TESTING THE HOMOGENEOUS INTEREST RATES ASSUMPTION BY PRINCIPAL COMPONENT ANALYSIS: THE EURO AREA CASE

Jerzy Stelmach*
Wroclaw University of Economics, Poland

Abstract. *Although there are many different interest rates in the economy, in theoretical and applied model building these distinctions are usually ignored by assuming that there is only one, "true" interest rate. Hence, the aim of this article is twofold. First, we empirically examine whether such assumption is plausible for the Euro area yield curve data. Second, using different time spans we try to assess the impact of the financial crisis on the validity of this assumption. For both purposes, the principal component analysis technique will be employed.*

Key words: *principal component, interest rates, bond, yield curve, macroeconomics*

Introduction

Interest rates are the fundamental elements of financial and economic activities, and their movements are the major risk factors driving the global capital flows. They are also a vital tool of monetary policy, an important variable in many macroeconomic models and the building block of financial mathematics.

There are many different interest rates in the economy. Despite the distinction between the nominal and the real ones, the interest rates can differ in their term, credit risk and even tax treatment. However, in theoretical and applied model building, these distinctions are usually ignored. This is because the various interest rates tend to move up and down together[1], revealing the significant amount of joint behaviour, which can be abstracted and summarized by the notion of single interest rate. For that reason, it seems theoretically legitimate to assume the existence of only one, "true" interest rate as the dominant and prevailing financial instrument in the economic system. The above-mentioned concept will be defined and referred to as the homogeneous interest rates assumption.

The homogeneous interest rates assumption is well known and widely used as it simplifies the complex behaviour of the interest rates structure. Almost all classical and Keynesian macroeconometric models, derivative-pricing algorithms, portfolio-manage-

[1] As was shown in (Fase, 1973) and (Fase, 1976).

* Wroclaw University of Economics, Komandorska 118/120, 53-345 Wroclaw, Poland
e-mail: jerzy.stelmach@ue.wroc.pl

ment tools, risk-measurement models make some assumptions about the interest rates, the most common being that interest rates are homogeneous. Notwithstanding the importance of this assumption, in most cases it is made only implicitly, without any formal diagnostic or testing procedures. Although a considerable amount of research[2] has been devoted to explain the co-movements among interest rates, none of these studies provide a verification of the homogeneous interest rates assumption, so still little is known about its reasonableness. Therefore, the main goal of this study was to empirically examine whether such assumption is plausible for the Euro area yield curve data. Moreover, using different time spans we tried to assess the impact of the present financial crisis on the validity of this assumption. For both purposes, in the role of a research tool, the principal component analysis was employed.

The remainder of the paper proceeds as follows. The next section describes the principal component methodology and provides its interpretation in the context of this study. The empirical results are presented in the third section. The final section concludes with a brief summary and directions for future research.

Methodology

As indicated in the previous section, our main concern was to identify and test the existence of the joint behaviour of a set of interest rates. To this end, we propose to employ the principal component analysis (PCA). This method is the well-known, standard statistical tool; therefore, in this paper we recall only a few points, focusing mainly on the interpretation, which is relevant to understand the empirical results.

PCA involves a mathematical procedure that linearly transforms a number of possibly correlated variables into a smaller number of uncorrelated variables called principal components. The first principal component accounts for as much of the variability in the data as possible, and each succeeding component accounts for as much of the remaining variability as possible. PCA is based on the eigenvalue-eigenvector decomposition of a covariance or correlation matrix. Its operation can be thought of as revealing the internal structure of the data in a way which best explains the variance in the data.

Let X denote a $T \times n$ matrix containing data on n correlated stationary time series each containing T observations at contemporaneous points in time and let $\mathbf{V}$ be the covariance (or correlation) matrix of X. The principal components of V are the columns of the $T \times n$ matrix P defined by

$$P = XW, \tag{1}$$

where W is the $n \times n$ orthogonal matrix of eigenvectors of V. Thus, the linear transformation defined by W transforms original data X on correlated random variables into a set

[2] See, for example, (Litterman et al., 1991), (Bliss, 1997), (Berk et al., 1999), (Moraux et al., 2002), (Soto, 2004), (Sethi, 2008), (Novosyolov et al., 2008).

of orthogonal random variables, i.e. the principal components. Since for the orthogonal matrix W its inverse is equal to its transpose, $W^{-1} = W^T$, we can turn (1) into a representation of the original variables in terms of principal components:

$$X = PW^T. \tag{2}$$

The mth principal component is the mth column of P, and if we denote the mth eigenvalue of V by $\lambda_{m,}$ the total variation in X is the sum of the eigenvalues of V: $\lambda_1 + ... + \lambda_n$. Hence, the proportion of this total variation, explained by the mth principal component, is

$$\varphi_m = \frac{\lambda_m}{\lambda_1 + \ldots + \lambda_n}, \tag{3}$$

and is used as an analogy for R^2 in the regression analysis, indicating the goodness of fit.

The major aim of PCA is to use only a reduced set of principal components to represent the variables X. For this purpose, the eigenvalues and their corresponding eigenvectors in W are ordered from largest to smallest, i.e. $\lambda_1 \geq ... \geq \lambda_n$. In a highly correlated system, the first eigenvalue will be much larger that the others, so the first principal component alone can explain a very large part of movements in the data, followed by the second component, and so on. Hence, most of variations in the system can be explained by using only the first few principal components, as these are the most important ones. This fact can be represented by adjusting equation (2) to an approximation of the original data in terms of the first k principal components only:

$$X \approx P^*W^{*T}, \tag{4}$$

where P^* is $T \times k$ a matrix with columns being the first k principal components, and W^* is an $n \times k$ matrix with k columns given by the first k eigenvectors. This equation lies at the core of all PCA models and allows extracting just the key sources of variation from the data.

In this article, the representation (4) also forms the basic motivation and justification of using PCA to test the homogeneous interest rates assumption. The logic adopted here is the following: if we are able to attribute a fairly large part of variation in the interest rates to the first principal component, we can consider the discussed assumption as plausible. Otherwise, the homogeneous interest rates assumption will be recognized as not supported by the available data over the given period.

The first principal component usually is interpreted as a common trend, and in the context of interest rates it can be seen as a parallel movement during which all rates change up or down by roughly the same amount (Jajuga, 2006). For this reason, the first component is often called the trend or level component (Soto, 2004). The interpretation of the second and higher order principal components does depend on having the natural ordering in the system. If the system is ordered, then the second principal component captures a change in the slope of the term structure (Alexander, 2008). If the elements of the second principal component are decreasing (or increasing) in magnitude and this component changes, then the interest rates move up at one end of the term structure and down at the other end. The third principal component is often called the curvature or convexity component, because if it changes, the interest rates move up (or down) at both ends of the term structure and down (or up) in the middle.

As in the case of any statistical method, PCA has also some limitations, which should be carefully recognized. The major limitations of PCA are that it provides a linear transformation only, and utilizes merely the second-order statistics (covariance) of the data. On the other hand, these limitations greatly facilitate the computation of PCA compared to nonlinear techniques such as neural networks. If the underlying data are Gaussian, nonlinear processing and the use of higher-order statistics do not yield any advantage over PCA type solutions.

Data and empirical results

The empirical analysis in this paper focuses on the euro area government debt market. The dataset contains historical zero-coupon spot yield curves for all euro area central government bonds with all issuers and all ratings included. The estimation of the yield curve is done by means of a modelling algorithm, namely the Svensson model[3], which minimises the sum of the quadratic difference between the yields which can be computed from the curve and the yields actually measured. The sample period extends from January 2007 to April 2010, providing in total 850 daily observations after excluding non-trading days. We will utilize yield curves for maturities between 3 months and 30 years with a three-month step size[4], thus the dimension of our input data matrix is 850 × 120. The series for the analysis are obtained from the European Central Bank webpage[5].

The sample constitutes a very valuable period, since the European economy was subject to the financial crisis. Figure 1 below shows the yield curves for all considered maturities. Three different regimes in interest rates are apparent from this graph. From January 2007 until mid-2007, the yield curves for different maturities were quite close to each other. In the second half of 2007 until September 2008, the span between the curves starts to widen. This process was intensified greatly following Lehman Brother bankruptcy on 15 September 2008, which triggered the world-wide economic downturn.

In finance, the profits and losses on fixed income portfolios are mapped to changes in interest rate risk factors measured in basis points (Alexander, 2008). Hence, the volatili-

[3] See (Svensson, 1994).

[4] In other words, we will use yield curves for the following maturities: 3M, 6M, 9M, 1Y, …, 29Y9M, 30Y.

[5] http://www.ecb.int/stats/money/yc/html/index.en.html.

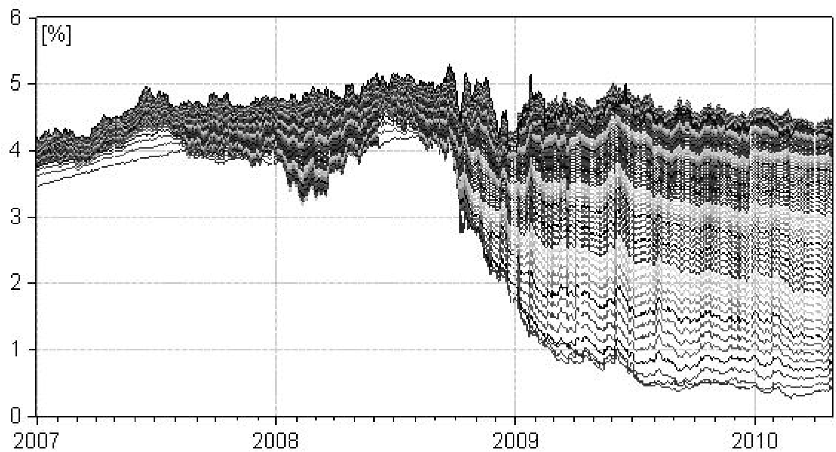

FIG. 1. **Euro area government zero-coupon yields, 01.2007-04.2010.**

ties and correlations of interest rates should refer to the absolute changes in interest rates in basis points. Figure 2 shows the volatility of the examined zero-coupon spot rates in the basis point per annum, plotted against the maturity of the spot rate. Volatility is lowest at the very short end (up to 1.5 years) and highest for rates longer than 28 years. Rates between 5 and 25 years' maturity have a rather stable volatility within the range of 60 bps as the lower and 70 bps as the upper boundary. Since the volatility of the majority of the rates is quite steady and does not fluctuate significantly over the sample, the results of applying the principal component analysis to the covariance matrix, which includes the volatilities of the rates, are expected to be similar to the results of applying PCA to the correlation matrix. Therefore, the subsequent analysis will utilize only the correlation matrix.

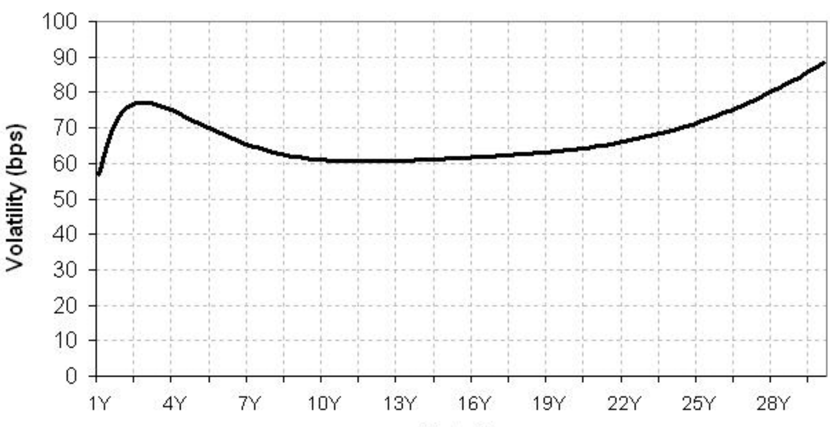

FIG. 2. **Volatilities of government zero-coupon yields, 01.2007–04.2010.**

We perform a PCA on the correlation matrix of the daily changes in each rate. The correlation matrix is a 120 × 120 matrix measured using the equally weighted average methodology. An extract from this matrix is shown in Table 1. The correlation matrix exhibits the usual behaviour for correlations in a term structure: correlations are higher for adjacent maturities and decrease as the maturity difference between the rates increases. Moreover, in this case the 1-year rate has the lowest correlation with the rest of the system overall, because this is the money market rate that is more influenced by government policies than the longer rates.

TABLE 1. **Correlation matrix of selected changes in spot rates**

Maturity	1Y	2Y	3Y	4Y	5Y	10Y	15Y	20Y	25Y	30Y
1Y	1.000	0.896	0.817	0.759	0.712	0.519	0.386	0.326	0.288	0.248
2Y	0.896	1.000	0.973	0.924	0.871	0.624	0.464	0.390	0.335	0.274
3Y	0.817	0.973	1.000	0.985	0.950	0.704	0.520	0.439	0.384	0.323
4Y	0.759	0.924	0.985	1.000	0.989	0.779	0.584	0.496	0.439	0.376
5Y	0.712	0.871	0.950	0.989	1.000	0.846	0.653	0.557	0.492	0.418
10Y	0.519	0.624	0.704	0.779	0.846	1.000	0.934	0.830	0.684	0.514
15Y	0.386	0.464	0.520	0.584	0.653	0.934	1.000	0.956	0.808	0.601
20Y	0.326	0.390	0.439	0.496	0.557	0.830	0.956	1.000	0.937	0.779
25Y	0.288	0.335	0.384	0.439	0.492	0.684	0.808	0.937	1.000	0.947
30Y	0.248	0.274	0.323	0.376	0.418	0.514	0.601	0.779	0.947	1.000

Given the distinct regimes in euro area interest rates, a PCA over the whole period only will not reflect the prevailing market circumstances. Therefore, below we conduct PCA over the whole sample and also over the three sub-periods identified above. Table 2 reports the main outcomes of the computations: the three largest eigenvalues λ_i along with the percentage of variation explained φ_i.

As one can see from Table 2, the eigenvalues are all positive because the correlation matrix is positive definite, and they have been sorted in a decreasing order of magnitude. Since we are dealing with the 120 × 120 correlation matrix, the sum of all eigenvalues should be equal to 120. The eigenvalues of the correlation matrix determine how much of the covariance in the system is explained by each principal component. For instance, over the whole sample period denoted by (A), the first eigenvalue is 90.19, which means that the first principal component explains 90.19/120 = 75.16% of all the variation. Taken together, the first three principal components account for 96.31% of the movements in the yield curve over the sample period (A). Moreover, if the first principal component shifts upwards, leaving the other principal components fixed, then all the rates will move upwards in an approximately parallel shift, and this type of shift explains 90.19% of

TABLE 2. **Eigenvalues and explanatory power of the first three principal components**

Period		Principal components			
		First $i = 1$	Second $i = 2$	Third $i = 3$	Sum
(A): January 2007 – April 2010	λ_i	90.19	18.11	7.26	115.56
	φ_i	75.16%	15.10%	6.05%	96.31%
(B): January 2007 – July 2007	λ_i	109.70	7.21	1.54	118.45
	φ_i	91.42%	6.01%	1.28%	98.71%
(C): August 2007 – 14 September 2008	λ_i	95.77	17.49	3.71	116.97
	φ_i	79.81%	14.57%	3.09%	97.47%
(D): 15 September 2008 – April 2010	λ_i	87.08	18.59	8.96	114.63
	φ_i	72.57%	15.50%	7.47%	95.54%

movements in the data. Similarly, the second or tilt component will produce 15.10% of the changes, and the third – curvature principal component – will affect 6.05% of the variation.

The following four figures (Figs. 3, 4, 5 and 6) show the first three eigenvectors plotted as a function of maturity of interest rate for each sub-period described in Table 2. In order to facilitate the visual comparison, the same vertical scaling and line styles were applied to all graphs. These figures indicate that the first three principal components have the standard stylized interpretation of level, tilt and curvature.

The first, and most important, eigenvector is practically a horizontal line because it has almost identical values on each maturity, and this feature is most apparent for the pre-crisis sub-period (B). As stated above, the short rates have a lower correlation with the system than the other rates. Hence, at the short maturities, the first eigenvector is not as flat as it is for the longer maturities.

The second eigenvector is very similar to a downward sloping line moving from positive to negative. Therefore, an upward shift in the second component, leaving the other components fixed, induces a tilt in the yield curve, with an upward move at the short end and a downward move at the long end. This type of movement accounts for at minimum 6% for sub-period (B) and maximum 15.5% for sub-period (D), meaning that in turbulent times we can encounter less parallel and at the same time less homogeneous movements of interest rates.

The third eigenvector, denoted on the graphs by the dashed line, has a smooth shape comparable to a quadratic function of maturity, being positive at the short and the long ends and negative for middle maturities. Hence, an upward shift in the third principal component (leaving the other components fixed) will change the convexity of the yield

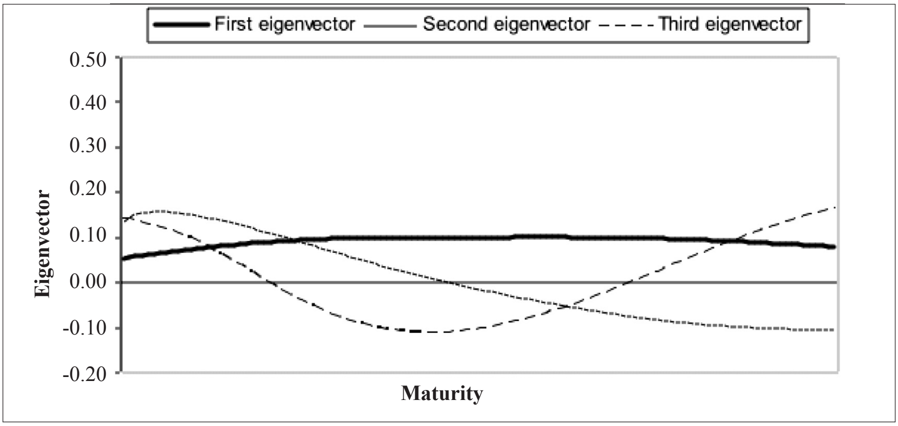

FIG 3. **Eigenvectors of the zero-coupon yields correlation matrix: sub-period (A)**

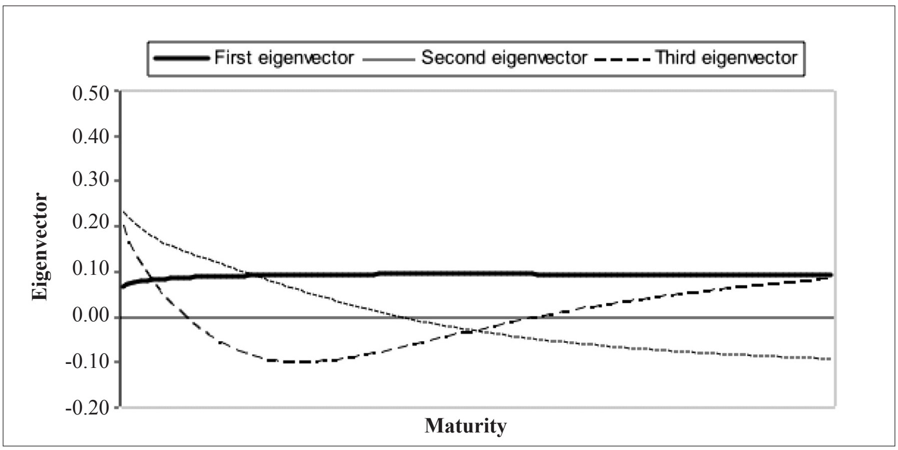

FIG. 4. **Eigenvectors of the zero-coupon yields correlation matrix: sub-period (B)**

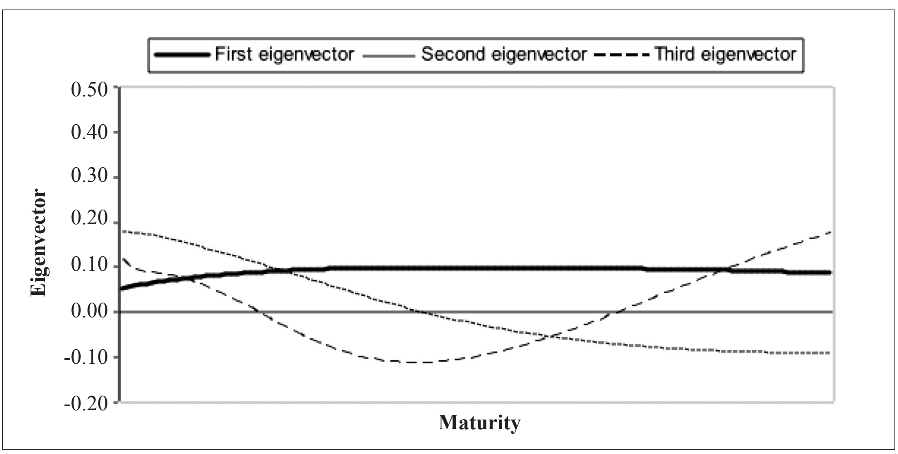

FIG. 5. **Eigenvectors of the zero-coupon yields correlation matrix: sub-period (C)**

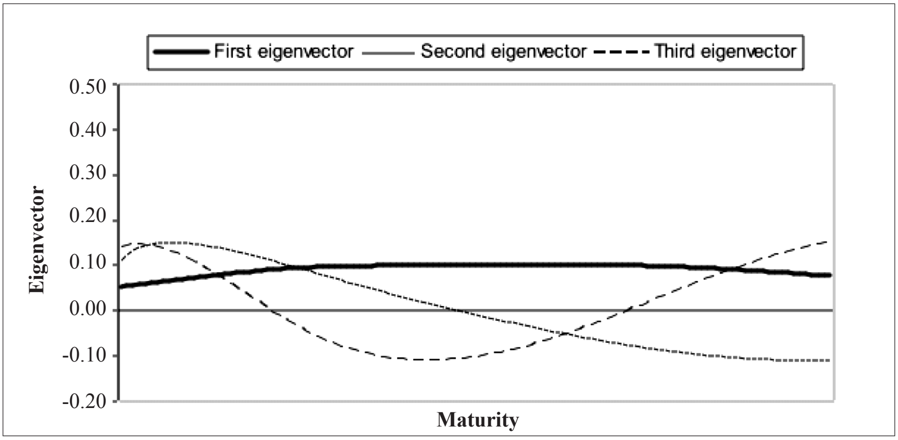

FIG. 6. **Eigenvectors of the zero-coupon yields correlation matrix: sub-period (D)**

curve. It will make a downward sloping curve more convex and an upward sloping curve less convex. This component accounts for only 1.28% of the variation in sub-period (B) for and 747% in sub-period (D).

From the perspective of testing the homogeneous interest rates assumption, the first principal component can be identified as the “true”, prevailing interest rate. In this sense, the more variation is captured by the first component, the more reasonable the homogeneous interest rates assumption becomes. Following such reasoning, we can empirically assess the validity of this assumption for different time periods indicated in Table 2. The conclusion to be drawn is that we can find a relatively strong justification for the homogeneous interest rates assumption only for fairly tranquil periods, such as period (B) before the beginning of the financial meltdown, during which the first principal component alone explains more than 91% of the variation. The results for the next two periods, namely (C) and (D), form the clear and expected pattern with a lower explanatory power φ_1 which is equal to 79.81% and 72.57% respectively. Since the difference between φ_1 for period (B) and period (D) is almost 19 percentage points, we can conclude that the financial crisis of 2008 had a pronounced and negative impact on the legitimacy of the homogeneous interest rates assumption. Thus, in general, we can state that during the periods of economic turmoil the investigated assumption is questionable and hardly acceptable because the rationale for it is limited.

Conclusions

In this paper, we consider the euro area historical yield curves in an attempt to evaluate the performance and validity of the homogeneous interest rates assumption. The adopted verification method is empirical and relies on the principal component analysis, which is applied to different time periods.

The conclusions regarding testing the homogeneous interest rates assumption are twofold. First, the “true”, prevailing interest rate can be identified as the first principal component which can be interpreted as a single indicator of the joint behaviour of the interest rates’ term structure. Second, while in the relatively stable economic conditions the first principal component reflects more than 91% of the co-movements in interest rates, during the period of turbulence this number is significantly smaller. As an immediate implication, it means that all the theoretical and applied models that incorporate the homogeneous interest rates assumption should be subjected to a critical review in this respect. Looking at it from a different perspective, we can also say that to have homogeneous interest rates in the whole euro area is fundamental for an effective functioning of the single monetary policy in the Eurosystem.

Taking into consideration the importance of the homogeneous interest rates assumption, further research in this area is needed. Future contributions and directions of studies can include the derivation of a formal statistical hypothesis test or extending the dataset by adding the other types of interest rates for different currencies.

REFERENCES

Alexander, C. (2008). Market Risk Analysis: Practical Financial Econometrics. Wiley & Sons.

Berk, J. M., Knot, K. (1999). Co-movements in long-term interest rates and the role of PPP-based exchange rate expectations. IMF Working Paper, No. 99/81, pp. 1–28.

Bliss, R. R. (1997). Movements in the term structure of interest rates. Economic Review, Federal Reserve Bank of Atlanta, Vol. 4, pp. 16–33.

Fase, M. M. G. (1973). A principal components analysis of market interest rates in the Netherlands, 1962–1970. European Economic Review, Vol. 4, pp. 107–134.

Fase, M. M. G. (1976). The interdependence of short-term interest rates in the major financial centres of the world: Some evidence for 1961–1972. Kyklos, Vol. 29, pp. 63–96.

Jajuga, K. (2006). Interest rate modeling and tools of inancial econometrics. Dynamic Econometric Models, Vol. 7, Nicolaus Copernicus University, pp. 7–13.

Litterman, R., Scheinkman, J. (1991). Common factors affecting bond returns. Journal of Fixed Income, Vol. 1, pp. 54–61.

Moraux, F., Perignon, C., Villa, C. (2002). Common factors in international bond returns revisited: a common principal component approach. Swiss National Science Foundation Working Paper.

Novosyolov, A., Satchkov, D. (2008). Global term structure modelling using principal component analysis. Journal of Asset Management, Vol. 9, pp. 49–60.

Sethi, R. (2008). New Zealand zero-coupon yield curves: a principal-components analysis. Bank for International Settlements. Available from: http://www.bis.org/repofficepubl/arpresearch200805.1.htm [Accessed on 02.05.2010].

Soto, G. M. (2004). Using principal component analysis to explain term structure movements: performance and stability. Progress in Economics Research, Vol. 8, Nova Science Publishers, pp. 203–225.

Svensson, L. E. (1994). Estimating and interpreting forward interest rates: Sweden 1992–1994. Centre for Economic Policy Research, Discussion Paper No 1051.