Estimating Credit Spreads using Constraints and Penalties

By: Sadanand Tutakne

Table of Contents

Abstract

Credit spreads refer to the difference in interest rates charged by borrowers to cover for default-related risks involved with any loan (or, any set of cash flows). The concept is relevant in the context of interest rate risk reporting too, because as per the US FAS, it is best to include a credit adjustment term when valuing/pricing any given set of cash flows. FAS guidelines allow this adjustment to be made either to the cash flows themselves (numerator), or to the discount rates used to calculate the present value of the cash flows (denominator). The latter is probably an easier approach and in this paper, we focus on the problem of adjusting discount rates for default risk. The paper describes several alternative approaches, including results from some hands-on exercises with Bloomberg data.

Acknowledgments

I would like to thank our customers, Sid Reddy, Kunal Sengupta, Kaustav Mukherjee, Somnath Mukherjee, Kaushik Das and Sushant Vimal for helpful discussions on the broad area of yield curve modeling and Treasury Analytics.

Introduction

Credit spreads refer to the difference in interest rates charged by borrowers to cover for additional default-related risks involved with any loan (or, any set of cash flows). The concept is relevant in the context of interest rate risk reporting too, because as per the US FAS, it is best to include a credit adjustment term when valuing/pricing any given set of cash flows. FAS guidelines (Concept 7, paragraphs 78-88) allow this adjustment to be made either to the cash flows themselves (numerator), or to the discount rates used to calculate the present value of the cash flows (denominator). In this paper, we focus on the problem of adjusting discount rates for default risk.

A higher or a lower discount rate would imply a different discounted present value for any given set of cash flows, and would therefore affect valuation and re-pricing results directly. Re-pricing schedules are a necessary part of any interest rate risk reporting structure, and therefore, it is desirable to have reasonable and accurate discount rates for these reporting tasks. This paper describes some key alternatives available, discusses conceptual and estimation-related issues and provides results for some based on Bloomberg data.

Alternative Approaches to Credit Spreads

The most common formula for credit spreads found in several freely available papers (in the context of zero-coupon bonds) is:

(1) Credit Spread = (-1/T) * ln(1 - PD*LGD),

where PD denotes the probability of default, LGD denotes the loss rate given default and T is the time to maturity of the zero-coupon bond. The main point to keep in mind in using this formula is that using “physical” estimates of PD and LGD (i.e., from default / loss prediction models) typically under-estimates actual market spreads. The reason is that market prices are determined by demand and supply and risk-averse investors typically demand premiums which are significantly more than what a risk-neutral investor (or even an investor with a standard utility function with some risk aversion built into it) would charge. Another reason mentioned in this context is the presence of other risks like liquidity risks. For the purposes of valuation and pricing, therefore, what we need is not the PD (or, LGD) from default and loss-prediction models but the “market-implied PD” (or LGD). It is commonly acknowledged in the financial modeling literature that market-implied PDs (or, risk-neutral PDs as they are called) are typically several times larger than the corresponding “physical” PDs (from statistical default prediction models), for the reasons mentioned above. The estimation of these market-implied PDs (or, risk-neutral PDs) from market data like bond prices and CDS spreads can be undertaken in several alternative ways, which have been taken up in a separate white paper. In fact, one of the methods involves estimating spreads from market data, and then applying standard formulae to extract risk-neutral PDs from them – given some reasonable assumptions on LGD. Of course, implied PDs can be derived directly from market data too (using reasonable assumptions on LGD). So while implied PDs can themselves be extracted from spreads, so can spreads be derived from market-implied (risk-neutral) PDs. Once we have these risk-neutral PDs from market data and some reasonable assumptions on LGD, we can use the above formula to derive spreads. This is one approach and is taken up in a separate white paper.

In closing the above discussion on deriving spreads from risk-neutral PDs, let us keep in mind that since there are several alternative ways of estimating risk-neutral PDs, each is also potentially a way to generate credit spreads too. These alternatives include - besides regression based approaches – approaches based on **structural credit models** and approaches based on calibrating **transition matrices** to market data too. In this paper, we outline the idea behind the transition matrix approach, but do not present results since these programs are still work in progress.

Alternatively, we could go back to first principles and recall that credit spreads are, by definition, simply the difference between risky and risk-free zero-coupon yield curves for different maturities. Therefore, if rating-specific zero-coupon yield curves are available, then simply subtracting one from the other provides us with the credit spread. On the other hand, if zero coupon yield curves for different credit grades are not available, then again, we could estimate **rating-specific yield curves** from market data, and then calculate the spreads. This is one approach attempted in this paper.

Given the similarity of the term structure estimation problem for spreads to the term structure estimation problem for risk-free instruments, it might be obvious that another method which can be used is a model of the "additional discount factor" required due to credit risk. The credit risk discount factor can then be estimated by any good model of the discount factor, just as the risk-free discount factor can. Examples of such specifications would be, e.g., cubic splines with constraints and curvature penalties, exponential splines, higher-order or mean-reverting polynomials, etc. In this paper, this is another "basic" approach that we attempt to study. In this approach, market prices are regressed on discounted cash flows multiplied by an additional discount factor to account for credit risk. The original discounting is done using a known risk-free yield curve, which is derived from a model for risk-free zero-coupon curves. The second discount factor is the credit discount factor, which is estimated via linear or nonlinear regression, as appropriate.

The main reason to explore several alternatives is that different alternatives can give results which are different in subtle ways, even if similar on the average. These additional considerations (e.g., how do the results for low maturities look like) are typically very important in choosing one method (or, model) over the other, as long as other major results are similar across methods (or, models).

Data and Estimation

We describe here, three approaches we have worked on using market data from Bloomberg.

1. Zero-Coupon Yield Curves for Risky Instruments:

This is perhaps the simplest approach - once we are comfortable with procedures to estimate zero-coupon yield curves for risk-free instruments (like treasury bills for the US). The idea is to use the same methodologies, but with data for risky instruments (from one particular risk grade) instead of data from risk-free instruments. The resulting zero-coupon curve will be, by definition, a risky zero-coupon curve and the difference between this curve and the risk-free zero-coupon curve will be the estimate of the credit spread applicable to that particular rating category. Repeating the exercise for each risk grade will provide us with an estimate of the average spread applicable to each rating grade.

The potential caveats of this method are similar to those involved in the modeling of risk-free yield curves. For one, the zero-coupon curve might have spikes at low maturities, might result in unintuitive (say, negative) spot or forward rates, or the resulting forward curves might display some curvature (jumps, spikes, etc.) which is difficult to defend. In addition to that, there is always the possibility that the models do not fit as well as expected in out-of-sample data.

The first of these problems can be dealt with in more than one way. One way is to use parametric forward rate functions (e.g., the Nelson-Siegel model or Svensson's extension of the same) which typically lead to smooth forward and spot curves and typically do not involve negative spots. The potential issue with this is that these models require non-linear estimation, which makes them more difficult to automate, because they typically involve trying different starting values and also require more time to run. A second approach would be to use linearized versions of the same parametric models, but before using these, we should ideally test whether the models perform similarly after linearization. A third approach to the problem is to use a cubic B-splines (or, say, exponential splines) model for the discount function, but add constraints (including non-negativity constraints) and curvature penalties which will work for typical data sets. These models can be expected to take lesser time to run and would not involve iterations on starting values. This third approach is attempted here. When exponential splines are used, we optimize over the additional exponential parameter using a grid search.

The other problem is that of out-of-sample performance. This is common to most models, and in most cases, there is no direct solution to poor out-of-sample performance other than understanding the errors and trying several corrective strategies, including changing the data and methodology. Periodic re-estimation is often advised in the context of yield curves, but this solves our problem only if the re-estimated models work significantly better than the older ones. Knowing error margins could help in some cases, but large error margins would make the models less useful for practical purposes. Therefore, the advice in this case would simply be to understand prediction errors and try alternative solutions to check what works in a particular context.

In general, yield curve models typically try to minimize price-errors, i.e., the distance between observed and predicted prices. For the penalized, constrained cubic B-splines case, the model is:

(1) Discount Factor = (B-Spline Matrix given Knots) * Spline Coefficients

(Tx1) (T x k) (k x 1)

Therefore,

(2) Price = (Cash Flow Matrix) * (B-Spline Matrix given Knots) * Spline Coefficients + error term,

(n x 1) = (n x T) (T x k) (k x 1) (nx1)

where T refers to the total number of cash flow dates for the bonds (or other instruments) used in the exercise together, n refers to the number of bonds used in the exercise and k refers to the number of spline coefficients. The number of columns in the spline matrix is given by k (=d1+d2+1), where d1 refers to the number of knots, d2 refers to the degree of spline specified (say, cubic, quadratic, etc.) and 1 gets added when splines "with intercept" are specified. The spline "intercept" column is not the same as an intercept in a regression and so should not be confused with that.

When curvature penalties and constraints are both applied, the problem changes from a simple OLS estimation problem to:

$$\text{Minimize}\,(Y - X\beta)'\,W\,(Y - X\beta) + \sum_i \lambda_i \beta'\,S_i\,\beta$$
$$\text{subject to}\ (i)\ C\beta = c\ \ \text{and}\ \ (ii)\ A\beta \succ b,$$

where C and c are appropriately sized matrices describing the equality constraints and A and b are matrices describing the inequality constraints. Y refers to the target variable, W refers to the matrix of weights and X refers to the matrix of explanatory variables. Curvature penalties are defined by constructing S(i) matrices appropriately. In our exercise, we use only one S matrix, which is derived from the matrix of second derivatives of the spline matrix. To be precise, if D is the matrix of second derivatives of the splines, we use D'D as the S(1) matrix. Thus, we penalize excessive curvature, as measured by the sum of the squared second derivatives of the spline matrix, while trying to fit the model to bond prices. This makes our solution similar to a smoothing splines solution. The problem is solved using the "pcls" function in the R package "mgcv."

Where curvature penalties are not required, estimation is done using the "lsei" function in the R package "limSolve." The optimization problem in this case is very similar to the problem above – except that there are no penalties and the inequality constraint is a weak inequality. Mathematically, the problem then reduces to:

$$\text{Minimize}\,(Y - X\beta)'\,W\,(Y - X\beta)$$
$$\text{subject to}\ (i)\ C\beta = c\ \ \text{and}\ \ (ii)\ A\beta \geq b,$$

where the notation is the same as for the previous problem.

The above discussion does not imply that no other model of the discount function gives acceptable results, or that cubic splines are the only way to model discount functions. There are several other ways, including parametric models of the forward curve (as in the Svensson model), exponential splines, etc. **Exponential spline** formulations of the discount function have a structure similar to the cubic B-splines formulation above, except that the splines are of the form exp(-k*alpha*t), where k=1,2,3, etc., alpha is an additional parameter typically chosen using grid searches and "t" denotes cash flow dates. The model therefore changes slightly to:

(1) Price = (Cash Flow Matrix) * (Exponential Spline Matrix) * Spline Coefficients + error term,

(n x 1) = (n x T) (T x k) (k x 1) (nx1)

2. Modeling the Term Structure of Credit Spreads:

This approach is different from the previous approach in that here we start with a matrix of discounted cash flows, i.e., cash flows already discounted by a previously estimated risk-free discount curve. The model, therefore, changes slightly to:

(1) Additional Credit Discount Factor = (B-Spline Matrix given Knots) * Spline Coefficients

(Tx1) (T x k) (k x 1)

Therefore,

(2) Price = (Discounted Cash Flow Matrix) * (B-Spline Matrix given Knots) * Spline Coefficients + error term,

(n x 1) = (n x T) (T x k) (k x 1) (nx1)

where the dimensions (T, k, n) are the same as for the previous model.

Since we observed a spike in the estimated credit spread curve at very low maturities, curvature penalties were imposed on the estimation exercise for this model, using the R packages described in the previous section. As the specification of the problem makes clear, we could use other models like **exponential splines** or the **Svensson** model for the risky discount function – it is not essential that we use the cubic B-splines formulation.

3. Credit Spreads using Generators of Transition Matrices:

In this approach, we again start with the basic idea that the price of risky instruments is derived by looking at (a) the price of risk-free instruments of the same maturity, (b) the probability of default (risk-neutral PD, as usual) and (c) the loss rate given default. The difference is that the PD is calculated using a transition matrix approach. In the transition matric approach, the dynamics of ratings over time are given by a transition matrix, say Q. If there are "k" credit "states" with the worst one being default, the dynamics is characterized by a Markov Chain as follows:

$$\begin{pmatrix} y_{1t} \\ y_{2t} \\ \cdots \\ y_{kt} \end{pmatrix} = \begin{pmatrix} q_{11} & q_{12} & \cdots & q_{1k} \\ q_{21} & q_{22} & \cdots & q_{2k} \\ \cdots & \cdots & \cdots & \cdots \\ q_{k1} & q_{k2} & \cdots & q_{kk} \end{pmatrix} \begin{pmatrix} y_{1,t-1} \\ y_{2,t-1} \\ \cdots \\ y_{k,t-1} \end{pmatrix},$$

where y(jt) represents the number of accounts / customers in rating category "j" at time "t" and q(ij) represents the probability of transiting from rating class "j" to rating class "i" between the two periods. The rating categories are ordered and the last (k-th) is the default category and is also an "absorbing" state, in the sense that accounts once in default are typically assumed to stay in default for the relevant time horizon. Therefore, q(kk)=1 and q(ik)=0 for all i.

Since transition matrices are calculated for a given horizon (say, 1 year), it is helpful to derive the generator of the transition matrix, so that the transition dynamics for any time horizon can be characterized (say, 1.2 years). Not all matrices have generators, but in most practical examples, either this problem is not observed or there are workarounds available from the literature, due to which the technique is useable. A time-invariant generator is a matrix (say, lambda), such that

$$Q(t,T) = \exp[\Lambda\,(T-t)].$$

Once lambda is estimated, the transition matrix for any horizon (T-t) can be estimated using the above formula. The last row of the transition matrix (or, last column, depending upon how the matrix is defined), gives the probabilities of default. The implied price of the risky cash flow due at time T (looked at from time point "t") is then given by:

$$v(t,T) = B(t,T)\,[1 - q(t,T) + q(t,T)\delta],$$

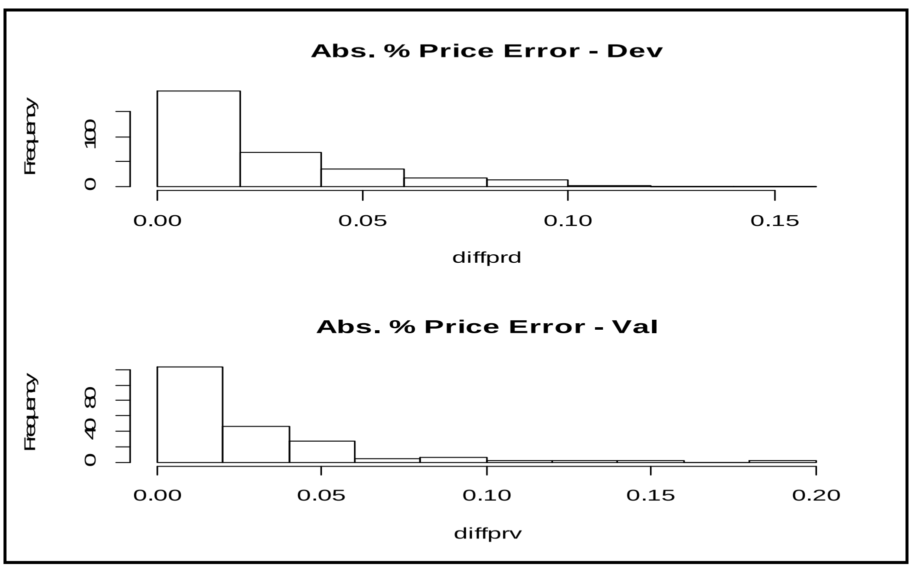
where "delta" is the recovery rate applicable to time "t". To calibrate the implied prices to actual market prices, we use non-linear regression to minimize the squared distance between observed and model-implied prices, with the elements of the generator matrix being the parameters to be estimated. While it is possible to simply use generators of actual transition matrices, it is likely that the PDs would under-estimate market prices (since physical PDs typically do). Therefore, it is advisable to calibrate implied prices to actual prices. To keep results relatively close to actual transition matrices, one can impose penalties on large deviations of estimated lambdas from the actual values. This is the approach taken by Arvanitis et. al. (1999). Since this approach is still work-in-progress, we do not present any results for this approach here.

Results: Spreads and Adjusted Cash Flows

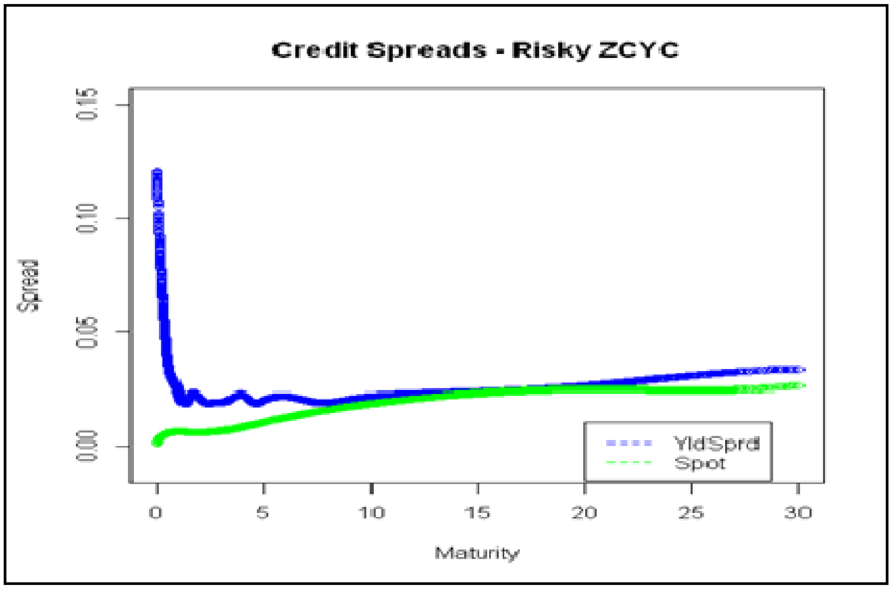
The main outputs we need from spreads models are (a) the spreads themselves, (b) the implied PDs, given the assumption on LGDs and (c) the resulting credit-adjusted cash flows, which provide us with an alternative way to value risky assets. In evaluating one method over another, the reasonableness of these results would be one factor, but another basic factor would be the fit of the models in in-sample and out-of-sample data. Since results for implied PDs and the resulting expected cash flows are taken up in a separate white paper, here we focus mainly on deriving the credit spreads using some of the methods above.

1. Results from Risky Zero-Coupon Yield Curves:

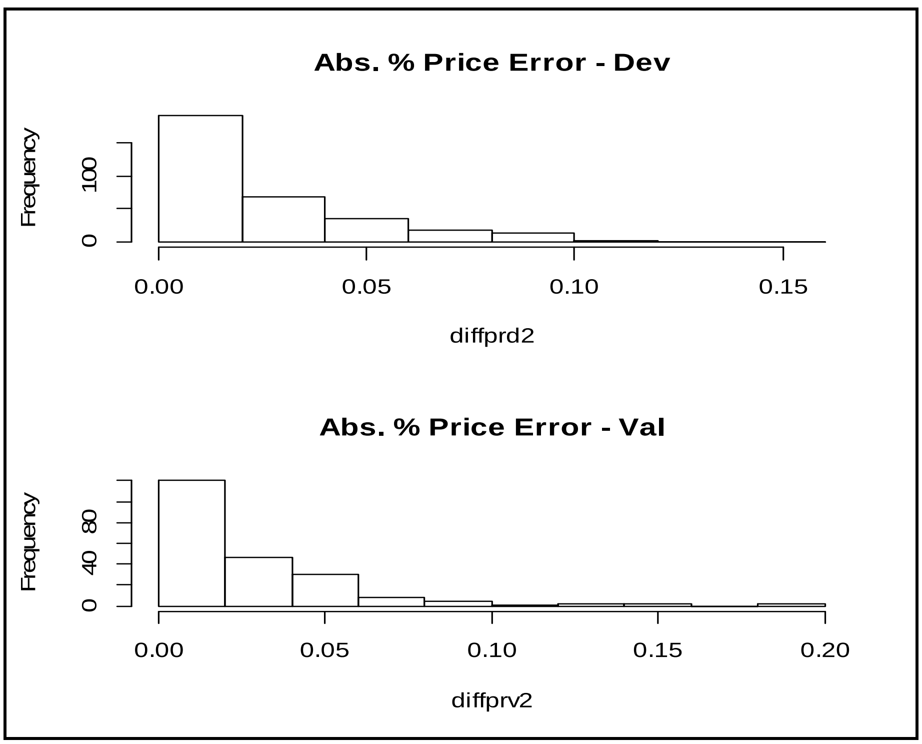
(a) The median absolute percentage price errors in development and validation data were 1.63% and 1.53% respectively, and the distributions of the errors were as follows, using constraints but no curvature penalties.

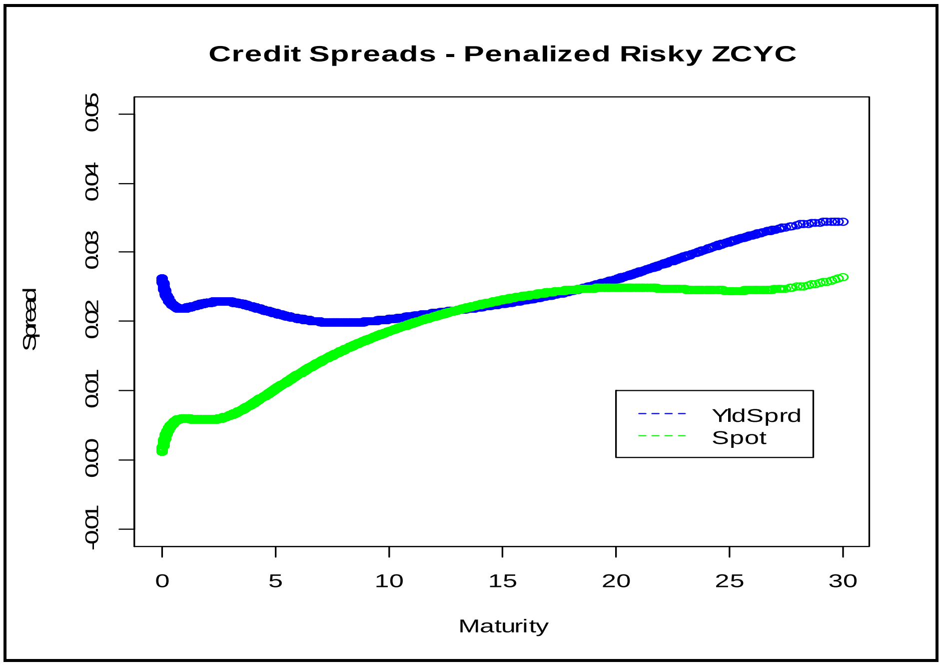


The resulting credit spreads are plotted below along with the risk-free discount curve, and the spike in the credit spread at lower maturities is evident (>8% for very low maturities).



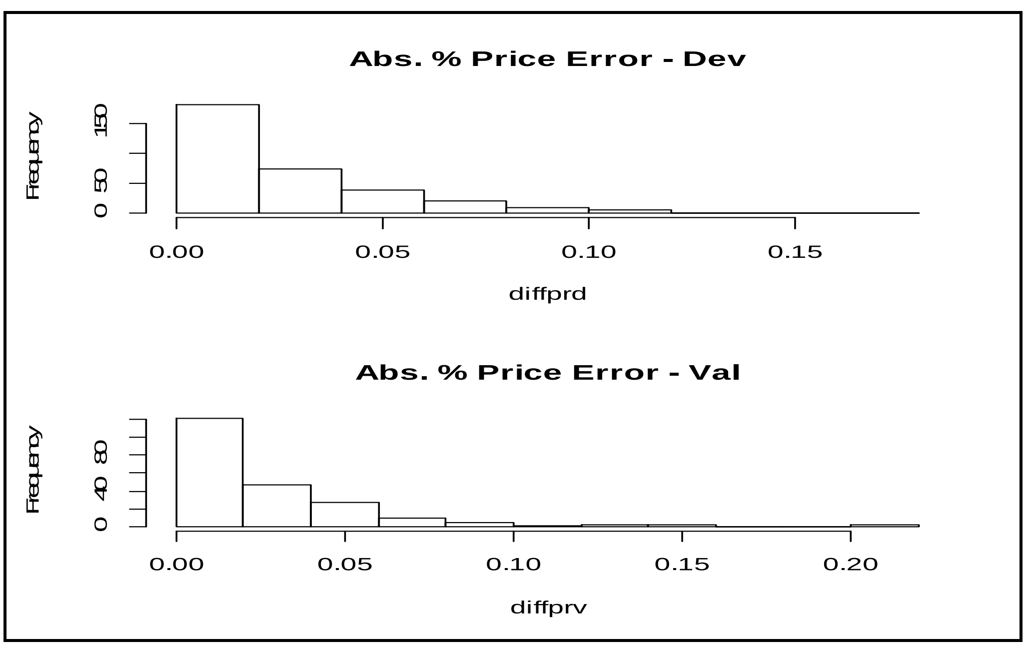
(b) When **curvature penalties** were added to the constrained estimation above, the resulting price errors and credit spreads looked like as follows.

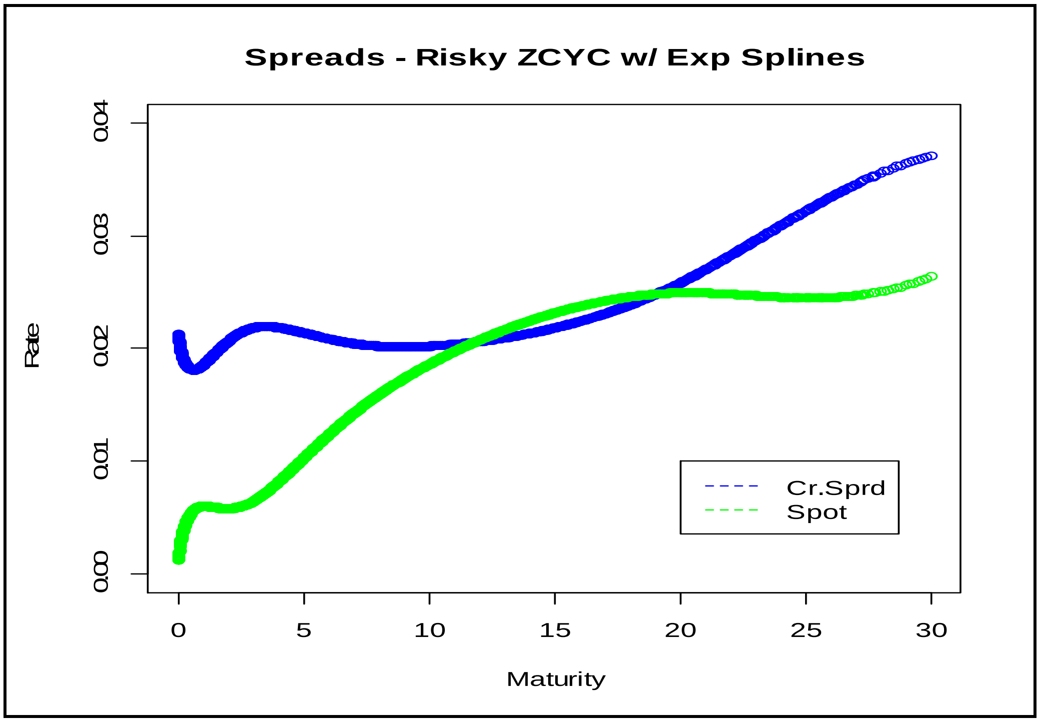




As we can see, **the huge spike at lower maturities gets corrected** when curvature penalties are applied to the risky zero-coupon curve. Since the risky curve does not show a spike (and neither does the risk-free curve we have estimated elsewhere), the spreads also do not show a spike anymore.

(c) In the above, the risky zero curve was modeled using cubic splines, with boundary constraints and with or without curvature penalties. As mentioned earlier, this is not necessary – other models of the risky discount curve can be used too. For example, using an exponential spline model (with alpha optimized using a simple grid search), the median absolute price errors turned out to be 1.67% and 1.68% respectively. The distribution of the absolute percentage price errors and the resulting credit spreads looked like as follows.

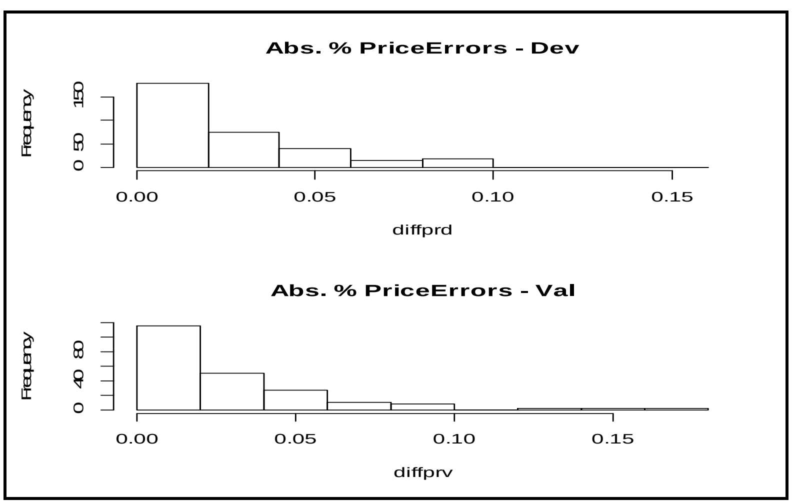


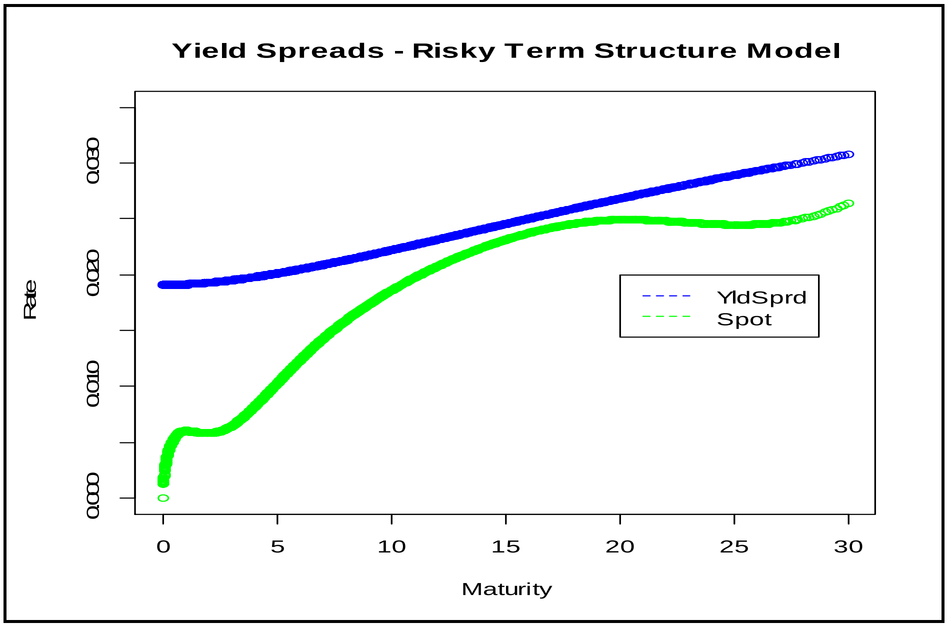


2. Results from Models of the Risk Term Structure:

Here we used an exponential splines model again, but this time for the risky discount curve. Therefore, cash flows were discounted using the risk-free rates (estimated elsewhere) before using them in the regressions. The regressions therefore directly provide us with an estimate of the "additional credit discount factor", and thereby, of the credit spreads.

Using this model, the median absolute price errors turned out to be 1.76% and 1.81% respectively. Some further intensive grid search on alpha could perhaps reduce this, but that is not the focus here, so we move to the other results.





3. A Note on Implied PDs:

The derivation of implied PDs has been discussed at length in a companion paper and is therefore not being repeated here. Suffice it to say that the most standard formula to obtain PDs from spreads can lead to PD estimated lying outside the [0,1] interval, and that is the reason why it is often better to model implied PDs in other ways. The standard formula which appears in several places is:

$$PD_t = \frac{[1 - \exp(-s_t t)]}{LGD},$$

Where PD(t) refers to the cumulative implied PD up to time "t", s(t) denotes the credit spread for maturity "t" and "t" denotes the term to maturity for a zero coupon bond. However, assuming recovery to be of the **"RM-type"**, we can use a different formula which gives implied PDs between 0 and 1. However, as we discuss in the companion paper, underlying the RM formula is a completely different assumption – i.e., the assumption that recoveries happen at the time of default itself and that the total claim of the lender at the time of default is the price of the bond just prior to recovery.

Concluding Remarks

In this paper, we begin with an overview of several ways in which credit spreads have been approached in the literature, and proceed by showing results from two different approaches. One good point about these approaches is that if the resulting spreads show unreasonable "spikes" at, say, lower maturities, then we can correct for these spikes by imposing curvature penalties in addition to the other constraints we might have imposed on the splines used for estimation. In the exponential splines case here, we did not need curvature penalties and inequality constraints, but we did impose the condition that the initial discount factor should equal unity. In other situations, inequality constraints (like the resulting splines must always be >=0) might need to be imposed, and these kinds of constraints and penalties can be imposed quite easily with the functions in the R packages "limSolve" and "mgcv". As for implied PDs and the transition matrix approach to estimating credit spreads, we have made significant progress on these too, but they are not added here because implied PDs have been taken up in a separate paper and the transition matrix approach is still under development.

Sadanand Tutakne
Senior Manager, Financial Services Analytics
GENPACT
e-mail: Sadanand.tutakne@genpact.com

About the Author

Sadanand Tutakne, Senior Manager, Financial Services Analytics, GENPACT, has about 11 years of corporate experience in behavior and credit scoring, loss forecasting, credit modeling for Basel and econometric modeling for Treasury models, e.g., yield curve modeling, modeling credit spreads and implied PDs and scenarios for yield curves. Prior to joining the corporate sector, he was a doctoral candidate (ABD) at the University of Maryland at College Park, where he worked on vector autoregressive models for quantifying the effects of interest rate policy shocks on macroeconomic variables.

References

1. Arvanitis, Angelo, Jonathon Gregory and Jean-Paul Laurent, "Building Models for Credit Spreads," The Journal of Derivatives 1999.6.3:27-43.
Link: http://www.iijournals.com/doi/abs/10.3905/jod.1999.319117
OR
http://www.ressources-actuarielles.net/ext/isfa/1226.nsf/9c8e3fd4d8874d60c1257052003eced6/7c1d935203184237c1257a4f006b127a/$FILE/building_models_for_credit_spreads.pdf

2. Berd, Arthur M., Roy Mashal and Peili Wang, "Estimating Implied Default Probabilities from Credit Bond Prices," Fixed Income Quantitative Credit Research, Lehmann Brothers, 7th August 2003.
Link: ftp://ftp.idc.ac.il/Faculty/Roy/Pub/QCRQ_Curves.pdf

3. Bezerianos, George and John Hatgioannides, "An Empirical Investigation of the Theory of Contingent Claims," Cass Business School.
Link: http://www.qass.org.uk/4thAFEQASSConferencePapers/Bezerianos.doc

3. Crosbie, Peter and Jeff Bohn, "MODELING DEFAULT RISK," Moody's KMV, December 18, 2003.
Link: http://solvay.ulb.ac.be/cours/farber/Ofin/Reading%206%20ModelingDefaultRisk.pdf

4. FASB, Original Pronouncements, As Amended, "FASB Concept Statement No. 7: Using Cash Flow Information and Present Value in Accounting Measurements," 2008.

Link:
http://www.fasb.org/cs/BlobServer?blobkey=id&blobwhere=1175820900214&blobheader=application%2Fpdf&blobcol=urldata&blobtable=MungoBlobs

5. Jarrow, R.A., D. Lando, and S.M. Turnbull (1997), "A Markov model for the term structure of credit risk spreads," Review of Financial Studies 10, 481–523.
Link:
http://forum.johnson.cornell.edu/faculty/jarrow/055%20Markov%20Model%20Credit%20Risk%20RFS%201997.pdf

6. Maclachlan, Iain C, "An empirical study of corporate bond pricing with unobserved capital structure dynamics," Munich Personal RePEc Archive (MPRA), The University of Melbourne, May 2007.
Link: http://mpra.ub.uni-muenchen.de/28416/1/MPRA_paper_28416.pdf