NSE RESEARCH INITIATIVE

PAPER NO.: 13

Idiosyncratic Factors in Pricing Sovereign Bonds:
An Analysis of the Government of India Bond Market

Research Team, National Stock Exchange of India Limited
(Gangadhar Darbha, Sudipta Dutta Roy and Vardhana Pawaskar)

National Stock Exchange of India Limited

Idiosyncratic Factors in Pricing Sovereign Bonds: An Analysis of the Government of India Bond Market

Research Team, National Stock Exchange of India Limited
(GANGADHAR DARBHA*, SUDIPTA DUTTA ROY AND VARDHANA PAWASKAR)

Research on term structure estimation has in recent years shifted from the specification and testing of alternative functional forms of the pricing equation to the analysis of factors over and above the present value relation that determine the pricing of a bond. Many studies have accordingly examined the importance of various security specific attributes for bond pricing. The present exercise contributes to the existing empirical literature with an analysis of the Government of India bond market. The term structure is estimated using the Nelson-Siegel functional form, using daily information on secondary market trades in Government of India securities from NSE-WDM for the period January 2000 to June 2001. We find that significant pricing errors exist when Government bonds are priced using the term structure alone. Security specific factors such as residual maturity, time since issuance or age, current yield and issue-size account for most of the variation in pricing errors.

JEL Classification: G12, G15

Keywords: Term Structure, Zero CouponYield Curve, Bond Pricing, Liquidity Premium

* corresponding author. E-mail: gdarbha@nse.co.in. We would like to thank Ajay Shah, Rangarajan Sundaram, Susan Thomas, Golak Nath and participants at the UTI institute of Capital Markets conference and the Money & Finance conference at IGIDR where earlier versions of the paper were presented. The authors are consultants with the National Stock Exchange. The views expressed are theirs and not those of the NSEIL.

Idiosyncratic Factors in Pricing Sovereign Bonds: An Analysis of the Government of India Bond Market

The term structure of interest rates - the relationship between interest rates in the economy and the term to maturity - forms the basis for the valuation of all fixed income instruments. Modeled as a series of cashflows due at different points of time in the future, the underlying price of any fixed income instrument can be expressed as the sum of the present values of the cashflows, with each cashflow discounted by the rate for the associated term to maturity. The term structure of interest rates thus forms the core factor determining the price of any fixed income instrument. Empirical literature for developed countries has demonstrated, however, that other economic factors also exist which cause individual bonds to be priced differently from that implied by the term structure. The effect of such factors is often quite significant. Explanations offered include the benchmark phenomenon (non-benchmark, illiquid bonds trading at prices lower than similar benchmark securities), coupon effects, effects of differential taxation of income and capital gains, tax clientele effects (different rates of tax for different entities) and eligibility of some classes of bonds for special purposes such as overnight repos and in lieu of estate taxes[1] . Also, it is recognised that distinguishing between the different phenomena is rendered difficult by the fact that securities are affected by multiple factors at the same time.

What the empirical literature has unequivocally established is that the role of these factors is non-trivial - even for the simplest, relatively homogenous category of Government bonds - and manifests in terms of large pricing errors when security prices are derived only as present value relations from an estimated term structure[2] . More importantly, the pricing errors are usually systematic in nature (Bliss [1996]), reflecting illiquidity premia over relatively liquid bonds. This in turn implies that, not properly accounted for, these observations can bias the estimated term structure itself. Further, analysis of the statistical significance of these factors and a measure of the magnitude of their impact is critical for arriving at estimated prices for each individual bond. In the light of these reasons, a study of the nature and extent of impact of non-present-value factors should form an integral part of an exercise that seeks to analyse pricing of bonds.

[1] These are all related to security-specific features and to that extent 'idiosyncratic', as distinct from the 'common' term structure underlying the pricing of all bonds. See McCulloch [1975], Elton & Green [1998] and Eom, Subrahmanyam & Uno [1998].

[2] See Elton & Green [1998] for an analysis in the US context and Eom, Subrahmanyam & Uno [1998] for the Japanese bond market.

The empirical literature has followed two approaches to analysing the impact of non-present-value factors - (i) testing the importance of security-specific factors on pricing and (ii) modeling the quantitative impact of these factors on pricing. The studies by Bliss [1996] and Eom, Subrahmanyam & Uno [1998] (ESU [1998] hereafter) for the US and Japanese bond markets respectively fall in the first genre. Bliss [1996] finds particular bonds being systematically over (under)-priced relative to the estimated term structure, implying a pattern attributable to security attributes. ESU [1998] establish the importance of security attributes by analysing the additional explanatory power that they have over and above the present value relation. The extent of impact of tax clienteles and differential tax rates has been analysed, mainly in the US context, by McCulloch [1971, 1975], Carleton and Cooper [1976] and Schaefer [1981].

Among studies that fall in the latter category, Subramanian [2001] controls for the possible impact of illiquid securities on estimation of the term structure for Government of India (GoI) bonds. The author identifies volume traded and number of trades in a security as proxies for relative liquidity and designs a weighting scheme in terms of a tan-hyperbolic function of the number of trades and total volume traded. Such approach, it must be mentioned, only controls for the impact of illiquidity on the estimated term structure; it does not explicitly model the impact in terms of either identifying the features that cause pricing anomalies for the latter category, nor analyses the nature and significance of such effects. Elton & Green [1998] (EG [1998] hereafter) test for the impact of non-present-value factors on pricing. The authors find that volume traded subsumes the impact of security-specific attributes like maturity and age on pricing, and accordingly account for the impact of non-present value factors by including volume traded as an additional variable in the pricing equation.

The focus of the present study is on testing for the importance of non-present value factors in terms of the proportion of variation in pricing errors explained by security-specific attributes. It is, in our view, the first comprehensive attempt to provide an analysis of the importance of various idiosyncratic factors in pricing of GoI bonds.

The data for the study are compiled from information on secondary market trades in Government securities reported on the Wholesale Debt Market (WDM) segment of the National Stock Exchange (NSE). The estimation framework is adapted from our earlier study (Darbha, Dutta Roy and Pawaskar [2000][3]) and takes into account various institutional details related to secondary market trading in India. The objective function is so specified as to enable estimation of the term structure in a manner that is robust to the non-availability of observations at long maturities. We find that mean absolute errors, while

[3] This version of our paper provides daily estimates of the term structure from February 1998 to December 2000.

comparable to those obtained by earlier studies in the Indian context, imply significant impact of non-present value factors. The importance of these factors is established through an analysis of the fitted price errors. Following the empirical literature for other countries and based on analysis of the data, we identify a set of security attributes - residual time to maturity, time since issuance, outstanding issue size and coupon - that could influence investor preferences. We find that a simple linear regression of errors on these factors explains most of the variation in the first stage errors. Further, the explanatory power of these security features is found to have increased significantly in the more recent period, indicating greater importance of non-present value factors in pricing.

Our scheme of presentation is as follows. Section II provides a brief description of the Indian Government bond market. Section III outlines the econometric methodology and the empirical specification of the estimated model. An account of the data and related estimation issues is presented in Section IV. Results and analysis are presented in Section V. Section VI concludes.

The Indian Government Bond Market

The Indian Government bond market comprises securities issued by the Government of India and the State Governments. Government of India securities include Treasury Bills (T-Bills) with maturity less than a year and dated Government securities (G'secs) with maturities exceeding a year. As on March 31, 2001, there were 116 G'secs outstanding with maturity dates ranging from 1 to 20 years. The total outstanding amount was Rs.3,87,854 crore (1 crore = 10 million). There were 54 T-Bills outstanding for an aggregate amount of Rs.16,980 crore. State Governments had an outstanding of 295 securities comprising Rs.43,176 crore.

The maturity distribution of outstanding G'secs as on March 31, 2001 reveals that over 50 percent of the outstanding issues have a residual maturity less than or equal to 5 years. About 30 percent of the securities lie in the maturity range of 5 to 10 years and the balance 20 percent have maturity beyond 10 years. There are, in fact, only 3 securities with maturity dates beyond 2016. During the financial year 2001-02, the Government of India issued, for the first time, securities beyond 20-year maturity.

The secondary market in Government securities is largely a telephone-based market, with trades subsequently reported on the Wholesale Debt Market segment of the National Stock Exchange (NSE-WDM)[4] and the Subsidiary General Ledger of the RBI (RBI-SGL)[5]. Secondary market activity in

[4] It is mandatory for broker-negotiated trades to be reported on the relevant exchange. Trading in G'secs was permitted on The Stock Exchange, Mumbai (BSE) in October 2000; however, BSE accounts for a very minor proportion of all trades. NSE-WDM, which became operational in June 1994, accounts for 60-70 percent of all trades.

[5] Settlement is done through RBI-SGL; this data set comprises the universe of secondary market trades.

Government securities witnessed an average growth of 91 per cent per annum during the period 1994-95 to 1999-2000. However, the size of the market is small compared to the amounts outstanding; the total turnover in 2000-01 was Rs.4,56,515 crore implying a turnover ratio of about 1.5. Trading in Government securities on the NSE-WDM was thin prior to 1998, but has grown significantly over the last 3 years. Trading is usually spread over the entire maturity spectrum, which, for the purpose of estimating a term structure, has the advantage that there are no gaps in the data at any maturity bracket.

Like in most other markets, secondary market activity is concentrated in a few securities, also referred to as benchmark securities. The identity of these benchmark securities changes over time, and it is difficult to identify the 'optimal' mix of characteristics that distinguish benchmark (liquid) securities from other illiquid Government bonds. Investor interest in these papers is partly driven by the Reserve Bank of India (RBI) policy of re-issuing certain securities at various maturities, which on the one hand increases the notional amount outstanding in these securities and on the other signals RBI preference for emergence of the benchmark. Barring few exceptions, other features – in addition to high outstanding amounts - common to actively traded bonds are a residual time to maturity that lies between 4-8 years and time since issuance not exceeding 3 years **(Table 1)**.

Table 1: Secondary market activity on NSE: most actively traded bonds in 2000-01

Security name	Volume (Rs. crore)	Percentage of total volume	Number of trades	Time since Issuance	Residual maturity	Outsta-nding (Rs. crore)
11.4% CG2008	56202.84	13.11	9589	0.17	7.83	6000
11.3% CG2010	39491.23	9.21	6580	0.21	9.79	9000
12.5% CG2004	35488.64	8.28	6084	6.27	3.73	11196.01
11.03% CG2012	28964.03	6.76	4642	0.24	11.76	9500
11.9% CG2007	20751.25	4.84	3448	2.09	6.91	13500
11.99% CG2009	19341.35	4.51	2980	1.23	8.77	13500
11.15% CG2002	16533.00	3.86	2435	2.83	2.17	5000
11.43% CG2015	12443.57	2.90	1896	0.19	14.81	12000
11% CG2006	10664.00	2.49	1757	0.21	5.79	3000
11.68% CG2006	10016.21	2.34	1555	1.22	5.78	7500
	249896.12	**58.3**	**40966**			

Note: *Residual maturity and time since issuance are in years; outstanding amounts are as at end-March 2001*
Source: *NSE*

A comparison of the attributes and secondary market activity in 2 securities with residual maturity of 10 years is illustrative of the role of non-present-value factors in influencing investor preferences and, in turn, volumes and prices (yields) in the secondary market **(Table 2)**. The 11.50% 2011 security was issued on August 5, 1991 and has a maturity date of August 5, 2011. The 11.50% 2011A was issued on November 23, 2000 and will be redeemed on November 23, 2011. The cashflow amounts of the 2 bonds are the same and the cashflow structures (times to coupon and redemption) almost similar. Secondary market activity in these papers on NSE-WDM reveals a concentration of activity in the more recent issue, resulting in turn in significant (price and corresponding) yield differentials between the two securities.

Table 2: Activity on NSE-WDM: Importance of security-specific factors

(Trade values are in Rs.crore)

	11.5% 2011A		11.5% 2011	
	Traded value	**Weighted YTM**	**Traded value**	**Weighted YTM**
Nov-00	1032.05	11.501	24.00	11.549
Dec-00	3172.00	11.243	16.00	11.165
Jan-01	2417.32	10.725	195.15	10.667
Feb-01	1764.08	10.385	156.00	10.460
Mar-01	915.02	10.299	70.20	10.516
Apr-01	2047.37	10.244	35.00	10.348

***Note**: Outright transactions only; ie. excluding repo trades*
***Source**: NSE*

Similar anomalies in pricing of liquid (benchmark) vis-à-vis illiquid securities exist in almost every maturity segment. A major objective of the present exercise is to analyse the extent of such anomalies to highlight their role in inter-security pricing differences.

Empirical Specification and Econometric Methodology

If the *spot rates of interest* (r_t) for every maturity period are known, then the *present value* of an m-period bond making a series of coupon payments C every period and with redemption value R is:

$$PV = \frac{C}{(1+r_1)} + \frac{C}{(1+r_2)^2} + ... + \frac{C+R}{(1+r_m)^m} \qquad ....(1)$$

The spot interest rate r_t is the interest rate applicable on a cash payment due in t periods. The set of spot rates is the *term structure of interest rates*. The factor $1/(1+r_t)^t$ that is used to discount the value of the future payment to the current period is the *discount factor*.

Empirical estimation of the term structure requires specifying a parametric relation between maturity and spot interest rates. Darbha, Dutta Roy and Pawaskar [2000] adopt the Nelson-Siegel formulation (Nelson & Siegel [1987]) for the derivation of such a relation[6]. Starting from a parsimonious representation of the forward rate function given by

$$f(m,b) = \beta_0 + \beta_1 * \exp(-m/\tau) + \beta_2[(m/\tau) * \exp(-m/\tau)] \qquad ....(2)$$

where 'm' denotes maturity and b=[β_0, β_1, β_2 and τ] are parameters to be estimated, the relevant spot rate function can be derived as

$$r(m,b) = \beta_0 + (\beta_1 + \beta_2) * [1 - \exp(-m/\tau)]/(m/\tau) - \beta_2 * \exp(-m/\tau) \qquad ....(3)$$

The implied long-term [as m $\rightarrow \infty$] and short-term [as m $\rightarrow$ 0] rates in this specification are given by β_0 and ($\beta_0+\beta_1$).

With the spot rate function specified as above, the PV relation is specified using the continuous form of the discount function given by:

$$d(m,b) = \exp\left(-\frac{r(m,b) * m}{100}\right) \qquad ...(4)$$

The estimated price ($\mathbf{p^{est}}$) for each bond is the sum of the present values of all its cashflows:

$$p_i^{est} = \sum_{j=1}^{k} d_j(m,b) * c_j \qquad ...(5)$$

It is common to observe market prices ($\mathbf{p^{mkt}}$) that deviate from the computed present value. For the purpose of the estimation exercise, it is postulated that the observed market price of a bond deviates from its underlying valuation by an error term e_i, which gives the estimable relation:

$$p_i^{mkt} = p_i^{est} + e_i \qquad ...(6)$$

The parameters b = [β_0, β_1, β_2 and τ] are estimated by minimising the sum of squared price errors. Earlier empirical studies (Svensson [1994], Bolder & Streliski [1999]) have found that minimising price errors results in fairly large errors in yield to maturity (YTM) for instruments with short maturities. This is on account of the fact that, since the elasticity of price with respect to one plus YTM equals the duration[7] of a bond, prices are relatively insensitive to yields for short maturities. The optimisation technique that seeks to minimse price errors will consequently lead to over-fitting of long–term YTMs at the expense of short-term YTMs. To correct for this in the empirical estimation, Bolder & Streliski [1999] suggest weighting of each price error by the inverse of its

[6] For other studies on choice of functional form in the Indian context, see Thomas & Saple [2000] and Subramanian [2001].
[7]Weighted average of the times to cashflow, weights being the discounted values of the cashflows.

duration. The appropriate weighting scheme is given by

$$w_i = \frac{1/D_i}{\sum_{j=1}^{n} 1/D_j} \quad ...(7)$$

where Di is the MacCauley duration of the bond given by

$$D(T,C)_t = \frac{\sum_{t=1}^{n} \frac{(CF_t * t)}{(1+YTM(T,C)_t)^t}}{\sum_{t=1}^{n} \frac{(CF_t)}{(1+YTM(T,C)_t)^t}} \quad ...(8)$$

In the empirical exercise, we use the duration-based weighting scheme to estimate equation (6) by minimising the weighted sum of squared price errors. The estimation can be carried out either by maximising the likelihood function or by minimising some other loss function defined over the errors (e_i). The former approach requires specification of the exact form of the distribution of the errors, while the latter takes a pure optimisation route without explicit distributional assumptions about the errors. The parameter estimates, it may be mentioned, are expected to be the same if the errors are normally distributed. However, parameters estimated under such loss functions, which are quadratic in nature, are sensitive to 'outliers', a real possibility in the present context. To reduce the impact of outliers on the parameter estimates, we specify a robust loss function [Beaton-Tukey loss function] that downweights large errors (say Re.1 deviation between model and market price) in the objective function (see Seber & Wild [1989] for details). With no prior knowledge of the error distribution, the choice of this approach also implies that standard statistical inference with regard to the parameter vector (b) may be weak.

The estimation is carried out using the constrained optimisation (CO) module in GAUSS (Schoenberg [1998]). The long-term and short-term rates are constrained to be non-negative, and so also the parameter τ. To the extent these constraints restrict the search procedure of the optimisation algorithm within a meaningful parameter space, they would reduce the overall search time. If the function also happens to be smooth in the relevant region, the constraints would increase overall speed of computation as well. Our experience has been that, when estimation is carried out without imposing these constraints, not only does the computational time increase substantially, but also the resulting parameter estimates are very often outside the meaningful range.

The steps followed in the estimation procedure are as follows:

i. A vector of starting parameters (β_0, β_1, β_2 and τ) is selected,
ii. The discount factor function is determined using these starting parameters,
iii. This is used to determine the present value of the bond cash flows and thereby to determine a vector of *starting* 'model' bond prices,

iv. Numerical optimisation procedures are used to estimate a set of parameters (under the given set of constraints *viz.* non-negativity of long run and short run interest rates) that minimise the specified loss function,
v. The estimated set of parameters is used to determine the spot rate function and therefrom the 'model' prices.

Data Details and Related Estimation Issues

The exercise uses data from the NSE-WDM[8], which constitute, on an average, about 60-70 per cent of the total trades negotiated and comprise those trades that are negotiated through member-brokers. The price information relates to 'traded prices' rather than 'quotes', and is not time-stamped. On every trade date and for each individual trade, we have information on the security traded, traded price[9], traded volume and settlement date[10]. Security details *viz.* date of issue, date of maturity and details of cashflows[11] for the bond, are available from a masterfile of securities available with NSE.

Bulk of the trading is in securities issued by the Central Government, ie. GoI securities; state government securities (SGS) account for a very small number of the trades conducted on any given day. It is useful to mention at this point that, state Governments being perceived as less credit-worthy than the Centre, SGS are issued and traded at a credit spread over GoI securities of same maturity. There are, in addition, differences in perceived credit-worthiness across states that is reflected in inter-state coupon (yield) differentials. To purge the estimated sovereign term structure of any of these effects, the dataset we use comprises only GoI securities.

A widely held perception in the Indian markets is that instruments with maturity less than a year, being traded as money market instruments, reflect pricing considerations different from that of longer-maturity securities. Further, pricing differences are observed between T-Bills and G'secs of the same residual maturity. Subramanian [2001] cites these as reasons for excluding such observations from the sample. Inasmuch as the objective of the current exercise is to analyse the nature and extent of such influences – in addition to providing daily estimates of the term structure – we do not apply any such prior filter on the dataset.

Volume weighted average prices are used, where the average is computed over trades with the same settlement date. This means that for each security, we

[8] Information on trades reported on RBI-SGL is publicly disseminated only on the day of settlement and could have trades conducted on different trade dates, which renders it difficult to use it for the exercise at hand.
[9] Clean price, ie. exclusive of accrued interest.
[10] Which lies within a range of 5 days from trade date in the T+0 to T+5 system allowed for broker-negotiated trades.
[11] Coupon rate and coupon payment dates.

have as many observations as the range of settlement horizons for trades negotiated on a given trade date.

Present value computations require information on time to coupon payments and redemption. These are calculated with reference to the settlement date. Market conventions require computation of accrued interest on a 30/360 basis for instruments with residual maturity exceeding a year and on actual/365 basis otherwise (this includes T-Bills), and these are adhered to in the computation of coupon accrual and time to cashflows.

There are various factors to which intra-security variation in prices can be attributed. First, the scope for price discovery in negotiated deals is limited, and even the dissemination of the transacted price is available to the market after a considerable lag, an outcome of the current state of the market where reporting rules are not very stringent. This may be an important factor contributing to the observed dispersion in prices across different trades in the same security. Further, within the T+5 settlement system, trades negotiated on a given day can have settlement dates varying from current date to 5 days hence. There are two mechanisms through which this exerts an impact on the price. First, expectations about the likely directionality of interest rates would be built into the contract if the term structure is expected to undergo a significant change by the time the deal is settled. To discount the cashflows for deals that do not settle on the current day, therefore, the appropriate rates to be used are those that are expected to prevail on the settlement date. We use implied forward rates - the best predictors of expected future spot rates - to discount these cashflows. The forward rates are derived from the estimated term structure using the relation

$$r_{t1}^{t2} = \left(\frac{(1+r_0^{t2})^{t2}}{(1+r_0^{t1})^{t1}} \right)^{1/(t2-t1)} - 1 \qquad \ldots(9)$$

where r_0^{t2} denotes the spot rate for maturity date t2 as on date 0, r_0^{t1} the spot rate for maturity date t1 as on date 0 and r_{t1}^{t2} denotes the spot rate for maturity date t2 expected to prevail at date t1.

Secondly, the negotiated price for a transaction that does not settle on the same day would need to incorporate the net cost of carry. From the point of view of the seller, the opportunity cost involved in settling a deal T days into the future is approximated by the foregone return in the call money market (say), while the return is given by the coupon that accrues for these days. If the net cost of carry is positive (negative), the negotiated futures price will be higher (lower) than the spot price. To compute the net cost of carry for the purpose of the empirical exercise, we proxy the overnight rate by the short-term rate ($\beta_0 + \beta_1$) derived from the estimated term structure. The cost of carry is added to the estimated spot price to arrive at the estimated futures price.

Results

Tests of model performance

This section presents the results of our empirical exercise. The period of analysis is from January 2000 to June 2001. **Table 3** presents summary statistics of residuals from the term structure estimation. We have reported mean absolute price errors (MAPE; deviations of market prices from the corresponding model prices), which provides a measure to gauge the performance of the chosen functional form and the corresponding estimate of the average term structure. The mean for each month is the average over all trades over all trading days in the month. Over January-December 2000, the mean ranges between 12 to 20 paise (1 paise is one-hundreth of a rupee). These are comparable to those obtained by earlier studies in the Indian context (TS [2000] and Subramanian [2001]) but are higher than those reported for the US (Bliss [1996]) and Japan (ESU [1998]), which is probably indicative of the importance of non-present value factors such as liquidity in the Indian context. Further, errors increase significantly in the more recent period (January-June 2001), probably indicative of the increasing importance of these factors. We have also reported the intra-month standard deviation as a measure of the variation in the mean absolute error across trading days within a month. High values of standard deviation relative to the mean indicate that there is significant intra-month variation in the fit of the model.

Table 3: Monthwise volume and price errors

Month	Volume (Rs. crore)	MAPE	Std. Dev
Jan-00	34771.34	0.192	0.256
Feb-00	41660.62	0.194	0.289
Mar-00	17968.03	0.203	0.399
Apr-00	33033.30	0.171	0.234
May-00	31308.65	0.189	0.269
Jun-00	16480.95	0.256	0.395
Jul-00	27141.13	0.192	0.280
Aug-00	14905.81	0.174	0.264
Sep-00	18492.79	0.169	0.258
Oct-00	22795.94	0.164	0.255
Nov-00	34518.64	0.178	0.331
Dec-00	29331.16	0.197	0.278
Jan-01	64401.93	0.241	0.330
Feb-01	53336.63	0.449	0.681
Mar-01	45212.94	0.579	0.803
Apr-01	44998.44	0.486	0.682
May-01	82109.84	0.668	0.846
Jun-01	84242.37	0.644	0.830
Average		0.332	0.525

A maturity-wise analysis of the model errors **(Table 4)** reveals high errors at the short end (< 1 year maturity), probably indicative of the importance of other factors - besides the term structure - that lead to pricing errors at these maturities. A consistent pattern observed in the data in this maturity segment, for instance, is the higher yields commanded by dated securities over and above T-Bills with similar residual maturity **(Table 5** provides an illustration**)**, presumably on account of the relative illiquidity of the former type of securities[12]. An increasing pattern beyond the 1-year maturity segment is consistent with similar findings of ESU [1998].

Table 4: Error analysis by maturity

Maturity	Mean	Std. Dev	Volume (Rs. crore)
0-1 years	0.099	0.170	54010.25
1-3 years	0.148	0.158	64740.82
3-5 years	0.254	0.364	89977.17
5-7 years	0.347	0.454	61889.41
7-10 years	0.548	0.786	209877.27
> 10 years	0.532	0.694	216215.58

Table 5: Yield difference between G'sec vs T-Bill: an illustration

Date (1)	Maturity (2)	Yield (3)	Maturity (4)	Yield (5)	Yield difference (6) = (3)-(5)
1-Jan-01	0.65	9.94	0.64	9.62	0.32
2-Jan-01	0.79	9.89	0.79	9.57	0.32
2-Jan-01	0.57	9.88	0.56	9.63	0.25
3-Jan-01	0.90	9.95	0.90	9.50	0.45
4-Jan-01	0.56	9.79	0.55	9.48	0.30
5-Jan-01	0.87	9.89	0.86	9.35	0.54
5-Jan-01	0.56	9.72	0.55	9.45	0.27
6-Jan-01	0.63	9.75	0.63	9.38	0.37
10-Jan-01	0.88	9.93	0.88	9.38	0.54
11-Jan-01	0.47	9.93	0.46	9.59	0.34
22-Jan-01	0.59	9.71	0.58	9.43	0.28
24-Jan-01	0.58	9.70	0.58	9.38	0.32

In view of the emerging nature of the Indian debt market, a question may naturally arise as to whether the negative correlation observed between residual maturity and errors is in reality a volume effect attributable to thin trading at

[12] Amihud and Mendelson [1991] report a similar finding in the US context.

the long end. This possibility is dispelled by the maturity-wise aggregate trading volumes reported in Column 4 of Exhibit 4. The figures reveal that, for the entire sample period, trading volumes have in fact been higher at longer maturities. However, the volumes are aggregated over multiple securities in each maturity bucket, each with possibly different trading volumes over the entire period. Two questions may then arise. First, across comparable trading volumes, do errors increase with maturity? If supported by the data, this may be indicative of investor preference in favour of short-medium bonds. Secondly, across comparable maturities, is there any pattern in pricing errors as securities witness higher trading volumes? Existence of a pattern may indicate that higher secondary market activity facilitates the price discovery process. The 2-way (maturity-wise/ volume-wise) analysis presented in **Table 6** is designed to address these issues. The figures reveal a broad pattern of errors increasing with maturity for comparable trading volumes. This may be reflective of investor preferences for short-medium instruments, which would imply an additional premium for longer maturity papers over and above the term structure.

Table 6: Volume-wise and maturity-wise analysis of price errors

Volume	Maturity					
	0-1 years	1-3 years	3-5 years	5-7 years	7-10 years	>10 years
1-5 crore	0.103	0.162	0.263	0.347	0.513	0.584
5-10 crore	0.090	0.152	0.240	0.373	0.531	0.566
10-25 crore	0.103	0.144	0.262	0.358	0.515	0.540
25-50 crore	0.101	0.136	0.270	0.337	0.530	0.524
50-100 crore	0.111	0.145	0.241	0.346	0.564	0.473
>100 crore	0.152	0.150	0.184	0.238	0.633	0.497

However, no clear pattern is evident in the mean errors across trading volumes for comparable maturities. This may seem puzzling, as most empirical studies use volume as a proxy for liquidity (EG [1998], Subramanian [2001]). It is also possible that any pattern between errors and traded volumes, even if it exists, could be blurred on account of the pooling of errors across securities with possibly different extent of secondary market activity but lying in the same maturity bucket. To gauge whether this is indeed the case, we next analyse month-wise volumes and errors for a select set of securities (**Table 7**). The securities have been so chosen as to represent a mix of actively traded and infrequently traded securities at different maturities. The figures reveal two interesting findings:

i. The extent of secondary market activity in any security varies over its lifetime. For any chosen security, a negative relationship is now clearly discernible between volumes and the extent of pricing error; ie. as volumes increase, pricing errors witness a decline;

ii. However, no such pattern emerges across securities, ie. securities with the highest trading volumes do not necessarily have the lowest errors.

Table 7: Pricing errors for select securities

(volumes are in Rs. crore)

	Jan-00		Feb-00		Mar-00		Apr-00		May-00		Jun-00	
	Volume	Error	Volume	Error	Volume	Error	Volume	Error	Volume	Error	Volume	Error
11.15% 2002	1035	0.065	795	0.087	491	0.132	1151	0.101	1705	0.142	755	0.172
11.10% 2003	185	0.124	235	0.121	215	0.182	140	0.166	680	0.093	215	0.181
12.50% 2004	1915.7	0.131	2011.5	0.197	1386.9	0.180	1457.0	0.142	3066	0.108	1941.9	0.146
10.20% 2005	0	0	0	0	0	0	0	0	0	0	0	0
11.00% 2006	0	0	0	0	0	0	0	0	0	0	0	0
11.90% 2007	915	0.323	930	0.159	689	0.201	992.5	0.230	460	0.336	5165	0.175
11.40% 2008	0	0	0	0	0	0	0	0	0	0	0	0
11.99% 2009	3960.1	0.145	4050	0.191	2342.9	0.122	2429.0	0.141	2209	0.108	625	0.376
11.30% 2010	0	0	0	0	0	0	0	0	0	0	0	0
12.32% 2011	2865	0.175	3020	0.183	865.5	0.278	2294.4	0.226	980	0.437	281	0.743
11.03% 2012	0	0	0	0	0	0	0	0	0	0	0	0
10.70% 2020	0	0	0	0	0	0	424.6	0.232	245.4	0.569	197.8	0.747

Contd.....

	Jul-00		Aug-00		Sep-00		Oct-00		Nov-00		Dec-00	
	Volume	Error	Volume	Error	Volume	Error	Volume	Error	Volume	Error	Volume	Error
11.15% 2002	1565	0.106	1600	0.085	1605	0.091	885	0.112	1230	0.116	1091	0.147
11.10% 2003	745.1	0.122	213	0.107	241.4	0.128	346	0.135	695	0.139	450	0.185
12.50% 2004	3102.1	0.124	4181	0.129	2925	0.297	2725.2	0.100	2299.3	0.077	1340	0.069
10.20% 2005	1812.3	0.469	190	0.277	20	0.171	20	0.327	55	0.227	5	0.073
11.00% 2006	930	0.174	2481.5	0.262	610.2	0.450	950.8	0.496	1769.1	0.422	915	0.323
11.90% 2007	5578	0.201	5	0.166	15	0.037	0	0	404.8	0.222	995	0.247
11.40% 2008	0	0	560	0.302	6730.2	0.224	7821.2	0.108	9403.5	0.218	4733.3	0.372
11.99% 2009	1739.9	0.158	150	0.250	116	0.390	56	0.278	1231	0.088	1965	0.079
11.30% 2010	125	0.174	67.5	0.164	0	0	3716.8	0.352	3397.8	0.139	3655	0.198
12.32% 2011	844.4	0.550	45	0.365	10	0.941	65	0.630	535	0.278	197.5	0.300
11.03% 2012	774.3	0.244	0.9	1.040	0	0	398	0.083	4710.0	0.230	5230.9	0.204
10.70% 2020	150.7	0.631	0.5	0.597	1.6	0.724	0	0	0	0	0	0

	Jan-01		Feb-01		Mar-01		Apr-01		May-01		Jun-01	
	Volume	Error	Volume	Error	Volume	Error	Volume	Error	Volume	Error	Volume	Error
11.15% 2002	1815	0.105	885	0.090	1641	0.104	1013	0.180	1510	0.286	1590	0.235
11.10% 2003	588.1	0.138	530	0.143	614	0.145	230	0.227	180	0.299	250	0.328
12.50% 2004	3359	0.070	3715	0.099	4069.6	0.178	2566	0.089	3425.1	0.164	3603.6	0.163
10.20% 2005	720	0.162	625	0.183	385	0.344	490	0.196	660	0.349	535	0.584
11.00% 2006	1040	0.296	815	0.499	827.5	0.587	470	0.553	1275	1.022	810	1.169
11.90% 2007	2505	0.244	2340	0.338	1571.0	0.378	1785	0.302	1772	0.644	1220	0.773
11.40% 2008	8997	0.501	7135.6	1.166	7743.5	1.137	7090.5	1.075	8827.6	1.024	10255	0.858
11.99% 2009	3771	0.236	2729.5	0.196	1070	0.470	1125	0.496	2293	0.751	2560	0.797
11.30% 2010	8177.1	0.408	10649.1	0.874	6053	1.378	2005	1.029	4472.1	0.903	4574.5	0.617
12.32% 2011	725	0.432	367.5	1.237	583	1.555	175	1.728	1545	1.888	2014.2	1.907
11.03% 2012	7321.0	0.335	4544.9	0.713	3297.5	0.889	1497.5	0.766	6237	0.882	7558	0.586
10.70% 2020	51.26	0.406	105.05	0.687	147.8	1.142	40	0.539	193.2	0.281	196	0.223

Together, these findings indicate that, while the extent of trading in a security facilitates price discovery, thereby reducing pricing errors, volume *per se* is an inadequate proxy for capturing liquidity differences in a cross-section of securities, at least in the Indian context. This finding is important since, as mentioned earlier, it is common in the existing literature to use volume as a proxy for liquidity. One plausible explanation for our finding is as follows. Amounts outstanding (issue size) differs significantly across securities; consequently, similar trading volumes imply entirely different degrees of liquidity for bonds with significantly different outstanding amounts. Our finding would suggest that a better proxy for liquidity may be the turnover ratio, the ratio of volumes traded over a period to the amount outstanding in the security.

Explaining the Errors: Importance of Idiosyncratic Factors

Our findings upto this point can be summarized as follows. The term structure estimates using the NS functional form provide a fairly good representation of the average term structure; however, pricing errors exist and, in fact, have increased in the recent period. The size of error is related to the residual maturity of a security and to the extent of secondary market activity it witnesses; however, for comparison across securities, volume is an inadequate proxy for liquidity. Volumes, of course, are an outcome of security-specific attributes that cause securities to be liquid/illiquid relative to each other, and is limited in its ability to provide an estimate of the illiquidity premium for each individual bond, even when it subsumes the impact of all security attributes for traded securities. Our next objective is to identify these attributes and gauge their quantitative importance in terms of their ability to explain inter-security variation in pricing errors.

Following existing empirical literature, we identify residual maturity, time since issuance (age), issue size and current yield (ratio of coupon to price) as attributes determining liquidity. To gauge the quantitative importance of these security-specific attributes, we run an OLS regression with these factors as independent variables. The estimated term structure represents average pricing; therefore securities that are characterised by average liquidity are priced close to the term structure. To determine the expected sign on these variables, let us now consider how each of these factors would influence the pricing of securities that are priced above – and are therefore relatively more liquid. It is reasonable to assume that investors would prefer short-medium bonds relative to long-maturity ones, a possible reason being the greater uncertainty related to long term interest rates. So lower is the maturity, more preferred would be the bond. This would raise its price above that predicted by the term structure and result in high positive errors. Likewise, bonds that have been in existence for a long time may not evince high investor interest, one possible reason being that their

coupons could be out of sync with the current interest rate structure. Independent of this too, investors might prefer to hold relatively newer bonds in their portfolio. Consequently, more recently issued is the bond, higher would be its price relative to that predicted by the term structure. To the extent amount outstanding (issue size) creates floating stock, larger is the issue size, more liquid would be a security and higher its pricing error.

Now consider securities that are priced below the term structure. Greater the residual maturity and time since issuance, more illiquid will be the security and lower will be its observed price relative to that predicted by the term structure. On the other hand, higher is the issue size, relatively more liquid will be the bond, so relatively closer will be its price to the term structure and lower the pricing error.

Following from the preceding discussion, the expected signs on residual maturity and time since issuance are negative, while that on issue size is positive, for both positive and negative errors. The dependent variable is therefore the first-stage errors from the term structure estimation. The independent variables are residual maturity (ttm) and time since issuance (age) and issue size (isz). We include square of age and residual maturity as additional explanatory variables to capture non-linearities in their relationship with pricing errors. Further, since the influence of maturity (age) is expected to differ with the value of age (maturity), we also include an interaction term between these variables[13].

Another variable that usually appears as an explanatory variable in regressions of this type is the current yield (ESU [1998]). The coupon effect normally manifests in terms of investor preferences in favour of high-coupon instruments in a regime with lower tax rates on interest income relative to capital gains (McCulloch [1975]). In the Indian context where the performance of investment managers is judged on the basis of the appreciation/depreciation of portfolios, the differential tax aspect is relatively less important. However, as coupon influences price volatility of a security, current yield may still be important in influencing pricing. We do not however have a prior on the sign of this coefficient.

The pricing errors are categorised into 4 maturity buckets defined as "<1 year", "1-5 years", "5-10 years" and ">10 years". This categorisation serves two purposes. First, it is reasonable to assume that the relationship between maturity and liquidity is not monotonic. Investors, for instance, may be relatively indifferent between securities of 2-year and 4-year residual maturity (say), but reveal a monotonically declining preference thereafter. While the precise nature of such a relationship is difficult to characterise, the categorisation into near-

[13] Diaz & Skinner [2001] use the ratio of age to original maturity as an explanatory variable to capture a similar effect.

homogenous maturity buckets enable us to purge the estimation of mis-specification problems on this count. Secondly, the ability of any chosen functional form to interpolate between and extrapolate beyond available maturity points depends on the scatter of the observed maturity points. While there is no obvious pattern in the pricing errors that could be attributable to the choice of the NS functional form, by categorising the data into maturity buckets, we are minimising the possible distortion that may arise in cross-sectional comparisons.

Table 8 presents the results of our exercise for the full sample period and for 3 sub-sample periods. Important findings that emerge are as follows:

i. An average of 20% of variation in errors is explained by the chosen specification. This compares favourably with the 8-18% explanatory power of Legendre polynomials in ESU [1998].

ii. Barring few exceptions, the signs of the coefficients on residual maturity, age and issue size are of the expected sign. Also, the coefficients are significant in most cases.

iii. Across maturities, there has been a consistent increase in explanatory power of the variables over the 3 sub-sample periods. This would corroborate our hypothesis that an increase in pricing errors in recent periods is on account of an increase in the impact of non-present-value factors.

Table 8: Gauging the importance of idiosyncratic factors

	Jan-00 to Jun-01		Jan-00 to Jun-00		Jul-00 to Dec-00		Jan-01-Jun-01	
	Co-eff	**t-stat**	**Co-eff**	**t-stat**	**Co-eff**	**t-stat**	**Co-eff**	**t-stat**
Maturity less than 1 year								
intercept	0.08	8.09	0.05	1.24	0.06	6.98	0.07	7.82
current yield	-0.75	-7.22	-0.51	-1.72	-0.46	-4.61	-1.28	-8.86
age	0.01	2.73	0.04	2.88	0.01	1.24	0.02	2.09
ttm	-0.12	-2.94	0.46	2.49	-0.16	-4.70	-0.03	-0.86
isz	0.00	-3.77	0.00	-4.59	0.00	-0.05	0.00	-2.40
age-sq	0.00	-1.83	0.00	-2.30	0.00	1.46	0.00	-0.59
ttm-sq	0.16	4.39	-0.28	-1.74	0.19	6.55	0.04	1.50
age * ttm	-0.02	-5.30	-0.07	-3.42	-0.03	-6.02	-0.03	-8.21
R-sq	0.15		0.11		0.33		0.49	
R-bar-sq	0.15		0.11		0.33		0.48	
Nobs	3614		853		1314		1447	
Maturity between 1 and 5 years								
intercept	1.87	24.21	0.65	5.71	1.27	7.38	2.93	20.77
current yield	-17.17	-24.02	-4.86	-4.73	-13.51	-7.36	-29.41	-22.12
age	0.03	3.32	-0.02	-2.66	0.05	2.36	0.11	6.77
ttm	-0.28	-14.68	-0.17	-7.08	-0.16	-3.34	-0.40	-11.92
isz	0.00	15.64	0.00	4.31	0.00	7.72	0.00	15.21
age-sq	0.00	-4.86	0.00	2.81	0.00	-2.50	-0.01	-7.71
ttm-sq	0.06	21.63	0.04	9.53	0.05	6.70	0.10	19.34
age * ttm	0.00	-2.64	0.00	-2.33	-0.01	-2.75	-0.01	-3.98
R-sq	0.36		0.23		0.24		0.59	
R-bar-sq	0.36		0.23		0.24		0.59	
Nobs	5521		2016		1545		1960	
Maturity between 5 and 10 years								
intercept	5.26	17.57	3.59	12.35	4.96	17.75	8.73	14.08
current yield	-31.00	-14.94	-25.80	-11.65	-40.01	-16.22	-51.91	-9.86
age	-0.45	-25.08	-0.10	-6.47	-0.20	-12.40	-0.85	-20.29
ttm	-0.27	-3.99	-0.14	-2.33	-0.04	-0.62	-0.52	-4.42
isz	0.00	8.82	0.00	1.10	0.00	6.81	0.01	13.50
age-sq	0.02	30.47	0.00	3.08	0.01	10.46	0.04	26.69
ttm-sq	0.01	1.47	0.01	1.71	0.00	-1.19	0.02	2.13
age * ttm	0.02	9.91	0.01	7.26	0.01	9.12	0.03	7.54
R-sq	0.28		0.19		0.43		0.48	
R-bar-sq	0.28		0.19		0.43		0.47	
Nobs	5651		1707		1398		2546	
Maturity greater than 10 years								
intercept	5.73	17.14	2.42	5.02	1.93	1.85	7.28	15.11
current yield	-30.36	-13.98	-17.13	-5.02	-15.82	-2.07	-41.70	-10.76
age	-0.26	-13.40	0.11	5.63	-0.01	-0.18	-0.28	-9.08
ttm	-0.15	-4.21	-0.06	-1.50	-0.03	-0.31	-0.17	-3.53
isz	-0.01	-19.39	0.00	-6.15	0.00	-3.78	-0.01	-13.68
age-sq	0.02	26.75	0.01	9.30	0.02	8.39	0.03	29.53
ttm-sq	0.00	2.17	0.00	1.98	0.00	1.07	0.00	1.31
age * ttm	-0.01	-5.99	-0.02	-12.20	-0.02	-4.45	-0.02	-9.80
R-sq	0.38		0.25		0.31		0.60	
R-bar-sq	0.38		0.25		0.30		0.60	
Nobs	3698		1186		421		2091	

Contd.....

In summary, these attributes explain significant variation in pricing errors across securities, implying that they do a good job of capturing inter-security liquidity differences. As against this, regressions using traded volume as the sole independent variable have an explanatory power of just about 10% (results available with authors). This is again a significant finding and is in sharp contrast to EG [1998] who find that, in the US context, volume adequately captures the impact of all security-specific attributes. Rather, our findings indicate that, in the Indian context, analysis and quantification of the liquidity premium should involve modeling structural features like age, time to maturity and issue size using a functional form that best explains pricing errors.

Conclusion

Research on term structure estimation has in recent years shifted from the specification and testing of alternative functional forms of the pricing equation to the analysis of factors over and above the present value relation that determine the pricing of a bond. Many studies have accordingly examined the importance of various security specific attributes that influence investor preferences, and consequently prices, for particular bonds. The present exercise contributes to the existing empirical literature with an analysis of the Government of India bond market. We find that significant pricing errors exist when Government bonds are priced using the term structure alone, a fact that has been well documented in earlier empirical studies for US and Japan and that could lead to distortions in the term structure estimation. Residual maturity, time since issuance, current yield and issue size are identified as security-specific attributes that account for most of these pricing discrepancies.

In terms of future work, our findings point out the need for estimating the term structure of interest rates jointly with a liquidity function that relates the idiosyncratic factors to bond prices. While EG [1998] specifies a relation through which interest rates and idiosyncratic factors affect bond prices simultaneously, Subramanian [2001] attempts to control for the impact of idiosyncratic factors through volume and number of trades based weights in the estimation of the term structure. Neither of these approaches, however, recognises the one-sided nature of the effect of illiquidity (and the factors affecting it), on bond prices. As a result, one cannot interpret the estimated term structure as the benchmark yield curve depicting the schedule of risk-free rates. To model the term structure of interest rates and the liquidity function jointly, the statistical estimation framework must explicitly recognise the one-sided nature of a component of the pricing relation. Darbha [2001] proposes a framework using the stochastic frontier function approach to estimate the term structure and illiquidity premia jointly from bond prices.

References

Amihud, Y. and H. Mendelson; "Liquidity, Maturity, and the Yields on U.S. Treasury Securities", Journal of Finance, 46 (September 1991), pp. 1411-1426.

Bliss, Robert R.; "Testing Term Structure Estimation Methods", Federal Reserve Bank of Atlanta Working Paper 96-12a, 1996.

Bolder, D. & D. Streliski; "Yield Curve Modelling at the Bank of Canada", Bank of Canada Working Paper, 1999.

Carleton, Willard T. and Ian A. Cooper; "Estimation and Uses of the Term Structure of Interest Rates", Journal of Finance 4 (1976), pp. 1067-1083.

Darbha, G.; "Estimating the Benchmark Yield Curve: a New Approach Using Stochastic Frontier Functions", NSE, 2001.

——————, S. Dutta Roy & V. Pawaskar; "The Term Structure of Interest Rates for India: the NSE Zero Curve", NSE, 2000.

Diaz, Antonio and Frank S. Skinner; "Estimating Corporate Yield Curves", Journal of Fixed Income, 2 (September 2001), pp. 95-103.

Elton, Edwin J. & T. Clifton Green; "Tax and Liquidity Effects in Pricing Government Bonds", Journal of Finance, 5 (1998), pp. 1532-1562.

Eom, Young H, Marti G. Subrahmanyam and Jun Uno; "Coupon Effects and the Pricing of Japanese Government Bonds: an Empirical Analysis", Journal of Fixed Income, (September 1998), pp. 69-85.

McCulloch; "Measuring the Term Structure of Interest Rates", Journal of Business, (January 1971), pp. 19-31.

——————; "The Tax Adjusted Yield Curve", Journal of Finance, 3 (1975), pp. 811- 830.

Nelson, Charles R. and Andrew F. Siegel; "Parsimonious Modeling of Yield Curves", Journal of Business, 4 (1987), pp. 473-489.

Robichek, Alexander A. & W. David Neibuhr; "Tax-induced Bias in Reported Treasury Yields", Journal of Finance, (1970), pp. 1081-1090.

Schaefer, Stephen M.; "Measuring a Tax-specific Term Structure of Interest Rates in the Market for British Government Securities", The Economic Journal, 91 (June 1981), pp. 415-438.

Schoenberg, R.; 'Constrained Optimisation – Application Module', GAUSS Applications, Aptech Systems Inc., 1998.

Seber, G.A.F and C.J. Wild; 'Nonlinear Regression', Wiley Series in Probability and Mathematical Statistics, John Wiley & Sons Inc., 1989.

Subramanian, K.V; "Term Structure Estimation in Illiquid Markets", Journal of Fixed Income, (June 2001), pp. 77-86.

Svensson, Lars E.O; "Estimating and Interpreting Forward Interest Rates: Sweden 1992-1994", NBER WP 4871, 1994.

Thomas, S. & V. Saple; "Estimating the Term Structure of Interest Rates in India", unpublished paper, IGIDR, 2000.