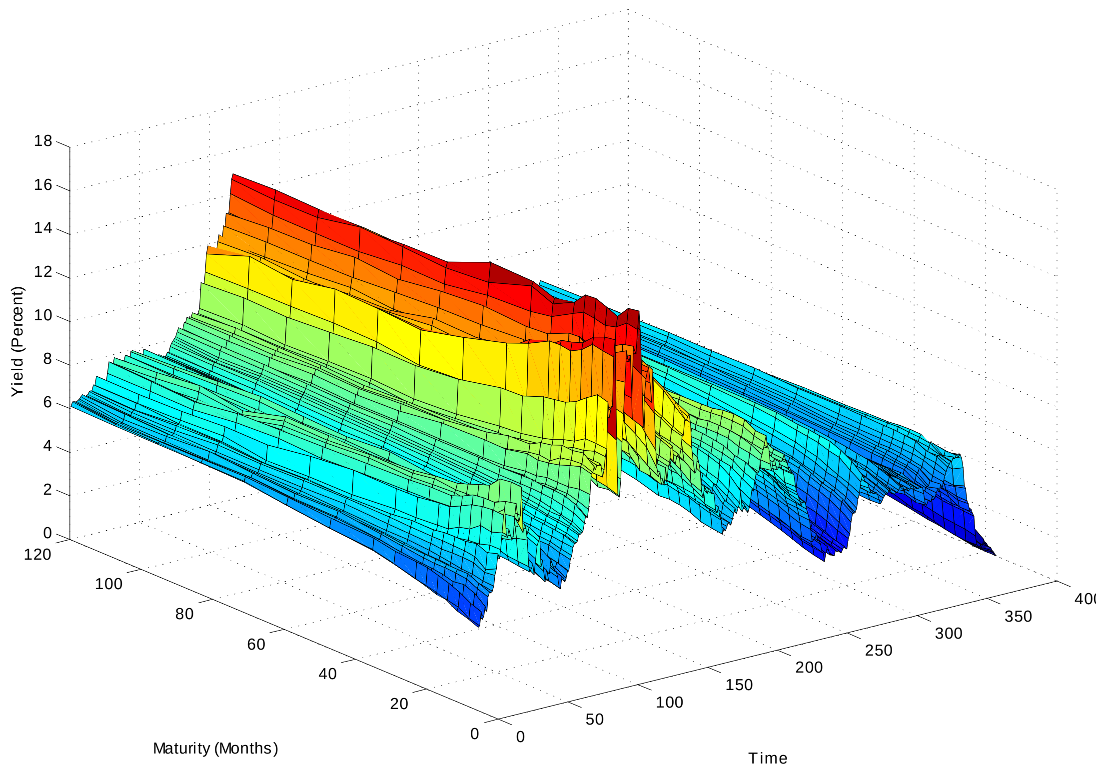


Figure 1: Fig. 1. Yield curves, 1972.01-2002.12. The sample consists of monthly yield data from January 1972 to December 2002. Maturities spanned include 3 months to 12 months at one month increment, then at three months increment to the 5 year maturity, and finally at six months increment to the 10 year maturity, where a month is defined as 30.4375 days.

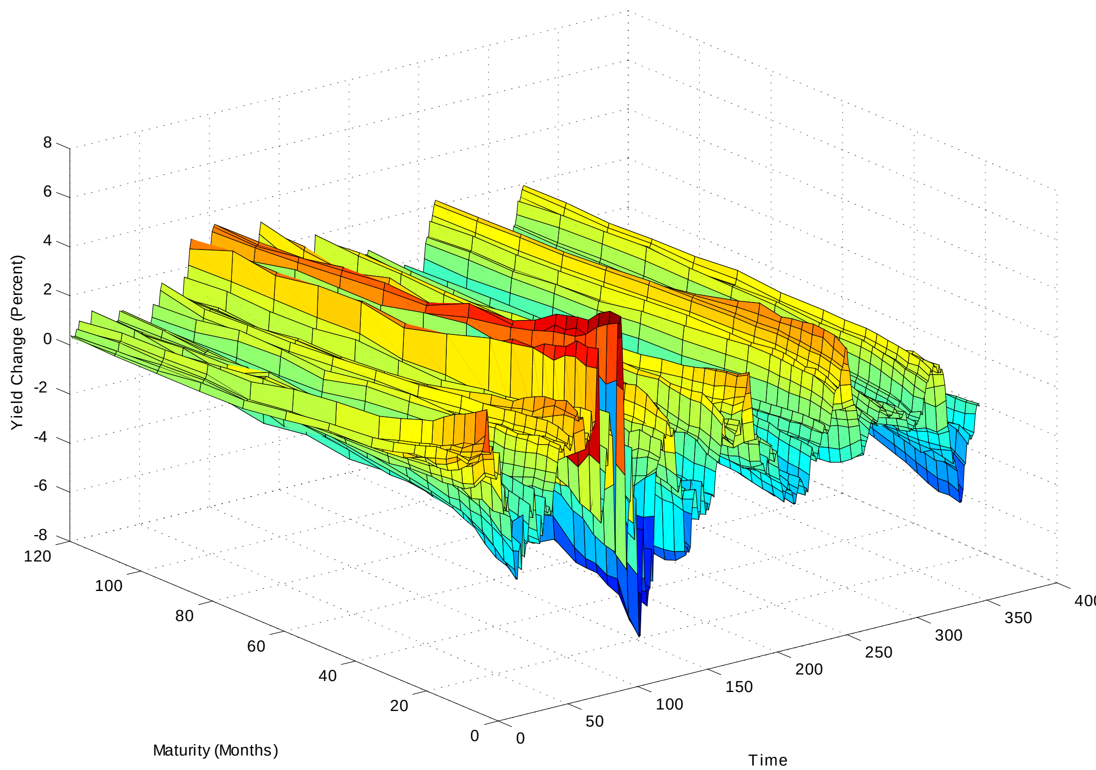


Figure 2: Fig. 2. Yield curve changes with a one year holding period, 1973.01-2002.12. The sample consists of monthly yield data from January 1973 to December 2002. Maturities spanned include 3 months to 12 months at one month increment, then at three months increment to the 5 year maturity, and finally at six months increment to the 10 year maturity, where a month is defined as 30.4375 days.

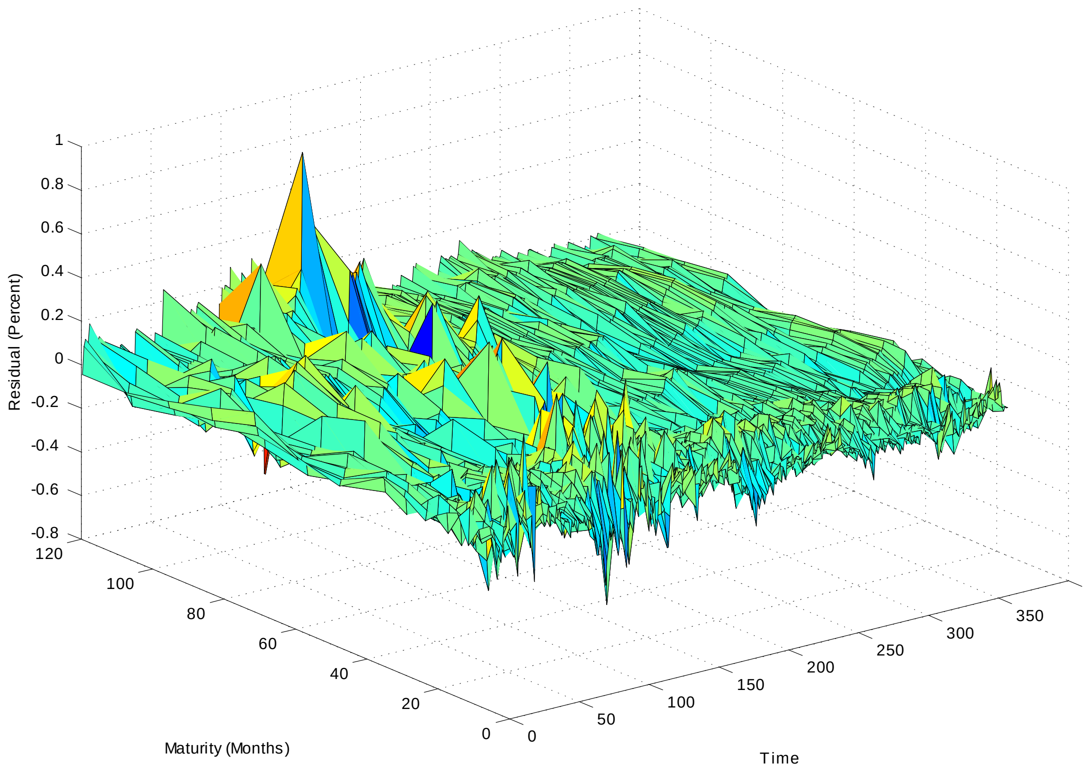


Figure 3: Fig. 3. Yield curve residuals, 1972.01-2002.12. We plot residuals from Nelson and Siegel curves fitted month-by-month with $\lambda = 0.0609$. The sample consists of monthly yield data from January 1972 to December 2002. Maturities spanned include 3 months to 12 months at one month increment, then at three months increment to the 5 year maturity, and finally at six months increment to the 10 year maturity, where a month is defined as 30.4375 days.

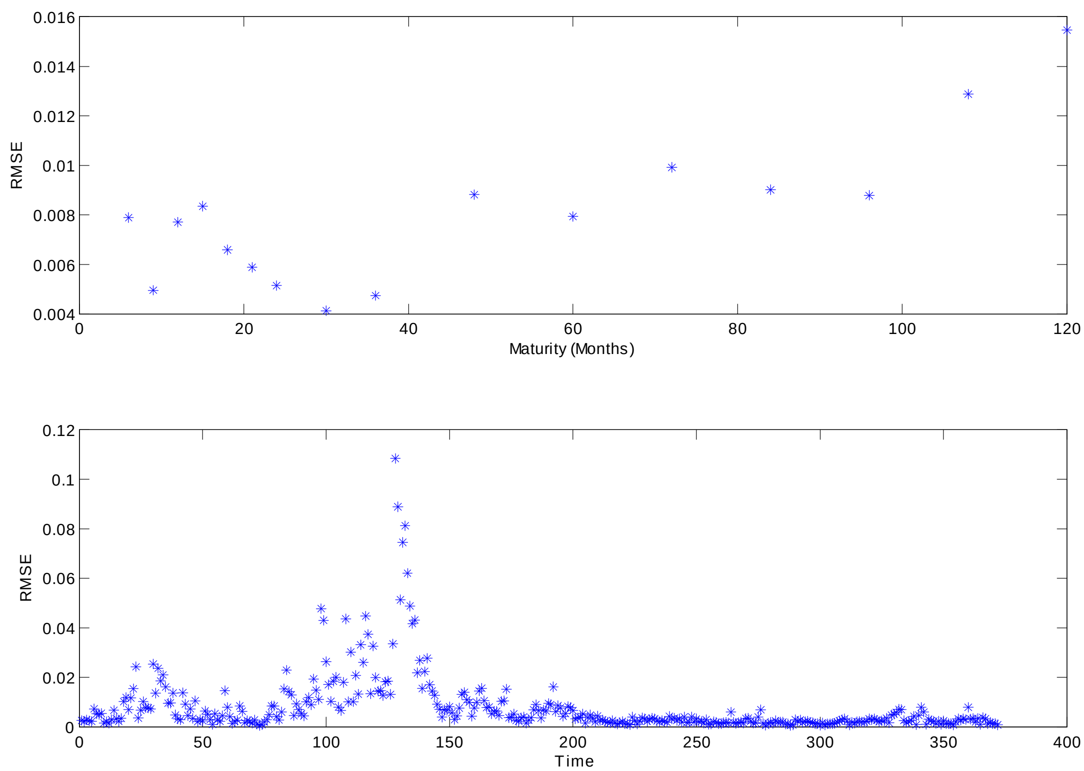


Figure 4: Fig. 4. Root Mean Square Errors, 1972.01-2002.12. We plot Root Mean Square Errors (RMSE) from Nelson and Siegel curves fitted month-by-month with $\lambda = 0.0609$. The sample consists of monthly yield data from January 1972 to December 2002. Maturities spanned include 3 months to 12 months at one month increment, then at three months increment to the 5 year maturity, and finally at six months increment to the 10 year maturity, where a month is defined as 30.4375 days.

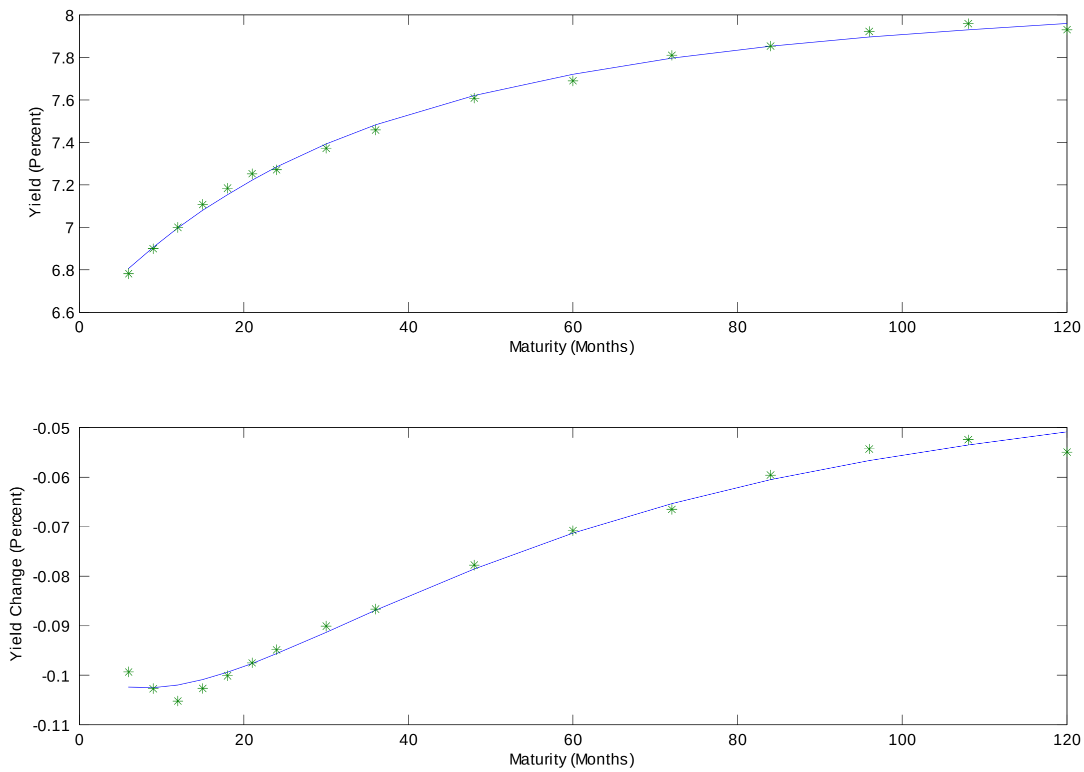


Figure 5: Fig. 5. Actual (data-based) and fitted (model-based) average yield curve and yield change curve (with a one year holding period). We show the actual average yield curve (resp. yield change curve) and the fitted average yield curve (resp. yield change curve) obtained by evaluating the Nelson-Siegel function at the mean values of the three factors (resp. factor changes). The sample consists of monthly yield data from January 1973 to December 2002. Maturities spanned include 3 months to 12 months at one month increment, then at three months increment to the 5 year maturity, and finally at six months increment to the 10 year maturity, where a month is defined as 30.4375 days.

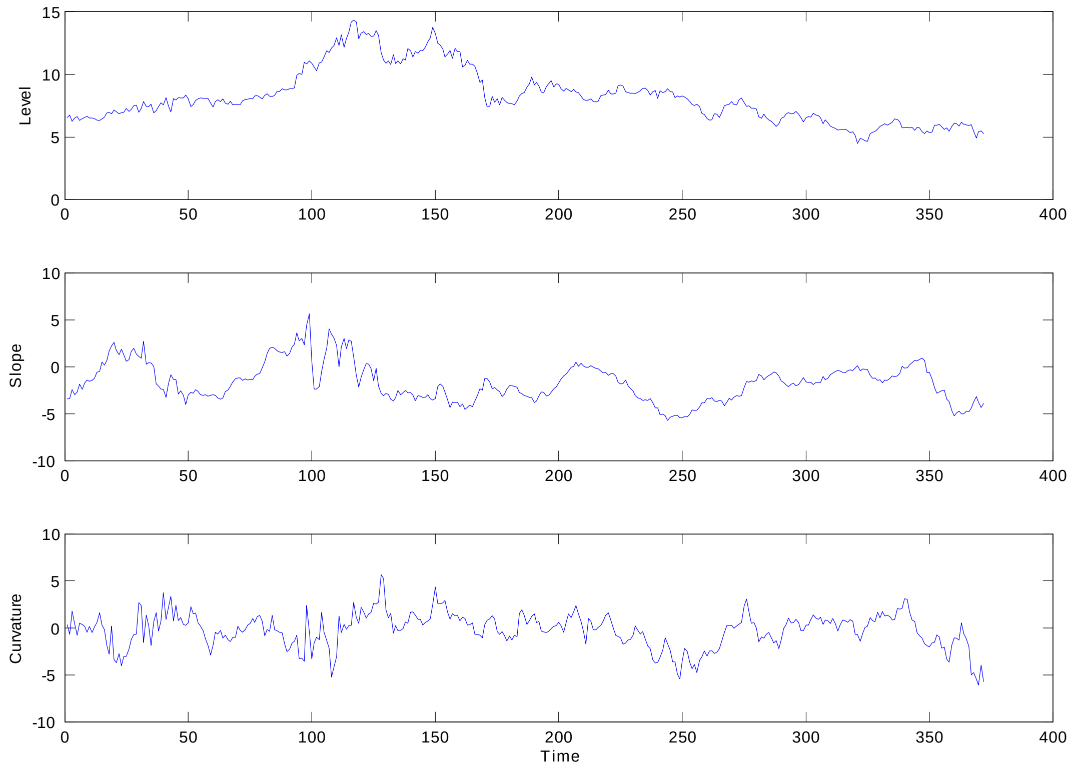


Figure 6: Fig. 6. Model-based level, slope and curvature (i.e. estimated factors) from Nelson and Siegel curves fitted month-by-month with $\lambda = 0.0609$. The sample consists of monthly yield data from January 1972 to December 2002. Maturities spanned include 3 months to 12 months at one month increment, then at three months increment to the 5 year maturity, and finally at six months increment to the 10 year maturity, where a month is defined as 30.4375 days.

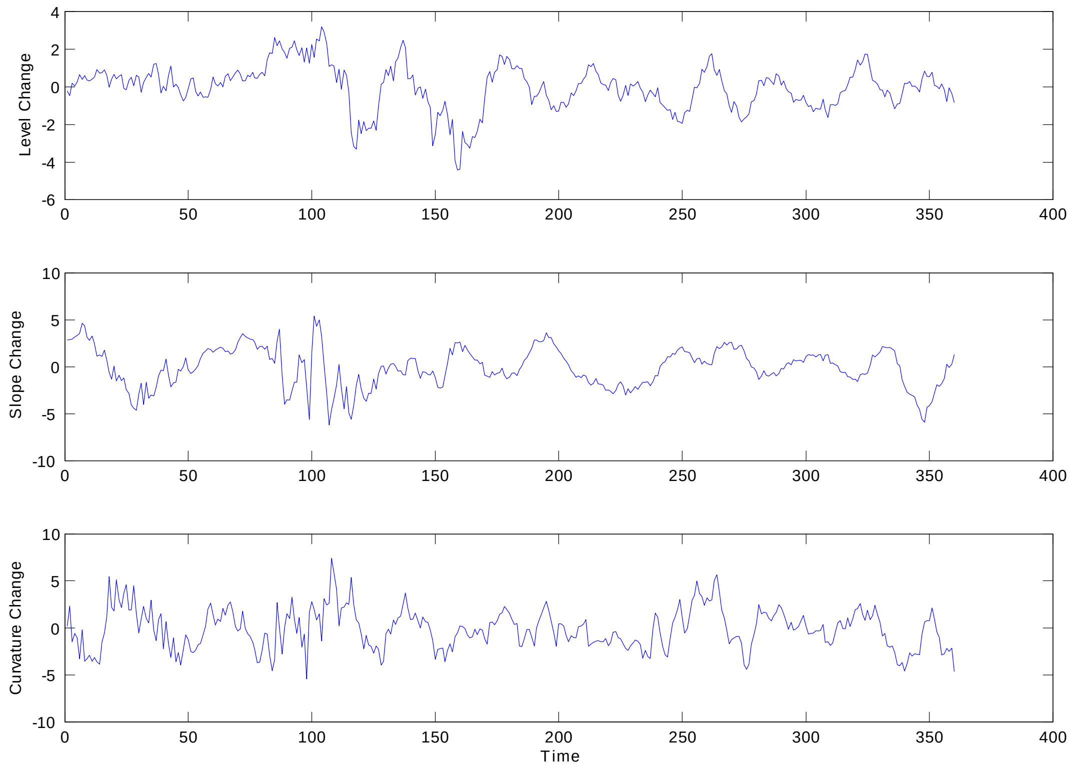


Figure 7: Fig. 7. Model-based level change, slope change and curvature change with an holding period of one year from Nelson and Siegel curves fitted month-by-month with $\lambda = 0.0609$. The sample consists of monthly yield data from January 1972 to December 2002. Maturities spanned include 3 months to 12 months at one month increment, then at three months increment to the 5 year maturity, and finally at six months increment to the 10 year maturity, where a month is defined as 30.4375 days.

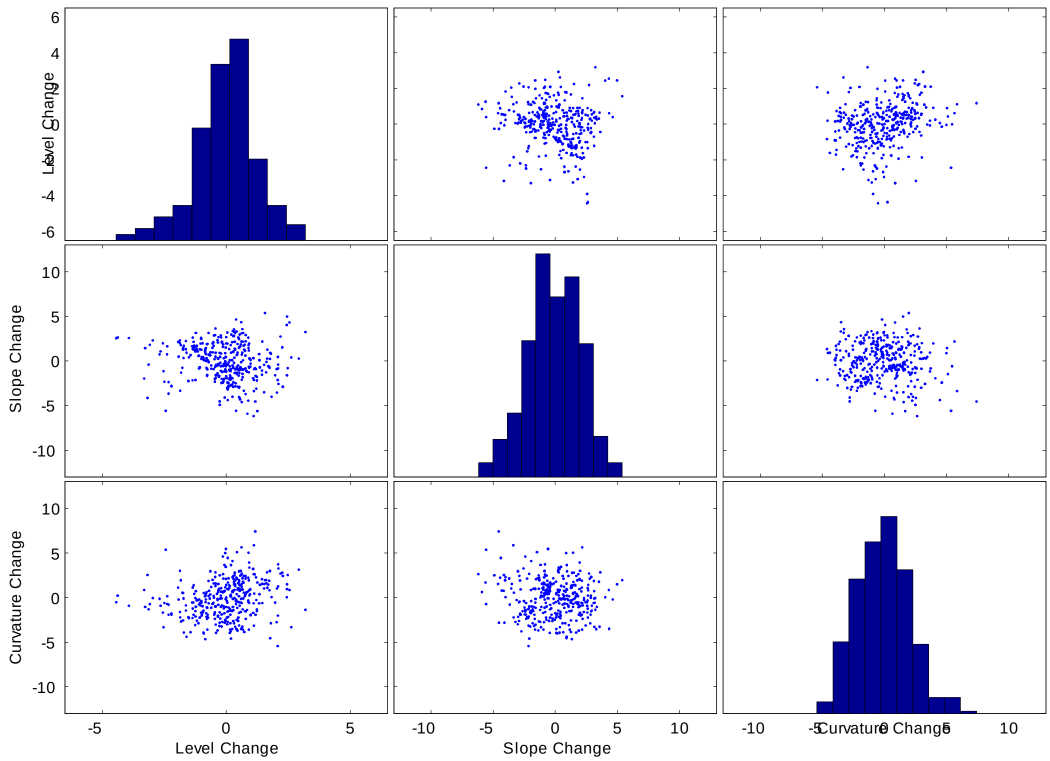


Figure 8: Fig. 8. Scatter plots of the changes of Level, Slope and Curvature Factors with an holding period of one year from Nelson and Siegel curves fitted month-by-month with $\lambda = 0.0609$. In the diagonal we plot an histogram that shows the distribution of each factor change. The sample consists of monthly yield data from January 1972 to December 2002. Maturities spanned include 3 months to 12 months at one month increment, then at three months increment to the 5 year maturity, and finally at six months increment to the 10 year maturity, where a month is defined as 30.4375 days.

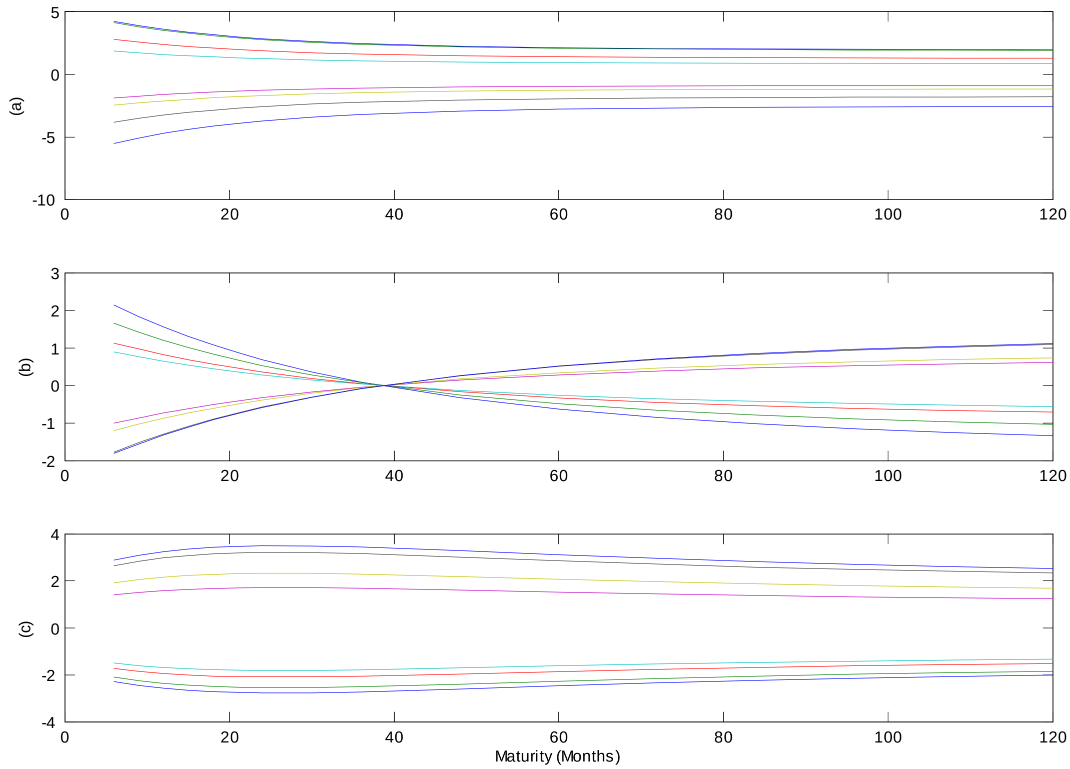


Figure 9: Fig. 9. Independent yield curve "Stress Scenarios". The following quantiles of the resulting distributions are reported: 0.5%, 1%, 5%, 10%, 90% 95%, 99%, and 99.5%. The sample consists of monthly yield data from January 1972 to December 2002. Maturities spanned include 3 months to 12 months at one month increment, then at three months increment to the 5 year maturity, and finally at six months increment to the 10 year maturity, where a month is defined as 30.4375 days.

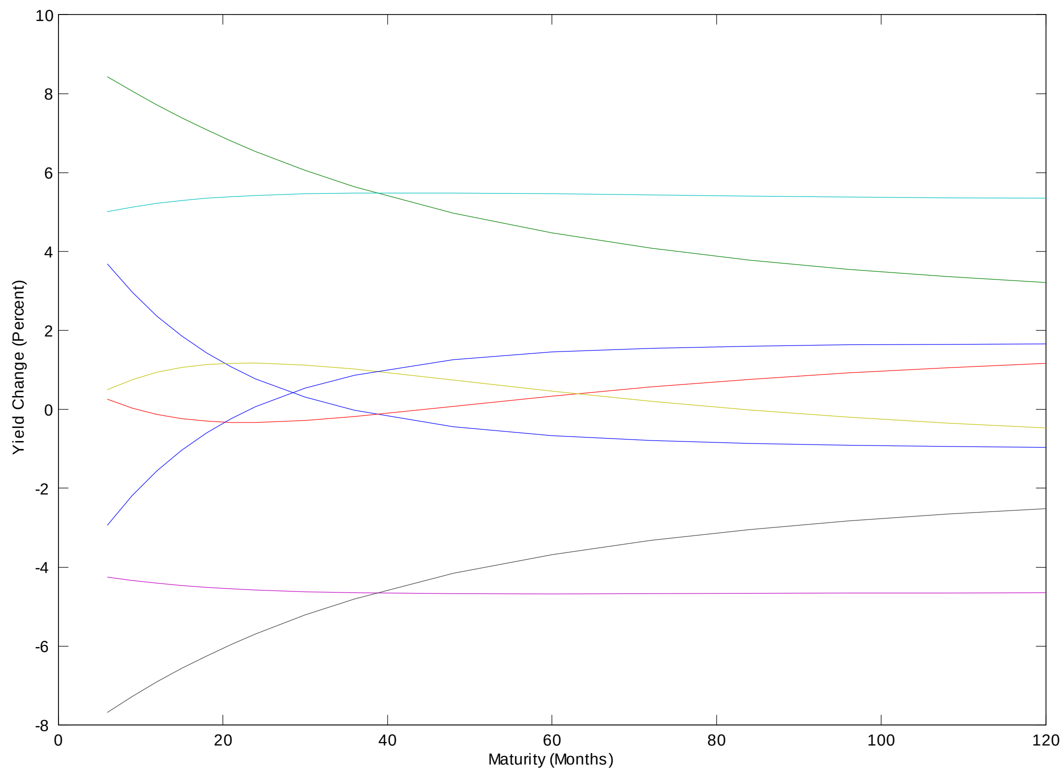


Figure 10: Fig. 10. Simultaneous shock to the three independent components. The following quantiles of the resulting distributions are reported: 1% and 99% which lead to $2^3 = 8$ scenarios. The sample consists of monthly yield data from January 1972 to December 2002. Maturities spanned include 3 months to 12 months at one month increment, then at three months increment to the 5 year maturity, and finally at six months increment to the 10 year maturity, where a month is defined as 30.4375 days.