ESTIMATING TERM STRUCTURE CHANGES USING PRINCIPAL COMPONENT ANALYSIS IN INDIAN SOVEREIGN BOND MARKET

Golaka C Nath[1], Manoj Dalvi[2]

[1]*Corresponding Author, Clearing Corporation of India, India. Email-id: gcnath@hotmail.com*
[2]*Associate Professor, Long Island University, New York.*

Abstract *This paper analyses the Indian sovereign yield curve shift from January 1997 to May 2012 to find out the principal factors affecting the term structure of interest rate changes. We apply Principal Component Analysis (PCA) on our data consisting of zero coupon interest rates derived from government bond trading using Nelson-Siegel functional form. This decomposition of the yield curve highlights important relationship between identified factors and metrics of the term structure shape. The empirical findings support statistical similarities between the Indian yield curve and term structure studies of major countries.*

Keywords: *Indian Sovereign Yield Curve, Principal Component, Interest Rates, Bond, Yield Curve, Macro-economics, and Term Structure of Interest Rates.*

1. INTRODUCTION

Yield curve estimation in emerging markets like India has been a challenging job. The sovereign bond market is characterized by illiquidity in terms of number of bonds traded vis-à-vis number of outstanding bonds, value traded as a proportion to outstanding bond issuances, activity concentrated on few bonds and the benchmark 10-year bond typically account for a large share in trading activity, low trading activity in major part of the yield curve. Hence estimation of the sovereign yield curves have to be done using sophisticated methods. Entities like National Stock Exchange (NSE) and Clearing Corporation of India Ltd (CCIL) have been doing a fair job by estimating the term structure on daily basis and releasing the same to the market. Before CCIL came to the arena and specifically after the introduction of anonymous order matching system in Gilts market in August 2005, NSE captured the trading data of the sovereign bond market through their Wholesale Debt Market (WDM) platform as major part of the deals in the market used to be transacted through brokers. It could not capture all deals in the market as some of the deals were transacted directly among market participants and settlement of all trades happened at Reserve Bank of India (RBI). It helped NSE to estimate the term structure on daily basis using Nelson-Siegel functional form and many banks and institutions adopted the valuation techniques using the information of the estimated term structure. The role of CCIL became very important after Reserve Bank of India mandated settlement of all Government securities deals through CCIL. Since all trades, including the brokered trades, have to be reported to a centralized system at RBI for final settlement through CCIL, it became the repository of all trades in Gilts in India. Unlike NSE, this helped CCIL in capturing the full market data in Gilts and since it has to provide guarantee of settlement, it estimated the term structure of interest rates on daily basis using Nelson-Siegel functional form.

Indian sovereign bond market is generally illiquid when we compare it to the developed markets. However, the well-functioning market microstructure helped it to have a great deal of market efficiency in pricing instruments traded in the market. The well-structured primary issuances market for Government bonds through Issuance calendars, availability of bonds in all maturities up to 30 years, higher level of outstanding issuances in many bonds of different maturities, passive consolidation of issuances through reopening issues and creating liquidity, a well-functioning primary dealers network, a central counterparty (CCP) based settlement system, availability of quality information to market participants on each and every bond through CCIL, creating an anonymous order drive system for sovereign bonds, a well-functioning money market for short term market using three different variants like Inter-bank call, inter-bank Repo, and a quasi-repo CBLO (Collateralized Borrowing and Lending Obligations), a well-designed Liquidity Adjustment Facility (LAF) of RBI to support the market to moderate the money supply using daily fixed rate Repo and Reverse Repo, etc. has helped the market in terms of price efficiency.

A reasonable estimation of the sovereign yield curve in an economy is important for several reasons, both at the macroeconomic level and at the level of private financial entities. The yield curve serves as a benchmark in the economy as private corporate entities raise funds by paying a credit spread for the risk inherent in them; investors use the sovereign yield curves to demand an appropriate price for their investment risk; banks and other financial institutions use the yield curves to not only price the illiquid securities in their books but also match the duration of their assets and liabilities; central banks use the information from secondary market yield curves to monitor the policy interest rate synchronization with the "economic effective rate" in the inter-bank market; at macroeconomic level, the yield curve has a predictive power for the state of economy. The yield curve modeling has become an important area for all financial markets. During the last few years, we could observe high volatility of interest rates. The yield of corporate and government bonds have increased significantly during the financial crisis. Due to current debt crisis on the periphery of European monetary union, bond yields remain at high level. In India, yields remain high for a long period as the inflation has remained high for good many months and liquidity shortage in the inter-bank market has been continuing unabated since July 2010.

Term structure estimation using models like Nelson-Siegel (NS) functional form has been in operation in India since 1999. The parameters estimated by this model helps us to calculate the spot interest rate (zero rate) for any term using the NS equation. The risk management practices like Value-at-Risk (VaR) heavily depend on the historical price behavior to estimate the possible future risk for having the sufficient amount of capital to cover market risk in those investments. It is paramount to simulate the historical price of the securities using the historical yield curves. The market observed price of the bonds cannot be used to compute VaR as a bond changes its structure every day (maturity comes down by 1 day on daily basis and hence a 10-year bond today was a 11-year bond one year back and hence its observed trading prices were on the basis of time to maturity and other factors). The purpose of this study is to understand the dynamics of the term structure of interest rate in India using Principal Component Analysis (PCA). The main purpose of this paper is to study the term structure dynamics and to figure out the common factors of the Indian term structure and its volatility as it helps to understand the pricing mechanism of various OTC and other underlying and derivative products. Corporate entities price their issuances on the basis of sovereign yield by adding a credit spread. Previous literature has focused on the term structure of interest rates (Litterman & Scheinkman, 1991; Dai & Singleton, 2000). These studies have concluded that a few common factors explain observed variation in historical bond prices. These three common factors in the term structure of interest rates are interpreted as level, slope, and curvature factors based on the factor loadings from principal components analysis (Díaz et al., 2010). This principal component analysis is a common method to analyse the bond valuation ability of alternative models on the first moment of interest rates (Litterman & Scheinkman, 1991; Piazzesi, 2005; Matzner-Løber & Villa, 2004; Pérignon et al., 2007; Cornillon et al., 2008; Olawale & Garwe, 2010; and Huang & Chen, 2011). Chandra (2008) studied Indian yield curve movements using PCA in order to identify factors which are responsible for changes in the yield curve.

The paper is divided into different sections: Section 2 provides the dynamics of historical term structure of interest rates; Section 3 provides the volatility of the term structure; Section 4 gives the use of PCA in studying dynamics of yield curve; Section 5 estimates the dynamics of term structure using PCA; Section 6 gives the conclusion and findings of the study.

2. HISTORICAL TERM STRUCTURE OF INTEREST RATE IN INDIA

Study of yield curve behavior has been an important part of financial market research as it provides us important information about the future expectation of growth, inflation, recession, etc. The slope change of the yield curves provides good information about the future of the economy (Estrella & Mishkin, 1996; Bernanke & Blinder, 1992; Mishkin, 1990). Indian sovereign bond market has seen many structural changes during last two decades or so. Many significant microstructure changes were introduced during last few years to strengthen Indian sovereign bond market. The issuance of sovereign bonds has become increasingly systematic with passive consolidation. Very few issues were new issues and RBI concentrated in reopening the issues to add liquidity as outstanding stocks increased due to re-issuances. The borrowings of the Government have considerably increased over time to fund a growing economy and reached INR30.5 trillion as of March'12 (Table 1).

The primary issuances of Government securities are managed by RBI as per a statute. The RBI also works as the central depository and record keeper of the Government debt. For historical reasons, the Government securities market was a typical Over the Counter (OTC) market where banks and financial institutions traded among themselves and settled at central bank money. A large financial market scam in 1992 involving Government securities, brokers, and Banks resulted in making the securities holding records into electronic book entry form from the physical form. The clear differentiation between constituent and proprietary positions and holding helped creating audit trail which helped the market in many ways in terms of transparency. The WDM segment of NSE started in June'94 and it revolutionized the transparency system in Government securities market.

Table 1: Government Securities Issuance

Year	Change over Previous Year (%)	Debt (INR Trillion)	Average Coupon (%)	Average Maturity (years)	Turnover Ratio
2006-07	19.05	12.97	8.66	10.10	78.76
2007-08	21.06	15.70	9.57	8.42	105.33
2008-09	18.22	18.56	8.22	9.91	116.37
2009-10	16.95	21.71	7.98	9.79	134.22
2010-11	14.76	24.91	7.84	9.76	115.24
2011-12	22.43	30.50	7.87	9.69	114.37

Note: Borrowing included dated securities, floating rate bonds, T-bills issued by Govt. of India and Turnover Ratio has been calculated as the ratio of 12 months total trading value and total outstanding debt.

NSE made it mandatory for brokers to report the deals to its electronic platform as most of the deals in Gilts were broker driven. Once the deals were reported to the platform, NSE initiated the dissemination of the same to the market on real time basis as well as the end of the day. It provided valuable information to the market in terms of clean data. NSE started using the deals to estimate yield curves and made the Zero coupon yield curves public from 1997.

The RBI introduced an electronic reporting system in Feb'02 making it mandatory for market participants (as most buyers and sellers are Banks and financial institutions) to report the deals within a limited time to its reporting system called Negotiated Dealing System (NDS). Once the deals were reported to the system, it could be consolidated for settlement using a Delivery versus Payment – II mechanism through CCIL which worked as a clearing house and a CCP. As a part of reforming financial market structure in India, RBI made it mandatory for all Banks and financial institutions to settle their deals in Government securities (outright and Repo) through CCIL from Feb'02. To add to transparency, RBI also introduced an anonymous order matching system sans brokers for Government securities in Aug'05. This resulted in a dramatic change in the market microstructure. Brokers became increasingly redundant as market participants started trading using the anonymous order matching system and within a very short span of time, about 80% of the market deals became deals without the convenience of the brokers. As all deals were being settled through CCIL, it started disseminating important information about the market to improve transparency in the market. CCIL also started estimating Zero curves and used the same for valuation and margining purpose. CCIL also introduced Delivery versus Payment – II mechanism in April'04 and added further comfort to the market.

Interest rate cycle in India moved from high interest regime to low interest rate regime and back to high interest regime during period under our study. There have been some important regulatory changes through introduction of Primary Dealers system and structured auction system using multiple pricing mechanisms. The Fiscal Responsibility and Budget Management Act (FRBM) helped RBI to move away from supporting primary auctions as devolvement of debt was shifted to Primary Dealers as they became underwriters of the Government securities issuances. The trading activity showed significant changes during the financial years from 2003-04 and 2011-12. It declined during three financial years while increased during other years for which we have used the data (Table 2) for our study.

Table 2: Trading Activity in Government Securities Market

Financial Years (Apr – Mar)	Change in Market Activity (%)
2003-04	46.37
2004-05	-27.99
2005-06	-23.76
2006-07	18.13
2007-08	61.90
2008-09	30.62
2009-10	34.89
2010-11	-1.47
2011-12	21.50

Note: Change in market activity is measured by growth of trading value over previous year

However, the Indian Government bond market remained relatively illiquid and the turnover ratio during April'03 and April'12 varied between 55% and 155% (Table 3).The market heavily depended on domestic institutions for its growth as investment from Foreign Institutional Investors (FII) was restricted with administrative caps. Trading was restricted to few securities and high concentration was on the 5 and 10 year benchmark securities though Government has been issuing securities up to 30 years of maturity.

Chart - 1: Movement of Short term and Long term Yields (Jan'97 - May'12)

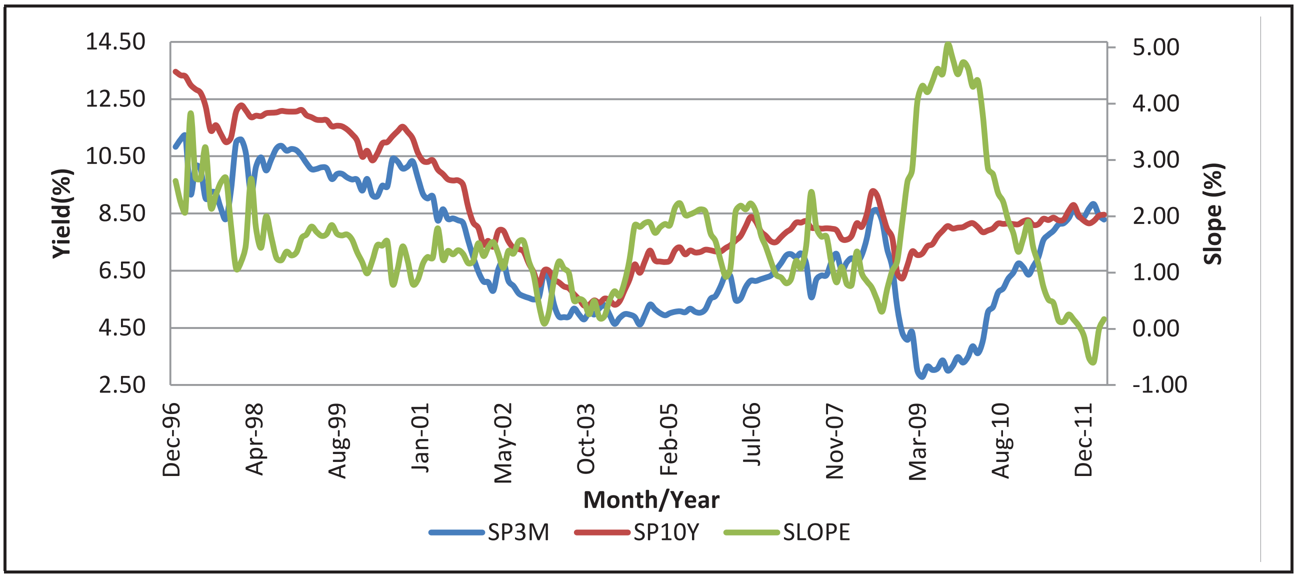


Table 3: Descriptive Statistics of Turnover Ratio (Apr'03 – Apr'12)

Parameters	Value (%)
Mean	103.36
Standard Error	2.53
Median	105.09
Standard Deviation	26.37
Minimum	54.96
Maximum	154.82
Months in data set	109

Note: Turnover ratio has been calculated as the ratio of 12 months total trading value and total outstanding debt.

NSE and CCIL (CCIL also uses NSS equation) have been using Nelson-Siegel functional form for estimation of spot yield curves. Nelson-Siegel functional form is a straight forward equation to estimate the yield of a particular term/ tenor/maturity suing the estimated 4 parameters. The simplistic N-S equation can be solved by an iterative method as it has 4 unknowns in one equation.

$$Y\ m\ = \beta_0 + \ \beta_1 + \beta_2 \quad 1 - \frac{e^{\frac{-m}{\tau}}}{\frac{m}{\tau}} \ - \beta_2 e^{\frac{-m}{\tau}}$$

We used the parameters β_0, β_1, β_2, and τ to estimate the appropriate rates for any term, *m*. We selected maturities, m's, ranging from 3-month to 30 years at appropriate terms like 3-month, 6-month, 1-year, 2-year, 5-year, 7-year, 10-year, 12-year, 15-year, 20-year, 25-year, and 30-year and calculated the time series of yields of these maturities from Jan'97 to May'12. For smoothing purpose, we converted the daily interest rate data into monthly data series by taking monthly averages. This resulted in about 185 monthly observations. We estimated slope of the curve by taking the difference between 10-year and 3-months spot yields (Chart-1).

Table 4: Descriptive Statistics of Historical term Structure of Interest Rate (%)

	3 Months	6 Months	1 Year	2 year	5 year	7 Year	10 year	12 year	15 year	20 year	25 year	30 year
Mean	7.0941	7.1637	7.3009	7.5561	8.1430	8.4132	8.7029	8.8447	9.0074	9.1946	9.3203	9.4104
Std Dev	2.2078	2.1519	2.0781	2.0153	2.0168	2.0424	2.0695	2.0790	2.0843	2.0817	2.0744	2.0669
Max	11.2150	11.3332	11.5540	11.9496	12.8606	13.1780	13.4580	13.5752	13.6955	13.8176	13.8913	13.9404
Min	2.7810	3.0948	3.6558	4.5466	4.8452	5.0052	5.2770	5.4565	5.6962	6.0105	6.2557	6.4463
Median	6.7543	6.8215	6.9070	7.0514	7.6225	7.8531	8.1186	8.2039	8.3110	8.4547	8.5275	8.5724
LP		0.0696	0.1372	0.2551	0.5869	0.2702	0.2897	0.1418	0.1627	0.1872	0.1257	0.0901

Note: LP is the liquidity premia – difference between two nearby rates in our study

Table 5: Descriptive Statistics of Changes in Historical term Structure of Interest Rate (%) and Correlation Coefficients

TYPE	Tenor	3M	6M	1Y	2Y	5Y	7Y	10Y	12Y	15Y	20Y	25Y	30Y
MEAN		-0.0001	-0.0002	-0.0002	-0.0002	-0.0003	-0.0003	-0.0003	-0.0003	-0.0003	-0.0002	-0.0003	-0.0003
STD		0.0046	0.0043	0.0039	0.0034	0.0028	0.0027	0.0027	0.0027	0.0028	0.0029	0.0030	0.0031
N		185	185	185	185	185	185	185	185	185	185	185	185
Correlation Coefficients (with Fisher's Z Transformation)													
		3M	6M	1Y	2Y	5Y	7Y	10Y	12Y	15Y	20Y	25Y	30Y
CORR	3M	1											
CORR	6M	0.9945	1										
CORR	1Y	0.9610	0.9845	1									
CORR	2Y	0.8719	0.9161	0.9707	1								
CORR	5Y	0.6937	0.7429	0.8184	0.9116	1							
CORR	7Y	0.6239	0.6648	0.7309	0.8250	0.9786	1						
CORR	10Y	0.5586	0.5888	0.6343	0.7214	0.9140	0.97++	1					
CORR	12Y	0.5317	0.5573	0.5999	0.6729	0.8709	0.9478	0.9940	1				
CORR	15Y	0.5053	0.5264	0.5608	0.6217	0.8143	0.9031	0.99717	0.9913	1			
CORR	20Y	0.4794	0.4965	0.5224	0.5678	0.7393	0.8334	0.9216	0.9506	0.9856	1		
CORR	25Y	0.4624	0.4774	0.4984	0.5326	0.6806	0.7225	0.8685	0.9113	0.9546	0.9907	1	
CORR	30Y	0.4486	0.4623	0.4800	0.5059	0.6326	0.7194	0.8176	0.8651	0.9173	0.9696	0.9937	1

We analyzed the descriptive statistics (Table 4) of the yields and found that the difference between maximum and minimum yield are far higher in the short term than the long term. This is due to the fact that the short term rates are more guided by monetary policy rates and liquidity factors. In the aftermath of financial crisis in 2007-08, RBI supported the market by infusing huge liquidity along with bringing down policy Repo rate and reserve ratios for the Banks. This helped in lower interest rates at the shorter end but the longer end remained more stable. The liquidity premia was highest for the 5 year security followed by 10 year and seven years. This replicates the market structure as large number of deals happens in the market within 5 to 10 year maturities.

The correlation among the yield changes (monthly return) of various selected maturities indicates that nearby yields are highly correlated (Table 5). We have also adjusted these correlation coefficients (r's) with Fisher's Z Transformation to see if the results change by estimating the covariance matrix using adjusted correlation coefficients (r). Though the results were not significantly different, in order to be statistically correct, we report the results with such adjustments. Fisher's r-to-Z transformation is an elementary transcendental function called the inverse hyperbolic tangent function. The reverse, a Z-to-r transformation, is therefore a hyperbolic tangent function. These transformations are needed to compute a weighted mean correlation coefficient and for hypothesis testing. It may be noted that averaged correlation coefficients are not computable directly from raw ***r*** values. Indeed, it is not possible to add, subtract, average, or take standard deviations out of raw ***r*** values. The sampling distribution of Pearson's ***r*** is not normally distributed. Fisher developed a transformation called "Fisher's z' transformation" that converts Pearson's r's to the normally distributed variable z'. The formula for the transformation is:

$$z' = .5[\ln(1+r) - \ln(1-r)],$$

where In is the natural logarithm. It is important to understand that there are two attributes of the distribution of the z' statistic: (1) It is normal and (2) it has a known standard error of:

$$\sigma_{z'} = \frac{1}{\overline{(N-3)}}$$

Fisher's z' is used for computing confidence intervals on Pearson's correlation and for confidence intervals on the difference between correlations.

Further some of the empirical stylized facts (Chart 2) about term structure of interest rate in India are:

1. Interest rates are mean reverting and changes have leptokurtic distributions (Chart 2).
2. Autocorrelation functions of interest rate changes are fast decaying – daily changes can be assumed to be

auto-correlated (Chart 3)

3. Autocorrelation functions of squared and absolute changes are slow decaying (volatility clustering and leverage effects).

Chart - 2: 10-Year Spot - Monthly Changes 1997-2012

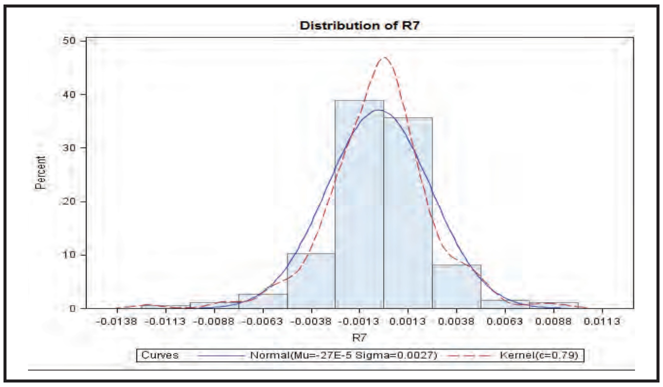
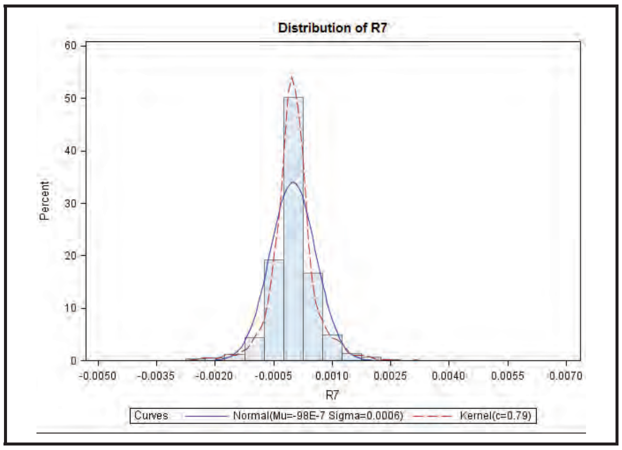


Chart - 3: 10-Year Spot - Daily Changes 1997-2012



2.1. Volatility of Term Structure of Interest Rate

Volatility is an internal part of the financial market, specifically the bond market. We estimated realized volatility of various maturities using an exponentially weighted moving average with a decay factor, l = 0.94. This form used for volatility is from the GARCH family and integrated to 1. The equation is widely used and mad popular as a risk measure by Risk Metrics.

The volatility is conditional one as it dynamically changes with new data coming into computation. As we have converted the daily data to monthly yields for various maturities, we also estimated the conditional volatility of these maturities using the above equation (Chart 4). Short term conditional volatilities (3 months and 1 year) have been higher compared to 5 year and 10 year maturities.

Volatility of 10 year yield has been the relatively lower since vis-à-vis other maturities as higher liquidity in 10-year benchmark might have helped to bring down the volatility with better price discovery mechanism. The 10-year benchmark securities remain the most liquid security in Indian sovereign bond market. During 2011-12, two 10-year securities maturing in 2021 (7.80% GOI 2021 and 8.79% GO 2021) combined together to take a market share of about 53% of the total trading activity in the market. Both these securities have very high turnover ratio vis-à-vis other securities in the market. While the long term rate volatility is generally influenced by major macro factors like growth opportunities in future, the short term rate volatility is more guided by monetary policy considerations, liquidity, inflation expectation, etc.

3. PRINCIPAL COMPONENT ANALYSIS (PCA) AND YIELD CURVE

Principal Component Analysis is a way of identifying patterns in data, and expressing the data in such a way as to highlight their similarities and differences. PCA is a powerful tool for analysing data. The other main advantage of PCA is that once you have found these patterns in the data, and you compress the data, i.e., by reducing the number of dimensions, without much loss of information. Since the PCA model explicitly selects the factors based upon their contributions to the total variance of interest rate changes, it may help in hedging efficiency when using only a small number of risk measures. Factor analysis is a general name denoting a class of procedures primarily used for data reduction and summarization. Factor analysis is an interdependence technique in that an entire set of interdependent relationships is examined without making the distinction between dependent and independent variables. Factor analysis is used in the following circumstances: to identify underlying dimensions, or factors, that explain the correlations among a set of variables; to identify a new, smaller, set of uncorrelated variables to replace the original set of correlated variables in subsequent multivariate analysis (regression or discriminant analysis); to identify a smaller set of salient variables from a larger set for use in subsequent multivariate analysis. Mathematically, each variable is expressed as a linear combination of underlying factors. The covariation among the variables is described in terms of a small number of common factors plus a unique factor for each variable. If the variables are standardized, the factor model may be represented as:

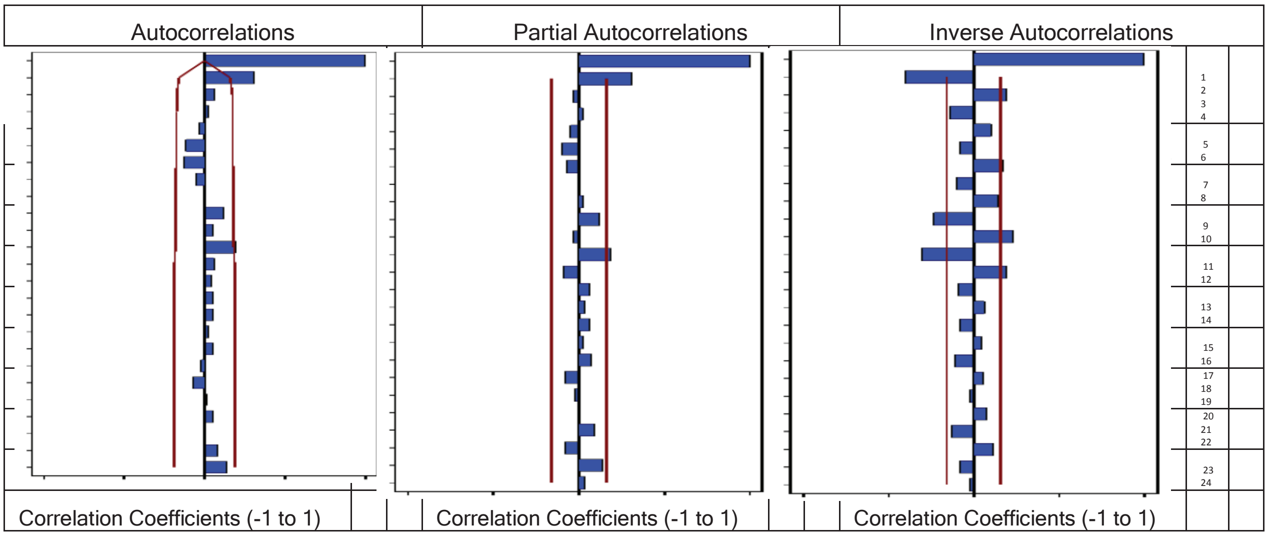
$$X_i = A_{i\,1}F_1 + A_{i\,2}F_2 + A_{i\,3}F_3 + \ldots + A_{im}F_m + V_iU_i$$

where

X_i = *I*th standardized variable

A_{ij} = standardized multiple regression coefficient of variable *i* on common factor *j*

Chart -3 : Auto Correlations upto 24 lags



F = common factor

V_i = standardized regression coefficient of variable *i* on unique factor *i*

U_i = the unique factor for variable *i*

m = number of common factors

The unique factors are uncorrelated with each other and with the common factors. The common factors themselves can be expressed as linear combinations of the observed variables.

$F_i = W_{i1}X_1 + W_{i2}X_2 + W_{i3}X_3 + \ldots + W_{ik}X_k,$ where,

F_i = estimate of *i*th factor

W_i = weight or factor score coefficient

k = number of variables

It is possible to select weights or factor score coefficients so that the first factor explains the largest portion of the total variance. Then a second set of weights can be selected, so that the second factor accounts for most of the residual variance, subject to being uncorrelated with the first factor. This same principle could be applied to selecting additional weights for the additional factors. For factor analysis to be efficient, it is important that an appropriate sample size should be used. As a rough guideline, there should be at least four or five times as many observations (sample size) as there are variables. In PCA, the total variance in the data is considered. The diagonal of the correlation matrix consists of unities, and full variance is brought into the factor matrix. Principal components analysis is recommended when the primary concern is to determine the minimum number of factors that will account for maximum variance in the data for use in subsequent multivariate analysis. The factors are called *principal components*.

4. APPLICATION OF PCA ON INDIAN SOVEREIGN TERM STRUCTURE OF INTEREST RATE

The PCA model assumes that the term structure movements can be summarized by a few composite variables. These new variables are constructed by applying PCA to the historical interest rate changes. The use of PCA in the bond markets has revealed that three principal components – height, slope and curvature of the yield curve are generally sufficient in explaining the variation in interest rate changes. The PCA approach to term structure assumes the following:

Where Δc_i is a set of realizations of principal components. The principal components, Δc_i are linear combinations of interest rate changes. And PCA tells us that not all the components, Δc_i, have equal significance. The first component explains the maximum percentage of the total variance of interest rate changes. The second component is linearly independent (i.e., orthogonal) of the first component and explains the maximum percentage of the remaining variance, the third component is linearly independent (i.e., orthogonal) of the first two components and explains the maximum percentage of the remaining variance, and so on. If yield curve shifts result from a few systematic factors, then only a few principal components can capture yield curve movements. Moreover, since these components are constructed to be independent, they also help in simplifying the task of managing interest rate risk. The principal components with low eigen values make little contribution in explaining the interest rate changes, and hence these components can be removed without losing significant information. This not only helps in obtaining a low-dimensional parsimonious model, but also reduces the noise in the data due to unsystematic factors.

Chart 4: Monthly Volatility of Term Structure (1997-2012)

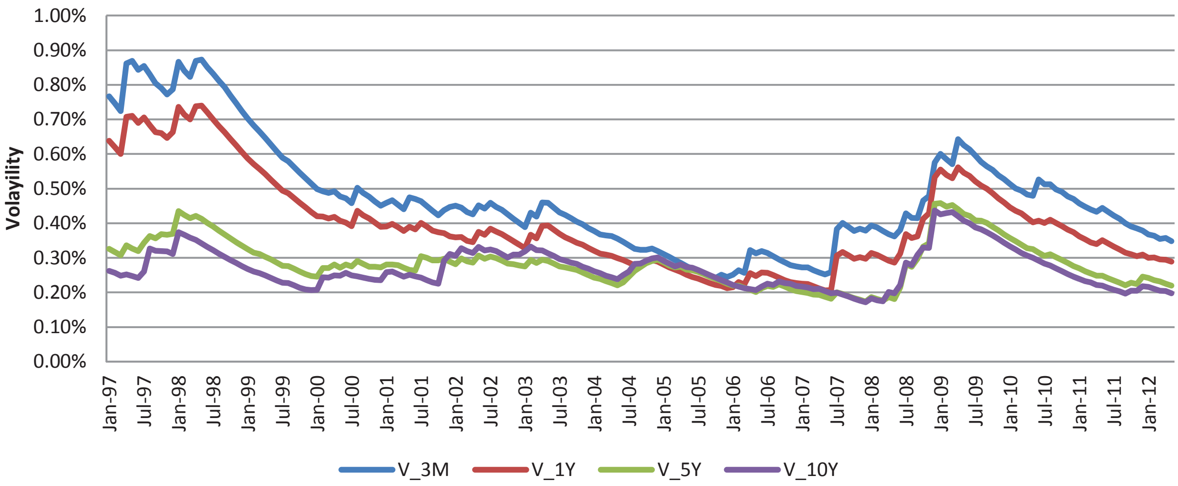


Table 4: Eigen values of the Covariance Matrix (Jan'97 to May'12)

Total Variance 0.0001303293				
Factors	**Eigen value**	**Difference**	**Proportion**	**Cumulative**
PC1	0.00009884	0.00007393	0.7584	0.7584
PC2	0.00002491	0.00001964	0.1911	0.9494
PC3	5.27008E-6	4.05678E-6	0.0404	0.9899
PC4	1.2133E-6	1.1136E-6	0.0093	0.9992
PC5	9.96959E-8	9.48736E-8	0.0008	0.9999
PC6	4.8223E-9	4.48966E-9	0.0000	1.0000
PC7	3.3264E-10	2.8511E-10	0.0000	1.0000
PC8	4.7531E-11	3.3076E-11	0.0000	1.0000
PC9	1.4455E-11	1.4061E-11	0.0000	1.0000
PC10	3.9384E-13	3.7873E-13	0.0000	1.0000
PC11	1.5111E-14	6.8081E-15	0.0000	1.0000
PC12	8.303E-15		0.0000	1.0000

PCA has been applied to the monthly yield changes data from Jan'97 to May'12 for the set of maturities discussed in Section 2. Table 6 gives the key factors of Indian sovereign term structure changes. The table reports the eigen vectors and eigen values of the covariance matrix of monthly changes in the Indian zero-coupon yields from Jan'97 through May'12. The first three principal components explain a major part of the total variance of interest rate changes. This result is consistent with other studies. The first factor accounts for 75.84% of the total variance, while the second and third factors account for 19.11% and 4.04%, respectively. In sum, the first three principal components explain 98.99% of the variability of the data, which indicates that these factors are sufficient for describing the changes in the term structure in India.

Chart 5 shows the shape of the eigen vectors corresponding to the first three principal components which explained most of the variances. These shapes give the impact of a unit change in each principal component on the term structure of interest rates. The change in the zero-coupon rates is plotted against the maturity terms with respect to each principal component. The first principal component basically represents a parallel change in yield curve, which is why it is usually named the level or the height factor. The second principal component represents a change in the steepness, and is named the slope factor. The third principal component is called the curvature factor, as it basically affects the curvature of the yield curve by inducing a butterfly shift.

An unit change of the i^{th} factor cause a change a_{jt} for each maturity t-year rate. Since factors are independent of each

Chart 5: Impact of Three most significant Components on Yield Curve

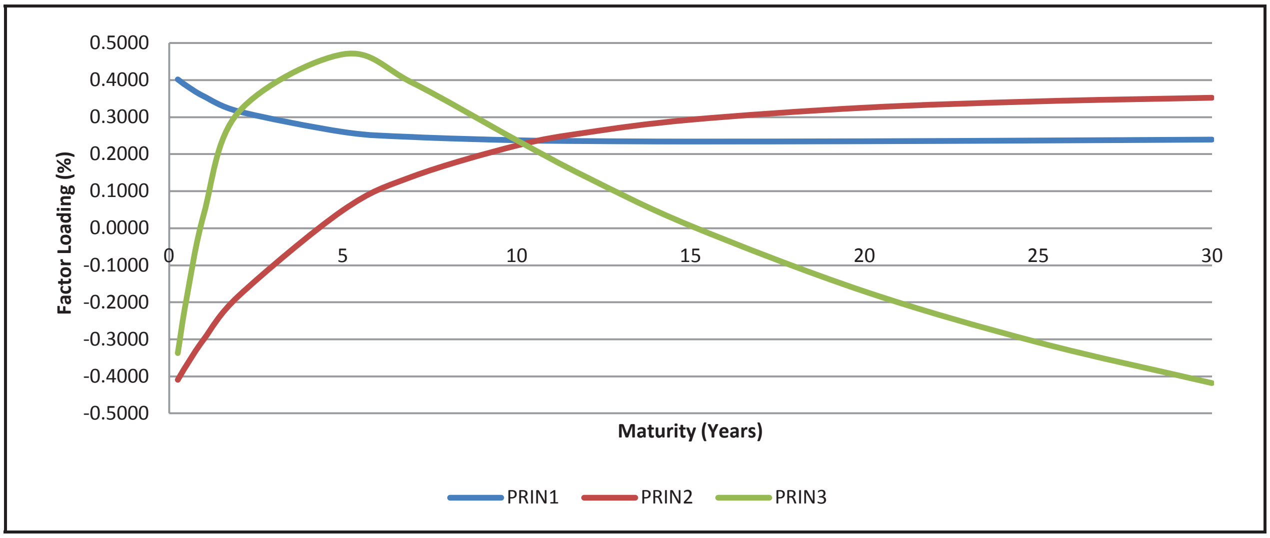


Chart 6: Scree Plot of Factor Loading

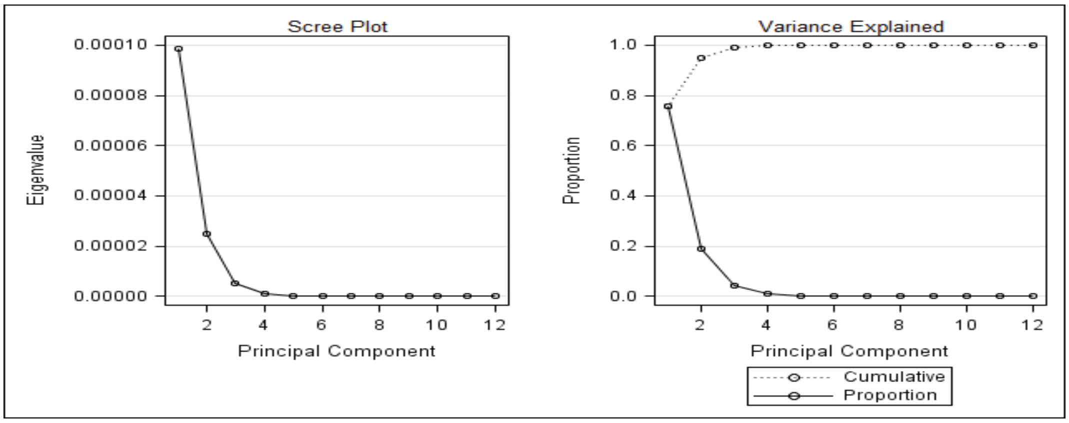


Chart 6A: Principal Component Scores for yields (Jan'97–May'12)

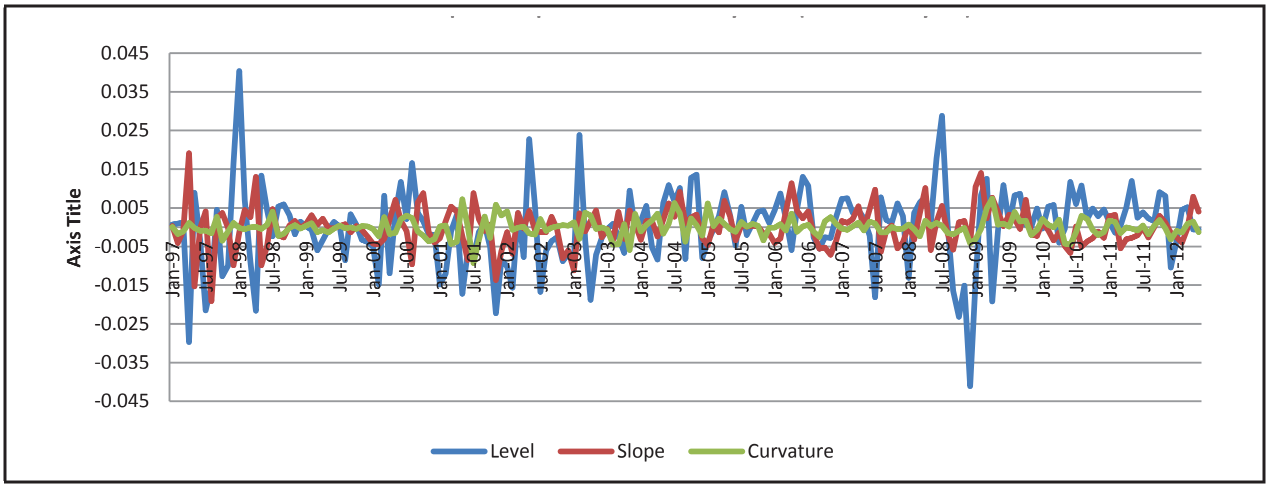
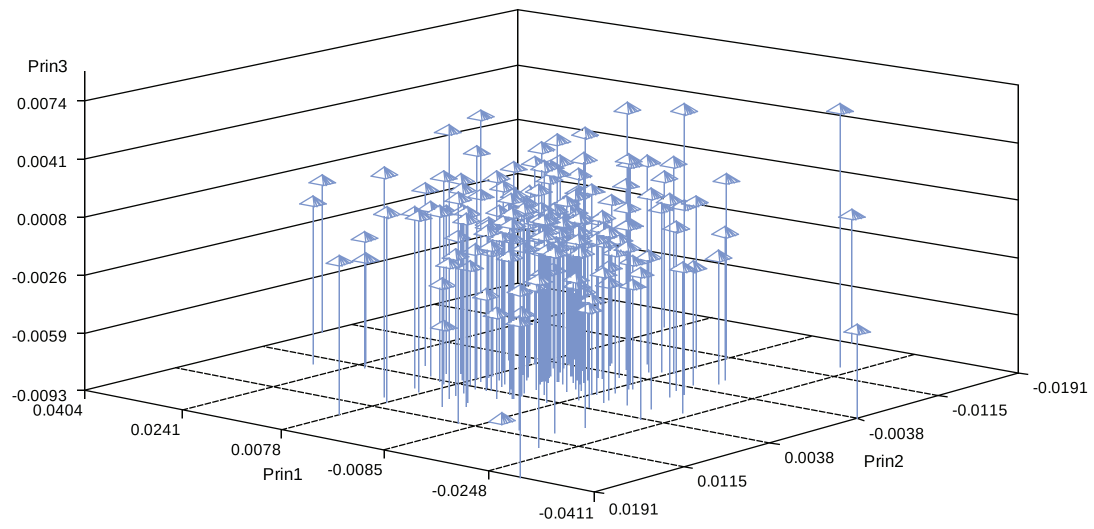


Chart-6B: 3D scatter plot of PC with 2 clusters



other, we may therefore express the total change of the random variable r_t by

$$\Delta r_t = \sum_{j=1}^{k} a_{jt}\,\Delta f_j$$

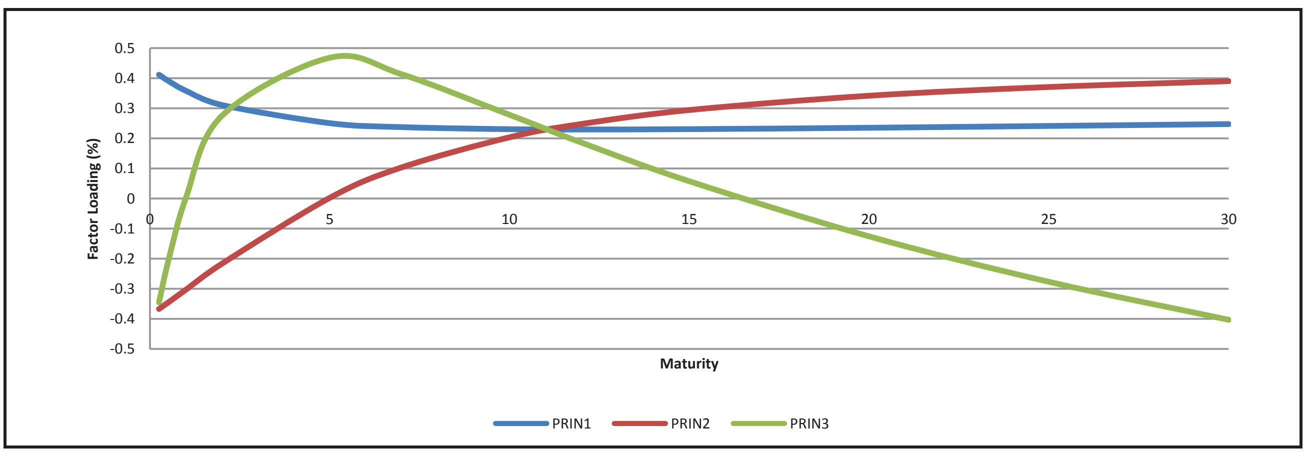
where, f_j is the j^{th} factor, k is the number of factors, a_{jt} is the coefficient, identified by eigenvector analysis, used to approximate the variance.

Our results show the coefficients for factor 1 is always positive, for factor 2, it is negative at start but turns to positive and for factor 3, it starts with negative values, then positive in the middle part of maturity and then turns to negative at the en part of the yield curve (Table – 7).

The result shows that $a_{1,10}$ as 0.2616 implying a unit change in factor 1 causes 0.2376 change in 10-year rate – if the 10-year rate is 8.50%, then it will become 8.52% due to a level factor change of 1 unit. For all factors, it will change to (0.2376%+0.2234%+0.2377% = 0.6987%) 8.56%.

A scree plot (Chart 6) is a plot of the Eigen values against the number of factors in order of extraction. Experimental evidence indicates that the point at which the scree begins denotes the true number of factors. Generally, the number of factors determined by a scree plot will be one or a few more than that determined by the Eigen value criterion. The examination of the Scree plot provides a visual of the total variance associated with each factor. The steep slope shows the large factors. The gradual trailing off (scree) shows the rest of the factors usually lower than an Eigen value of 1. In choosing the number of factors, in addition to the statistical criteria, one should make initial decisions based on conceptual and theoretical grounds. However, at this stage, the decision about the number of factors is not final.

Chart 7A: Impact of Three most significant Components on Yield Curve (Phase 1)



In our result, the coefficients for the first principal component are all positive, so that an increase in the score of the first principal-component results in an increase in all yields. The first principal component can therefore be regarded as a level factor. Since the coefficients are not all equal, a change in the score of the first principal component does not result in a parallel shift; instead, the short end of the curve moves more than the long end.

Table 7: Eigen vectors of 3 Principal Components

Eigenvectors			
Maturity	**PRIN1**	**PRIN2**	**PRIN3**
0.25	0.4019	-0.4094	-0.3372
0.5	0.3849	-0.3703	-0.1914
1	0.3563	-0.2998	0.0383
2	0.3156	-0.1828	0.3139
5	0.2602	0.0485	0.4698
7	0.2464	0.1396	0.3928
10	0.2376	0.2234	0.2377
12	0.2353	0.2584	0.1385
15	0.2342	0.2931	0.0067
20	0.2349	0.3256	-0.1705
25	0.2369	0.3427	-0.3082
30	0.2394	0.3524	-0.4182

The coefficients for the second factor are negative at the short end and then increase to a positive value at the long end. Hence, a change in the score of the second principal component results in an opposite effect on the two ends of the yield curve, and this factor can be viewed as causing a change or twist in the slope of the yield curve. The third principal component has a positive effect on medium yields and a negative effect on short and long-term yields and hence can be interpreted as a hump factor or butterfly. Chart 6A and 6B illustrate the principal-component scores for the first three principal components from Jan'97 to May'12.

Table 8 : Eigen values of the Covariance Matrix (Phase 1)

Total Variance 0.0001493234				
Factors	**Eigen value**	**Difference**	**Proportion**	**Cumulative**
PC1	0.0001106	0.0000803	0.7409	0.7409
PC2	0.0000303	0.0000241	0.2032	0.9441
PC3	0.0000062	0.0000042	0.0415	0.9857
PC4	2.00E-06	1.80E-06	0.0133	0.9990
PC5	1.50E-07	1.50E-07	0.0010	1
PC6	0.00E+00	0.00E+00	0	1
PC7	0.00E+00	0.00E+00	0	1
PC8	0.00E+00	0.00E+00	0	1
PC9	0.00E+00	0.00E+00	0	1
PC10	0.00E+00	0.00E+00	0	1
PC11	0.00E+00	0.00E+00	0	1
PC12	0.00E+00	0.00E+00	0	1

The study also tried to explore if the structure of the market has gone through any significant changes after structured settlements system, anonymous trading platform, etc. were introduced by RBI. We divided the period first into two phases – Phase 1 being the period from Jan'97 to Dec'04 (96 months) and Phase 2 being the period from Jan'05 to May'12 (89 months). We found that the Phase 1 has behaved

Chart 7B: Impact of Three most significant Components on Yield Curve (Phase 2)

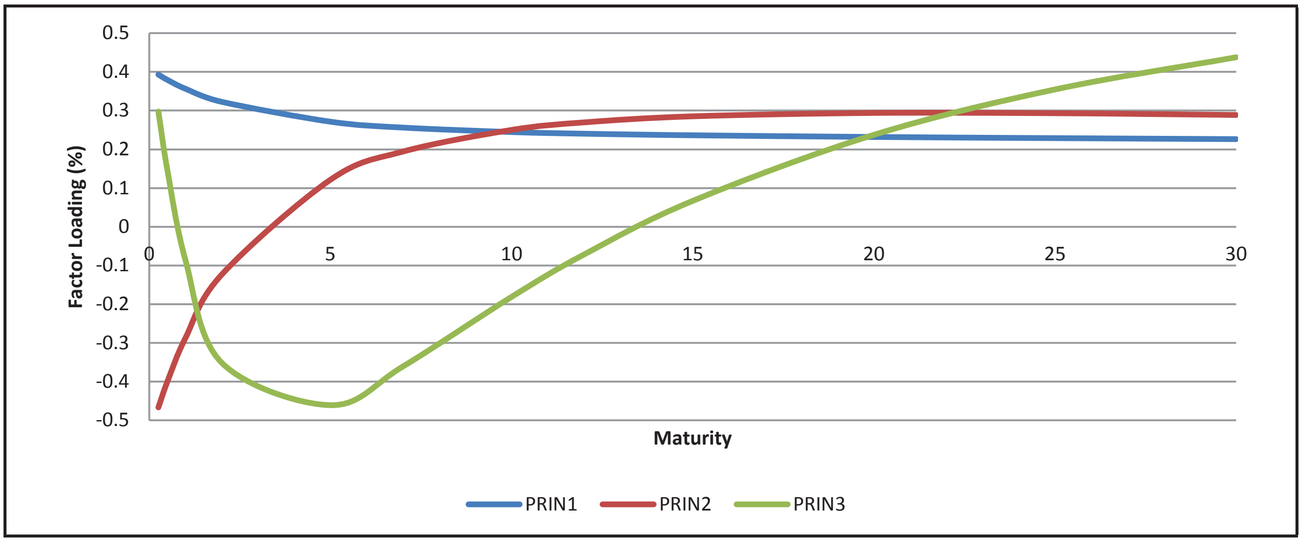
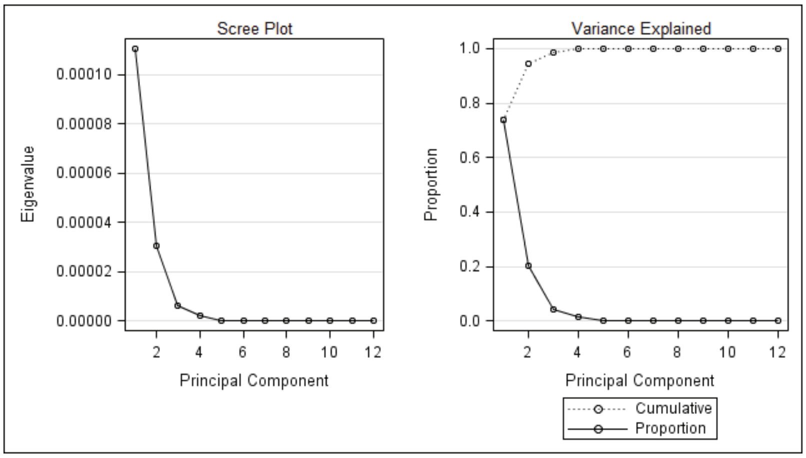


Chart - 8: Scree Plot of Factor Loading (Phase I)



in more or less the similar manner with regard to explanation of sovereign yield changes like the entire period. The first 3 factors explained 98.57% (Table – 8) of the variations. The first factor accounts for 74.09% of the total variance, while the second and third factors account for 20.32% and 4.15%, respectively which indicates that these factors are sufficient for describing the changes in the term structure of interest rate in Indian sovereign bond market.

Table 9: Eigen vectors of 3 Principal Components (Phase 1)

	Eigen vectors		
Maturity	**PRIN1**	**PRIN2**	**PRIN3**
0.25	0.411641	-0.367430	-0.345230
0.5	0.391616	-0.346295	-0.212406
1	0.357989	-0.303032	0.002799
2	0.310831	-0.216876	0.274927
5	0.250418	0.002200	0.468091
7	0.237246	0.103402	0.412998
10	0.230330	0.203500	0.278067
12	0.229438	0.247780	0.185660
15	0.230512	0.294094	0.057247
20	0.235167	0.341680	-0.126064
25	0.241112	0.370601	-0.277158
30	0.247294	0.390032	-0.403303

Our results (Table 9) show the coefficients for factor 1 is always positive, for factor 2, it is negative at start but turns to negative at 5 year and for factor 3, it starts with negative values, then positive in the middle part of maturity and then turns to negative at the en part of the yield curve.

The result shows that $a_{1,10}$as 0.2303 implying a unit change in factor 1 causes 0.2303 change in 10 year rate – if the 10-year rate is 8.50%, then it will become 8.52% due to a level factor change of 1 unit. For all factors, it will change to (0.2303%+0.2035%+0.2781% = 0.7119%) 8.56%. Chart – 7A shows the shape of the eigenvectors corresponding to the first three principal components which explained most of the variances. These shapes give the impact of a unit change in each principal component on the term structure of interest rates. The change in the zero-coupon rates is plotted against the maturity terms with respect to each principal component.

The scree plot (Chart 8) also explains the above results.

For Phase 2, we found that first 3 factors explain 99.62% of the changes in the term structure of interest rate in India (Table 10). The results are in line with the Phase 1 and the full period under our analysis. The results also in line with major studies on use of PCA to study term structure of interest rate changes in other global markets.

Our results (Table 11) show the coefficients for factor 1 is always positive, for factor 2, it is negative at start but turns to positive afterwards and for factor 3, it starts with positive values, then negative in the middle part of maturity and then turns to positive at the en part of the yield curve.

Chart 7B shows the shape of the eigen vectors corresponding to the first three principal components which explained most of the variances. These shapes give the impact of a unit change in each principal component on the term structure of interest rates. The change in the zero-coupon rates is plotted against the maturity terms with respect to each principal component.

The result shows that $a_{1,10}$ as 0.2447 implying a unit change in factor 1 causes 0.2447 change in 10-year rate – if the 10-year rate is 8.50%, then it will become 8.52% due to a level factor change of 1 unit. For all factors, it will change to (0.2447%+0.2502%-0.1805% = 0.3137%) 8.53%.

Using eigen vectors, we can estimate the possible interest rate changes in various terms (tenors). Suppose, we have a

Table 10 : Eigen values of the Covariance Matrix (Phase 2)

Total Variance 0.0001064713				
Factors	**Eigen value**	**Difference**	**Proportion**	**Cumulative**
PC1	0.000083	0.000063	0.7760	0.7760
PC2	0.000020	0.000016	0.1866	0.9626
PC3	3.57251E-6	3.21154E-6	0.0336	0.9962
PC4	3.60963E-7	3.19747E-7	0.0034	0.9996
PC5	4.12156E-8	3.77727E-8	0.0004	1
PC6	3.44293E-9	3.30771E-9	0	1
PC7	1.3522E-10	1.3223E-10	0	1
PC8	2.9902E-12	2.9005E-12	0	1
PC9	8.9714E-14	8.2093E-14	0	1
PC10	7.6212E-15	8.249E-16	0	1
PC11	6.7963E-15	1.8302E-15	0	1
PC12	4.966E-15		0	1

Table 11: Eigen vectors of 3 Principal Components (Phase 2)

Eigen vectors			
Maturity	**PRIN1**	**PRIN2**	**PRIN3**
0.25	0.393104	-0.466570	0.297748
0.5	0.379436	-0.399575	0.145611
1	0.356373	-0.286736	-0.087204
2	0.322723	-0.123404	-0.351440
5	0.271403	0.121505	-0.460850
7	0.256227	0.194300	-0.360906
10	0.244658	0.250202	-0.181103
12	0.240418	0.269615	-0.071652
15	0.236323	0.285349	0.066882
20	0.232020	0.294065	0.237510
25	0.228944	0.293347	0.354730
30	0.226447	0.288872	0.437885

situation in which we expect 2 year rate to change by 5bps and 10 year rate to change by 2bps. We believe that first two components explain most of the variations in the curve. Using the PCA eigen vectors we can estimate the rate change expected in one year by solving the following equation:

Using Cramer's rule, we can solve the above equation to get factor scores for both Principal Components which will be common to all maturities. In this case these are 0.135014 and -0.05209. Hence, the expected increase in 1 year rate would be 0.356373 * 0.135014 +0.286736 * 0.05209 = 0.063051.

We also wanted to study if the recent financial crisis had any impact on the structure of changes in term structure dynamics. We divided the entire period again into two parts - Phase 1 up to 2007 (before Financial crisis) and after 2007. The data up to 2007 reveals that first three factors explain about 98.74% of the variations in the term structure while data after 2007 shows that first 3 components explain 99.57% of the variations in the term structure (Table 12). The results are in line with whole period as well as other two sub-periods we already analyzed in the paper.

Table 12: Factors explaining Term Structure shifts

Period	Level	Slope	Curvature	Total Explained

Jan'97 to Dec'07	0.7289	0.2158	0.0427	0.9874
Jan'08 to May'12	0.8234	0.1448	0.0275	0.9957

The results for the whole period of our study as well as for both phases are more or less in similar lines and in sync with results from major studies in the literature. However, there is significant variation in factor explanation if we consider the year-wise analysis of PCA (Table 13). On an average, level shifts explained about 79% of the shifts in term structure while slope changes explained about 17% and curvature changes explained about 4% of the term structure changes in India. However, for some years, level shifts explained changes in term structure with very low value at 58% (2006) while for some years, it was very high at 95% (2008). Curvature (twists) changes are very few in Indian cases. For 2006, the slope changes explaining the term structure shift has been highest at 38%.

Table 13: Year-wise Factors explaining Term Structure shifts

Year	Level	Slope	Curvature	Explained by 3 Factors
1997	0.7949	0.1975	0.0075	0.9999
1998	0.9123	0.0789	0.0088	1.0000
1999	0.8136	0.1525	0.0338	0.9999
2000	0.7130	0.2586	0.0235	0.9951
2001	0.7078	0.1648	0.0903	0.9629
2002	0.9200	0.0677	0.0118	0.9995
2003	0.8070	0.1250	0.0633	0.9953
2004	0.8567	0.0841	0.0585	0.9993
2005	0.7427	0.1580	0.0951	0.9958
2006	0.5815	0.3810	0.0370	0.9995
2007	0.8673	0.1278	0.0048	0.9999
2008	0.9501	0.0448	0.0046	0.9995
2009	0.7727	0.1560	0.0687	0.9974
2010	0.6835	0.2168	0.0930	0.9933
2011	0.7789	0.2055	0.0126	0.9970
2012	0.6759	0.2996	0.0224	0.9979
Average	0.7861	0.1699	0.0397	0.9958
Median	0.7869	0.1570	0.0287	0.9986
STDDEV	0.1006	0.0893	0.0334	0.0090
Max	0.9501	0.3810	0.0951	1.0000
Min	0.5815	0.0448	0.0046	0.9629

We also extracted the 6-month forward rates (6 X 12, 12 X 18, 18 X 24, …….114 X 120) up to 10 years and run a PCA to see if the characteristics are on desired lines. The forward rates are extracted from the spot rates (Nelson – Siegel parameters). These zero curves (spot) are modeled ones: they are generated by some parametric model, and consequentially have significant built in smoothness associated. In fact, the N-S method is tailor made for PCA - it is a parametric method, with the entire curve described by a single set of parameters representing the long-run level of interest rates, the slope of the curve and humps in the curve. We found that for the forward curves, first 3 factors explain all variations – the first principal component (level shift) explaining about 76% of shifts in the curves, the second component (slope) explaining about 19% of shifts and the third component explaining about 4% of the shifts in the curves (Table 14).

Table 14: Year-wise Factors explaining shifts in Forward Curves

Year	Level	Slope	Curvature	Explained by 3 Factors
1997	0.5176	0.4630	0.0192	0.9998
1998	0.7994	0.1738	0.0265	0.9997
1999	0.7899	0.1644	0.0455	0.9998
2000	0.7433	0.2306	0.0210	0.9949
2001	0.7987	0.1172	0.0697	0.9856
2002	0.9221	0.0632	0.0137	0.9990
2003	0.7825	0.1725	0.0394	0.9944
2004	0.9271	0.0413	0.0314	0.9998
2005	0.7428	0.2256	0.0274	0.9958
2006	0.8000	0.1442	0.0552	0.9994
2007	0.5692	0.4107	0.0200	0.9999
2008	0.9551	0.0421	0.0027	0.9999
2009	0.6354	0.2344	0.1254	0.9952
2010	0.5876	0.3040	0.0788	0.9704
2011	0.8623	0.1113	0.0253	0.9989
2012	0.7693	0.2034	0.0260	0.9987
Average	0.7626	0.1939	0.0392	0.9957
Median	0.7862	0.1732	0.0270	0.9990
STDDEV	0.1290	0.1200	0.0306	0.0077
Max	0.9551	0.4630	0.1254	0.9999
Min	0.5176	0.0413	0.0027	0.9704

5. CONCLUSION

Principal Component Analysis has been widely used to study the shift in the term structure of interest rate. We have used PCA to identify the factors which are responsible for changes in yield curve. The results indicate that the three factors provide us most of the variations in the term structure shift in India market. The study finds that the first three principal components explain a major part of the total variance of interest rate changes. This result is consistent with other studies. In sum, the first three principal components explain about 98% of the variability of the data, which indicates that these factors are sufficient for describing the changes in the term structure in India. Even after dividing the period into two phases, the results were more or less the same. About 98% of the changes are explained by three factors – level, slope and curvature. The results were similar to most of the studies in developed markets. The dynamics of the change in term structure of interest rate has not changed much after the onset of the financial crisis.

REFERENCES

Adrian, T., & Wu, H. (2009). The Term Structure of Inflation Expectations, Working Paper. Federal Reserve Bank of New York.

Ang, A., Boivin, J., Dong, S., & Loo-Kung, R. (2010). Monetary Policy Shifts and the Term Structure. *Review of Economic Studies*, Forthcoming.

Buraschi, A., Cieslak, A. & Trojani, F. (2010). Correlation Risk and the Term Structureof Interest Rates, Working Paper. University of Lugano.

Campbell, J. Y., Sunderam, A., & Viceira, L. M. (2011). Inflation Bets or Deflation Hedges? The Changing Risk of Nominal Bonds, Working Paper. Harvard Business School.

Chan, K. C., Karolyi, G. A., Longstaff, F. A., & Sanders, A. B. (1992). An Empirical Comparison of Alternative Models of the Short-Term Interest Rate. *Journal of Finance*, 47(3), pp. 1209-1227.

Cieslak, A., & Povala, P. (2011). Understanding Bond Risk Premia, Working Paper. Northwestern University and University of Lugano.

Cox, J. C., Ingersoll, J. E., & Ross, S. A. (1985). A Theory of the Term Structure of Interest Rates. *Econometrica*, 53(2), pp. 373-384.

Dai, Q. and Singleton, K. J. (2000), Specification Analysis of Affine Term Structure Models. The Journal of Finance, 55(5), pp. 1943–1978

Fleming, M. J. (1997). *The Round-the-Clock Market for U.S. Treasury Securities*. FRBNY Economic Policy Review.

Fontaine, J. S., & Garcia, R. (2011). Bond Liquidity Premia. *Review of Financial Studies*. Forthcoming.

Haubrich, J., Pennacchi, G., & Ritchken, P. (2011). Estimating Real and Nominal Term Structures using Treasury Yields, Inflation, Inflation Forecasts, and Inflation Swap Rates, Working Paper. Federal Reserve Bank of Cleveland.

Hautsch, N., & Ou, Y. (2008). Yield Curve Factors, Term Structure Volatility, and Bond Risk Premia, Working Paper. Humboldt University of Berlin.

Hu, X., Pan, J., & Wang, J. (2011). Noise as Information for Illiquidity, Working Paper. MIT Sloan School of Management.

Jones, C. M., Lamont, O., & Lumsdaine, R. L. (1998). Macroeconomic News and Bond Market Volatility. *Journal of Financial Economics*, 47(3), pp. 315-337.

Kim, D. H., & Singleton, K. J. (2011). Term Structure Models and the Zero Bound: An Empirical Investigation of Japanese Yields, Working Paper. Yonsei University and Stanford University.

Litterman, R., & Scheinkman, J. (1991). Common Factors Affecting Bond Returns. *Journal of Fixed Income*, Vol. 1(1), pp. 54-61.

Morgan, J. P. (2011). The Domino Effect of a US Treasury Technical Default, US Fixed Income Strategy.

Maitland, A. J. (2002). Interpolating the South African Yield Curve using Principal - Component Analysis: A Descriptive Approach. *South African Actuarial Journal*, 2, pp. 129-45.

Mishkin, F. (2006). *The Economics of Money, Banking and Financial Markets*. Addison-Wesley.

Nath, G. C., Gaurav, Y., & Wagle, A. (2006). Estimating a Reliable Benchmark Sovereign Yield Curve in an Emerging Bond Market (CCIL).

Nelson, C. R. & Siegel, A. F. (1987). Parsimonious Modeling of Yield Curves. *Journal of Business*, 60(4), pp. 473-489.

Phoa, W. (1997). Can You Derive Market Volatility Forecast from the Observed Yield Curve Convexity Bias? *Journal of Fixed Income*, 7(1), pp. 43-53.

Price, K. V., Storn, R. M., & Lampinen, J. A. (2005). Differential Evolution: A Practical Approach to Global Optimization. Berlin: Springer.