Value at Risk for Corporate Bond Portfolios

P.C. VENKATESH

P.C. VENKATESH
is a financial economist with the Comptroller of the Currency in Washington, DC.
pc.venkatesh@occ.treas.gov

Risk measurement systems (particularly daily VaR systems) typically reflect a compromise between the competing objectives of delivering cost-effective and timely, but reasonably accurate, results. Judicious choices with respect to data used and computational aspects can be made to reduce the overall costs and computation time. A natural question then is how these design choices affect the quality of the risk measures produced.

I explore this question in the context of producing a daily value at risk (VaR) for portfolios of credit-risky bonds and loans. Employing historical simulation (HS) as the basic framework, I consider several variants that differ with respect to data or computation time and complexity.

Broadly speaking, the variants tested differ with respect to: 1) number and identity of the risk factors included; and 2) treatment of the risk factors within the calculations, e.g., stochastic properties of the risk factors included or history used.

The basic variant tested relies on a single risk factor derived solely from rating-level indexes. Rating-level indexes are an intuitive and not very data-intensive choice—it is relatively easy to obtain current and historical data, of good quality, for rating-level indexes. The adequacy of such a risk factor depends on the relative importance of idiosyncratic risk (which we expect to vary with, e.g., credit quality or size) in corporate bond portfolios. To accommodate this, I also test a second variant, which augments the first risk factor with a generic idiosyncratic factor.

The second differentiating feature is the treatment of the risk factors within the calculations. Computation costs can be reduced by ignoring data properties that may be time-consuming to estimate or simulate.[1] Examples of such data properties include conditional heteroscedasticity (suggesting a GARCH-type model for the dynamics) and leptokurtosis (suggesting GARCH or mixtures and jumps).[2]

I compare the performance of variants that do and do not incorporate such properties of the risk factor data. Alternative histories are also considered. Finally, the experiments are carried out for test portfolios varying in credit quality (from investment-grade to high-yield) and in size.

Of primary interest are portfolios constructed from individual bonds, with target characteristics (e.g., credit quality, size). But rating-level indexes, which can be viewed as an externally defined large portfolio, provide a valuable complement. Thus, rating-level indexes serve a dual role, as risk factors and as test portfolios.

The principal findings are as follows. Rating-level index returns are found to exhibit time series properties consistent with GARCH-like effects—this is relevant in their role as risk factors. For rating-level indexes as test portfolios, incorporation of these time series properties leads, on balance, to improved

VaR forecasts. This roughly parallels findings elsewhere. For test portfolios constructed from individual bonds, the performance of variants employing a single risk factor (based on rating-level indexes) leaves substantial room for improvement. Incorporating a second generic idiosyncratic risk factor leads to significantly improved VaR performance.

I. OUTLINE OF METHODOLOGY

The methodology is fairly standard and straightforward, consisting of general steps as follows. First, I generate a forecast distribution (tailored to the characteristics of the test portfolio) of possible returns over the next day (or some other target horizon such as the next week). Upper and lower VaR thresholds are typically chosen as the upper and lower tail quantiles of the forecast distribution.[3]

As suggested in the introduction, each methodological variant gives rise to a different forecast distribution, and hence to different VaR thresholds. VaR thresholds are generated for each day (or alternative horizon) over some test period (a calendar year). Depending on the methodological variant, the thresholds may be static or vary dynamically through time.

Second, an exceedance is recorded whenever the test portfolio's realized return (over the target horizon) exceeds the VaR threshold.[4] From the time series of exceedances over the test period, two performance metrics are constructed: the frequency of exceedance (the number of exceedances divided by the total number of comparisons/observations in the test period), and the average exceedance (the sum of the absolute exceedances divided by the total number of observations in the test period).[5] Successive test periods, spanning one calendar year each, from 1997 through 2001, are used in the analysis.

Third, for portfolios constructed from individual bonds, a cross-sectional component is introduced by repeating steps 1 and 2 for numerous test portfolios (with the same common attributes—e.g., size, credit quality). That is, the *same* cross-section of test portfolios is evaluated under the different methodological variants.

The standard binomial test is used to assess whether the observed exceedance frequency is statistically different from the prediction (as per the VaR threshold). The patterns of statistically significant exceedances generated by the different variants are compared. In addition, the relative performance is also easily assessed via the cross-sectional distributions of pairwise matched differences.[6]

Methodological Variants

The variants/alternatives examined differ in: 1) the risk factors included; and 2) the treatment of the risk factors—e.g., the history used, or the stochastic properties assumed.

Risk factors. When only a single risk factor is used, it is constructed from rating-level index values (as reported by Merrill Lynch). To begin with, a rating is imputed to each bond in the test portfolio, according to the bond's yield relative to the rating-level yields at the beginning of the test period. The rating imputation algorithm is very basic; the midpoint of adjacent rating-level yields is the threshold, and bonds with yields higher than the midpoint yield are assigned the higher rating (and vice versa). If the test portfolio includes only bonds of a single common rating, the risk factor is just that of the rating-level index; if not, it is a weighted average composite of the rating-level index values, with weights equal to the fraction of bonds in each rating.[7]

Note that yields are used only for the rating imputation; the core calculations are based on returns (log of price relatives). Specifically, the rating-level index values used are total return index values, reflecting both capital gains (losses) and coupon income. Returns for the bonds in the test portfolios are also based on logs of price-relatives.

In general, test portfolios will exhibit idiosyncratic variability relative to a set of risk factors based on rating-level indexes alone. Therefore, a variant incorporating an idiosyncratic risk factor is developed. The properties of a generic idiosyncratic risk factor are estimated from the idiosyncratic variation of an arbitrary set of several test portfolios (the estimation portfolios). In the testing phase, the forecast distribution for this variant is built from the rating level and the idiosyncratic risk factors together.

Treatment of risk factors—stochastic properties. The basic treatment is to assume that daily rating-level returns are independently and identically distributed (i.i.d.). As shown by other studies and confirmed here, this is only approximately true. Of particular relevance here is the persistence of volatility shocks (e.g., manifested in positive serial correlation of squared returns).

Applying a GARCH-type model to the rating-level series, one can estimate the relevant parameters to generate forecasts of conditional mean and volatility. The residual, or innovation, series from the GARCH model can be handled in various ways, depending on its properties and design criteria. I follow procedures suggested by, e.g., Barone-Adesi and Vosper [2000] and McNeil and

Frey [2000], to incorporate these properties into a historical simulation framework.[8]

Treatment of risk factors—history used. A recurring question in VaR applications is the sample length or history to be used in estimating the stochastic properties of the risk factors. I examine two variants. The basic variant uses all available data, i.e., the entire history. This is a useful experimental alternative even though it appears to presume prescience. The other variant uses the cumulative history from the start of the sample to the beginning of the test period.

Construction of Test Portfolios

To begin with, a price/yield data set for a large pool of bonds representing a variety of issuer and bond characteristics—industry membership, issuer/bond rating, bond maturity, coupon type (fixed versus floating)—is created. Bonds with certain types of embedded options are excluded. Test portfolios are created by sampling from this master data set.

In general, test portfolios can be created either by random sampling or to reflect specific portfolio characteristics (e.g., 80% investment-grade). The main results reported rely on test portfolios constructed on the basis of size and credit quality (imputed rating level) specification.

II. RELATED LITERATURE

The literature on value at risk is vast and growing. A website at www.gloriamundi.org/var/provides a substantial bibliography. D'Vari and Sosa [2000] and Vlaar [2000] investigate the performance of alternative VaR methodological variants. In my parlance, their primary interest is in the incorporation of time series properties into a single risk factor framework; that is, they do not explore the role of idiosyncratic risk.[9]

In the credit risk literature, the research of greatest interest deals with the stochastic properties of rating-level indexes and with idiosyncratic variability.

Pedrosa and Roll [1998] examine daily data (from October 1995 through March 1997) from Bloomberg on a variety of rating-level and sector level spreads. Their main conclusions are: 1) spreads can be assumed to follow a random walk; 2) there is a (latent) common systematic factor as evidenced by cointegration among the different spread series; 3) changes in spreads are fat-tailed and can be modeled as mixtures of normals; and 4) spread changes show GARCH effects.

In the context of portfolio credit risk models, Kiesel, Perraudin, and Taylor [1999] study daily (April 1991–November 1998) data from Bloomberg on rating-level spreads. They characterize high credit quality spreads as incorporating a mean-reverting component that dies out, while low credit quality spreads are like random walks. Thus, spread volatilities vary by rating categories depending on the holding period—note, however, that these effects are more relevant for the longer horizons typically used in portfolio credit risk models.

Kim [2000] studies credit migration correlations based on rating, and concludes that, for short risk horizons (one day to two weeks), correlations between obligors and correlations between credit events and systematic market factors are negligible. This suggests there is a highly idiosyncratic component in rating transitions.

Collin-Dufresne and Martin [2000] use monthly data on individual corporate bonds and seek to explain the time variation of their spreads. They find that theoretically motivated firm-specific variables account for only about 25% of the variation. Running a principal components analysis on the residuals, they find that the first component accounts for 58% of the variance—they interpret this component as a common systematic factor.[10]

III. DATA

It is worth reiterating that a distinguishing feature of my research is that I use daily data on individual bonds. Much of the empirical work on credit issues has relied on either monthly data on individual bonds, or daily data on aggregates (e.g., rating or industry sector).

Sample Selection

The master sample of individual bonds is created using Bloomberg's built-in search routines, and successively applying various filters. The final sample consists of fixed-rate dollar-denominated bonds issued by U.S. companies.

Most types of bonds, notes, and debentures (senior/junior; unsubordinated/subordinated) are included. Examples of excluded types are government- or locally guaranteed bonds, mortgages (and related instruments), asset-backed instruments, and collateralized equipment trusts. Bonds with call and sinking fund provisions are retained, while bonds with other types of explicit options (e.g., convertible, puttable) are excluded.

To avoid overrepresentation of heavy issuers, I include a maximum of two bonds per issuer, based on remaining maturity—the shortest and the longest.[11]

Price and Yield Data

Daily price and yield data are downloaded using C programs via *Open Bloomberg*. The sample spans January 2, 1997, through February 25, 2002, but data for individual bonds will vary depending on, e.g., date of issue, default.

Initial data checks. By default, the downloaded data are so-called Bloomberg fair value (hereafter, BFV) prices, which are roughly akin to matrix prices.[12] The quality of these data is assessed via spot checks against Bloomberg generic prices (BGN) and, where available, dealer quotes; quality is found to be very good.[13]

Returns. Because bonds trade infrequently, a continuous stream of daily prices is the exception rather than the rule. To rectify this, missing prices are overwritten by the most recent price within the previous ten days.[14]

Returns are computed as logs of price-relatives. Suspicious reversals and spikes are flagged for further scrutiny, and corrected where possible; otherwise they are set as missing. Specifically, for days with returns of higher than 10% ($|r_t = \log(p_t/p_{t-1})| \geq 0.10$), a series of checks are carried out to ensure that this is not caused by erroneous prices at $t-1$ and/or t.[15]

Checking test portfolio returns. Days of absolute portfolio returns higher than 5% are also flagged. The constituent returns are examined for abnormalities, and an attempt is made to see if there were any corroborating news stories.

Index Data

Merrill Lynch makes available on Bloomberg time series data on a variety of government and corporate bond indexes. I downloaded data on rating-level indexes—AAA, AA, A, and BBB, and BB, B, and C (high-yield series)—from January 2, 1997, through February 25, 2002.

IV. STATISTICAL PROPERTIES OF TEST PORTFOLIOS

Rating-level indexes can be viewed as an externally defined large single-rating test portfolio.

Summary Statistics

Exhibit 1 provides basic time series statistics of the index daily returns over the entire sample period (January 2, 1997–December 31, 2001, equaling 1,305 observations). The investment-grade series (AAA, AA, A, and BBB) shows very similar daily return properties over this period. In the non-investment-grade series, there are differences relative to the investment-grade series and among ratings. There is sharply greater (negative) skewness and kurtosis than in the investment-grade series; the standard deviation, however, is actually lower for the BB and the B series. The B and C series are roughly similar to each other, but somewhat different from the BB series.

The results of correlation and principal components (PC) analyses are shown in Exhibit 2. Panel A (correlations) shows that: 1) the AAA, AA, A, and BBB indexes are highly intercorrelated; 2) the B and C indexes are moderately intercorrelated; 3) the B and C indexes are essentially uncorrelated with the AAA, AA, A, and BBB indexes; and 4) the BB index is quite highly correlated with the AAA, AA, A, and BBB indexes but only weakly so with the B and C indexes.

Panel B of Exhibit 2 shows that the first four principal components explain close to 99% of the generalized variance. Examination of the factor loadings (not shown) suggests that the first PC is a risk driver common

EXHIBIT 1
Summary Statistics for Indexes

Series	Median	Mean	S.Dev.	Skewness	Kurtosis	Min.	Max.
AAA	0.034	0.028	0.296	-0.30	4.3	-1.38	1.07
AA	0.033	0.029	0.280	-0.24	4.1	-1.17	1.03
A	0.034	0.028	0.297	-0.31	4.4	-1.29	1.22
BBB	0.032	0.026	0.313	-0.45	5.1	-1.91	1.15
BB	0.037	0.024	0.185	-1.64	15.1	-1.92	0.67
B	0.028	0.007	0.237	-3.25	35.0	-3.14	1.02
C	0.032	-0.010	0.392	-2.47	25.9	-4.92	1.66

EXHIBIT 2
Correlations and Principal Components (PCs) of Indexes

Panel A. Correlations of Daily Returns

	AAA	AA	A	BBB	BB	B	C
AAA	1.0	0.98	0.97	0.96	0.69	0.03	-0.02
AA		1.0	0.99	0.98	0.70	0.02	-0.02
A			1.0	0.99	0.70	0.04	0.00
BBB				1.0	0.71	0.07	0.03
BB					1.0	0.49	0.38
B						1.0	0.68

Panel B. Cumulative Percentage of Generalized Variance Explained by the PCs

PC1	PC2	PC3	PC4
61.06	93.12	97.57	98.98

EXHIBIT 3
Time Series Properties of Indexes

		Index Series						
	Parameter	AAA	AA	A	BBB	BB	B	C
Mean	γ_0	0.036	0.036	0.036	0.041	0.028	0.029	0.010
	γ_1	0.127	0.100	0.095	0.116	0.153	0.461	0.233
	γ_2					0.106	0.088	0.110
	γ_3						0.058	0.070
	γ_4						0.090	0.102
Variance	ω	0.001	0.001	0.001	0.000	0.001	0.001	0.002
	α_1	0.039	0.051	0.033	0.104	0.056	0.199	0.060
	α_2				-0.066		----	
	β_1	0.952	0.957	0.954	0.959	0.901	0.797	0.924

to the investment-grade indexes (AAA, AA, A, and BBB); and the second PC loads heavily on the B and C indexes, positively on the BB, and close to zero on the other indexes. Together, these results suggest that: 1) the set of AAA, AA, A, and BBB indexes has a common component; and 2) the B and C high-yield series have another common component, distinct from the first component.

Time Series Properties

These results provide the general background. Exhibit 3 shows the parameter estimates for capturing serial correlation and GARCH effects in these series, over the entire sample period.[16] The best fits, based on statistical significance of the estimated coefficients and the properties of the residuals, are shown.

The sum of the ARCH/GARCH coefficients in the variance equation is close to unity, consistent with the persistent volatility shocks. With one exception, terms beyond a GARCH (1, 1) are not statistically significant.

From the mean equation, serial correlation is quite mild and restricted to the first order for the investment-grade series (the R^2 are about 0.003, and the Durbin-Watson statistics are very close to 2.0). For the non-investment-grade, however, lags up to the fourth order are significant—the R^2 are 0.05, 0.27, and 0.13 for the BB, B, and C series, respectively (D-W statistics remain close to 2.0). Possible reasons for the extended serial correlation are lagged adjustment or infrequent trading of the constituent bonds.

Tail Quantiles, Year by Year

Tail quantiles are a key measure in this research, so it is useful to get a sense of their general behavior for rating level indexes. Exhibit 4 shows the tail quantiles (the first percentile (*Q01*) and the 99th percentile (*Q99*) of the distribution of daily index returns, compiled separately for each calendar year.

It is evident that the tail quantiles vary a fair amount from year to year. There is also some suggestion that the interyear variation rises with declining credit quality.

As with other properties reviewed so far, the investment-grade indexes behave very similarly to each other, except in 1998. For the non-investment-grade, it is difficult to deduce any clear patterns. The B and C series do show larger downside quantiles in certain years, somewhat confirming the intuition that they are riskier.[17]

The behavior of the BB series is somewhat counterintuitive (at least at first glance), in that its tail quantiles are often smaller than those of even the investment-grade series. Detailed investigation of these minor puzzles is left for future research. (As a precaution, however, I compared the Merrill Lynch BB series to the BB series from (Salomon Smith Barney, and confirmed the integrity of the data.)

Test Portfolios of Individual Bonds

Each bond's first valid yield for a calendar year is used to impute credit quality (using the rating imputation algorithm I have described), and this initial credit quality is used to create the test portfolios. As not every bond may have valid price data for every day, the number of valid contributions to a test portfolio's value may vary from day to day.

The total numbers of candidate bonds available for each calendar year, for each rating group, are shown in

EXHIBIT 4
Tail Quantiles of Returns of Indexes (%)

		Index						
Year	Quantile	AAA	AA	A	BBB	BB	B	C
1997	0.01 (*Q01*)	-0.85	-0.82	-0.82	-0.86	-0.59	-0.51	-0.82
	0.99 (*Q99*)	0.86	0.82	0.79	0.82	0.47	0.24	0.64
1998	0.01	-0.83	-0.80	-1.14	-1.20	-0.73	-1.27	-1.98
	0.99	0.87	0.82	0.82	0.87	0.53	0.45	0.81
1999	0.01	-0.86	-0.79	-0.85	-0.84	-0.50	-0.35	-0.80
	0.99	0.68	0.68	0.76	0.72	0.38	0.27	0.68
2000	0.01	-0.66	-0.60	-0.64	-0.72	-0.42	-0.90	-1.20
	0.99	0.68	0.68	0.68	0.71	0.32	0.75	0.66
2001	0.01	-0.95	-1.01	-1.07	-0.86	-0.56	-0.93	-1.19
	0.99	0.78	0.75	0.87	0.88	0.42	0.86	1.32
1997–2001	0.01	-0.81	-0.74	-0.78	-0.88	-0.56	-0.93	-1.19
	0.99	0.78	0.74	0.75	0.77	0.43	0.61	0.81

EXHIBIT 5
Sample Sizes

	Rating Group					
Year	AAA and AA	A	BBB	BB	B & C	Total
1997	465	211	403	125	135	1236
1998	871	210	222	183	141	1453
1999	435	472	689	293	180	2017
2000	184	342	1069	428	237	2212
2001	844	362	278	162	190	1509

EXHIBIT 6
Tail Quantiles of Returns of Test Portfolios

	Cross-Sect.	AAA and AA		A		BBB		BB		B and C	
Year	Stat.	Q01	Q99	Q01	Q99	Q01	Q99	Q01	Q99	Q01	Q99
1997	Median	-0.66	0.60	-1.05	0.95	-1.38	1.23	-0.91	0.76	-0.74	0.58
	Extreme	-0.81	0.72	-1.15	1.03	-1.51	1.33	-1.06	0.92	-0.90	0.69
1998	Median	-0.81	0.67	-1.61	1.16	-1.46	1.04	-0.94	0.64	-1.49	0.90
	Extreme	-1.00	0.80	-1.82	1.28	-1.68	1.30	-1.23	0.77	-2.34	1.30
1999	Median	-0.55	0.35	-0.94	0.63	-1.30	0.88	-0.96	0.71	-1.40	1.06
	Extreme	-0.68	0.49	-1.20	0.84	-1.58	1.07	-1.39	0.87	-2.94	1.65
2000	Median	-0.66	0.23	-0.45	0.39	-0.91	0.78	-0.81	0.61	-2.03	0.76
	Extreme	-2.40	0.72	-0.70	0.51	-1.25	1.00	-2.27	0.80	-3.52	1.16
2001	Median	-0.76	0.62	-1.49	1.22	-1.09	0.85	-1.40	1.01	-3.35	1.46
	Extreme	-2.05	0.77	-1.77	1.50	-1.85	1.18	-3.14	1.29	-7.32	2.23

Exhibit 5.[18] To be included in a given year, a bond must have been outstanding the previous year.

A random sample of $nSize = 50$ bonds from each cell provides a test portfolio of a given rating for that year. This is repeated 100 times, thus creating 100 different test portfolios, for each year and rating group. For the subsequent analysis, the 100 different test portfolios provide the basis for a cross-sectional analysis, within each year-rating cell.

As with the rating-level indexes, it is helpful to get a sense of the properties of the tail quantiles of the test portfolios. For each year-rating cell, the median and extreme of the cross-sectional distribution of tail quantiles (the smallest value for *Q01* and the largest value for *Q99*) are shown in Exhibit 6.

There are broad similarities to the indexes, in that there is a fair amount of interyear variation and that the variation appears to be greater for lower credit quality portfolios. Variations also appear to be generally on the same order of magnitude, but it is evident that tail quantiles of the test portfolios are larger, in absolute value, than tail quantiles of the indexes for lower credit quality.

One possible interpretation is that the indexes are better diversified than the test portfolios (recall that each test portfolio consists of 50 securities). To put it differently, the level of idiosyncratic risk (relative to the single index-based risk factor) is greater for portfolios of lower credit quality. Cross-sectionally, a comparison of the median to the extreme indicates that the tail quantiles vary considerably across test portfolios of the same credit quality. This cross-sectional variation is also consistent with the presence of idiosyncratic risk.

Finally, it can be noted that the lower tail quantiles (*Q01*) tend to be larger than the upper tail quantiles (*Q99*). That is, the return distributions are left-skewed.

V. VaR ANALYSIS

My narrative description is oriented toward the test portfolios constructed from individual bonds, but each rating-level index can be viewed as an externally defined, large, single-rating test portfolio, and this is how I treat the indexes.

V0 is the basic reference variant. It uses a single risk factor—for single-rating test portfolios (i.e., A, BBB, and BB), the risk factor is just the index return of the corresponding rating; otherwise (i.e., for AAA and AA, B, and C), the risk factor is a weighted average (weights equal to the number of bonds of each rating) of the corresponding index returns.

The entire history of index returns (January 2, 1997, through December 31, 2001) is used to develop the forecast upper and lower VaR thresholds/quantiles. So, for single-rated test portfolios, this variant leads to forecast VaR thresholds that are constant through time and invariant across test portfolios. For multirating test portfolios, the forecast VaR thresholds will depend on the exact composition of the test portfolio; for a given test portfolio, however, the forecast VaR threshold remains constant within a calendar year, as test portfolios are reconstituted only at the beginning of the calendar year.

V1 is the same as *V0*, except that only the cumulative index history up to the beginning of the calendar year under analysis is used in the construction of forecast VaR. So, for single-rated test portfolios, this variant leads to forecast VaR thresholds that are static within a calendar year but vary from year to year. Requiring a minimum of one year of history means that calendar year 1997 cannot be analyzed under this variant.

V2 uses a single risk factor that incorporates the time series of the rating-level index returns, as summarized in Exhibit 3. In other respects, it is the same as *V0*. In this inherently adaptive scheme, the forecast VaR thresholds will vary through time.

V3 augments the risk factor employed in *V0*, with a generic idiosyncratic risk factor.[19] For the idiosyncratic term, only histories prior to the calendar year under analysis are sampled. A fresh idiosyncratic term is sampled each day and for each test portfolio. Therefore, forecast VaR thresholds will vary from day to day for all test portfolios. Requiring a minimum of one year of history means that calendar year 1997 cannot be analyzed under this variant.

Each of these variants assumes, in different ways and to different degrees, some form of stationarity or stability in the stochastic properties of the indexes' returns. The year-to-year variation seen in the tail quantiles of the indexes' returns (Exhibit 4) raises some doubts about such an assumption. Yet, given the relatively short spans of data available, and my limited purposes, extended explorations along these lines are beyond the scope of this article.

For each test portfolio: 1) the exceedance frequency of the VaR threshold is the number of exceedances in a test period divided by the number of observations in the test period; and 2) the average exceedance is the sum of the absolute exceedances (difference between portfolio return and the VaR threshold, provided there is an exceedance; zero otherwise) divided by the total number of observations in the test period. The expected exceedance frequency, p, is 1% under the null hypothesis that a variant's VaR threshold is the true descriptor of empirical reality. That is, the null hypothesis is H_0: $p = 0.01 = p_0$. The standard binomial test can be applied.

For the relevant sample sizes (ranging from 250 to 257), the null can be rejected in favor of the alternative $p > 0.01$, for several combinations of observed exceedance frequencies and significance levels: $\geq 1.94\%$ (significance level of 10% or better), $\geq 2.00\%$ (5% or better), and $\geq 2.80\%$ (1% or better).[20]

Results for Rating-Level Indexes

Only variants *V0*, *V1*, and *V2* are applicable for rating-level indexes as test portfolios. The results are shown in Exhibit 7. To facilitate comparison across the VaR variants, the results are shown separately for each rating group. Exceedance frequencies and average exceedances relative to the lower VaR threshold are shown under the heading *Q01*; results relative to the upper VaR-threshold are shown under the heading *Q99*.

General observations based on V0. The results for the basic variant, *V0*, show noticeable year-to-year variation in the exceedance frequencies, but coupled with a tendency to be higher than the null/expected value of 1%. Rejections of the null at exceedance frequencies ≥ 1.94%: 1) predominate in 1998 and 2001, periods of considerable activity in the credit markets; and 2) are more common for lower credit quality indexes.

Average exceedances in Exhibit 8 are generally greater in 1998 and 2001, and are generally higher for lower credit quality indexes.

Performance of the variants. *V1* performs somewhat worse than *V0*—there are more rejections of the null, and the exceedance frequencies are higher. *V2* offers some improvement over *V0*. There are fewer rejections, even when the null under *V0* is rejected; the results for 2001 are particularly illustrative.[21]

It is interesting to see that even when the past history includes a turbulent year (1998), method *V1* does not perform very well for a subsequent volatile year (2001). This raises some questions about the adequacy of the historical simulation framework based on past history for VaR analysis of corporate bonds, indicating the need for further research. One can assert more confidently that incorporation of the time series properties of the risk factor does lead, on balance, to improved VaR forecasts.

Results for Test Portfolios of Individual Bonds

Each test portfolio consists of 50 securities, and the analysis is carried out for 100 test portfolios for each rating-year group. Exhibits 9 and 10 show the results for exceedance frequencies and average exceedances. They are structured much the same as Exhibit 7, with two main differences. One difference is the addition of another variant, *V3*. The second difference is that the analysis is carried out for 100 test portfolios for each rating-year group. From the cross-sectional distribution (across the 100 test

EXHIBIT 7
VaR Performance for Indexes—Exceedance Frequencies

A. Rating = AAA

	Q01			*Q99*		
	Variant #			Variant #		
Year	V0	V1	V2	V0	V1	V2
1997	1.19	NA	0.80	1.19	NA	1.20
1998	0.80	0.80	0.40	2.00	0.80	2.80
1999	0.80	0.80	1.20	0.40	0.00	0.00
2000	0.40	0.40	0.40	0.40	0.00	0.40
2001	1.55	1.55	1.94	0.78	0.78	0.39

B. Rating = AA

	Q01			*Q99*		
	Variant #			Variant #		
Year	V0	V1	V2	V0	V1	V2
1997	1.58	NA	0.79	1.19	NA	1.19
1998	0.79	0.79	1.19	2.38	0.79	1.98
1999	0.80	0.80	1.20	0.40	0.00	0.40
2000	0.00	0.00	0.00	0.00	0.00	0.80
2001	1.16	1.16	1.54	0.77	0.77	0.39

C. Rating = A

	Q01			*Q99*		
	Variant #			Variant #		
Year	V0	V1	V2	V0	V1	V2
1997	1.19	NA	0.80	0.79	NA	1.20
1998	1.20	1.20	1.20	1.99	1.20	1.99
1999	1.19	0.79	1.19	0.79	0.40	0.79
2000	0.00	0.00	0.00	0.00	0.00	0.40
2001	1.16	1.16	1.54	1.16	1.16	0.39

D. Rating = BBB

	Q01			*Q99*		
	Variant #			Variant #		
Year	V0	V1	V2	V0	V1	V2
1997	0.79	NA	0.79	1.19	NA	1.59
1998	1.19	1.59	1.19	1.59	1.19	1.19
1999	0.79	0.79	0.79	0.40	0.00	0.40
2000	0.00	0.00	0.40	0.40	0.00	0.80
2001	1.94	2.33	1.55	1.16	1.16	0.78

E. Rating = BB

	Q01			*Q99*		
	Variant #			Variant #		
Year	V0	V1	V2	V0	V1	V2
1997	0.79	NA	1.19	1.19	NA	1.59
1998	1.59	1.20	1.59	2.39	1.59	1.99
1999	0.40	0.40	1.20	0.40	0.00	0.40
2000	0.00	0.00	0.00	0.00	0.00	0.00
2001	1.95	2.33	0.78	0.78	0.78	0.78

F. Rating = B

	Q01			*Q99*		
	Variant #			Variant #		
Year	V0	V1	V2	V0	V1	V2
1997	0.00	NA	0.79	0.00	NA	0.40
1998	0.80	4.78	1.99	0.00	8.37	1.20
1999	0.00	0.00	0.79	0.00	0.00	0.00
2000	0.79	2.78	0.40	1.19	1.59	1.19
2001	3.09	4.63	0.77	3.47	6.56	1.93

G. Rating = C

	Q01			*Q99*		
	Variant #			Variant #		
Year	V0	V1	V2	V0	V1	V2
1997	0.00	NA	0.40	0.40	NA	1.19
1998	1.60	4.00	2.00	0.80	2.80	0.00
1999	0.00	0.00	0.40	0.00	0.40	0.79
2000	0.79	3.17	1.19	0.00	0.79	1.19
2001	2.32	3.47	0.77	3.47	5.79	1.54

EXHIBIT 8
VaR Performance for Indexes—Average Exceedance

A. Rating = AAA

	Q01			Q99		
	Variant #			Variant #		
Year	V0	V1	V2	V0	V1	V2
1997	0.14	NA	0.19	0.17	NA	0.21
1998	0.27	0.24	0.05	0.22	0.10	0.27
1999	0.14	0.12	0.08	0.01	NA	NA
2000	0.02	0.02	0.01	0.01	NA	0.02
2001	0.26	0.27	0.09	0.11	0.10	0.14

B. Rating = AA

	Q01			Q99		
	Variant #			Variant #		
Year	V0	V1	V2	V0	V1	V2
1997	0.18	NA	0.23	0.17	NA	0.16
1998	0.22	0.16	0.19	0.22	0.09	0.19
1999	0.13	0.13	0.10	0.02	NA	0.06
2000	NA	NA	NA	NA	NA	0.04
2001	0.27	0.28	0.13	0.12	0.11	0.12

C. Rating = A

	Q01			Q99		
	Variant #			Variant #		
Year	V0	V1	V2	V0	V1	V2
1997	0.14	NA	0.15	0.15	NA	0.17
1998	0.49	0.44	0.26	0.28	0.20	0.25
1999	0.17	0.10	0.09	0.05	0.01	0.01
2000	NA	NA	NA	NA	NA	0.03
2001	0.32	0.32	0.10	0.20	0.22	0.15

D. Rating = BBB

	Q01			Q99		
	Variant #			Variant #		
Year	V0	V1	V2	V0	V1	V2
1997	0.09	NA	0.22	0.16	NA	0.28
1998	0.67	0.70	0.48	0.24	0.17	0.18
1999	0.10	0.08	0.05	0.03	NA	0.01
2000	NA	NA	0.01	0.01	NA	0.04
2001	0.28	0.41	0.16	0.17	0.18	0.16

E. Rating = BB

	Q01			Q99		
	Variant #			Variant #		
Year	V0	V1	V2	V0	V1	V2
1997	0.13	NA	0.23	0.13	NA	0.14
1998	0.38	0.35	0.24	0.19	0.12	0.16
1999	0.03	0.01	0.06	0.01	NA	0.01
2000	NA	NA	NA	NA	NA	NA
2001	0.87	0.96	0.65	0.08	0.07	0.01

F. Rating = B

	Q01			Q99		
	Variant #			Variant #		
Year	V0	V1	V2	V0	V1	V2
1997	NA	NA	0.12	NA	NA	0.02
1998	0.69	1.58	0.68	NA	1.00	0.01
1999	NA	NA	0.04	NA	NA	NA
2000	0.13	0.55	0.01	0.20	0.51	0.09
2001	1.29	2.19	1.04	0.46	1.40	0.40

G. Rating = C

	Q01			Q99		
	Variant #			Variant #		
Year	V0	V1	V2	V0	V1	V2
1997	NA	NA	0.09	0.01	NA	0.13
1998	1.26	2.14	0.85	0.09	0.40	NA
1999	NA	NA	0.08	NA	0.01	0.03
2000	0.26	0.78	0.22	NA	0.05	0.16
2001	2.57	2.68	1.48	1.26	1.74	0.45

portfolios), the median and the extreme (highest value) are shown in the tables.

General observations based on V0. Exceedance frequencies of the cross-sectional median for the basic variant, *V0*, in Exhibit 9 suggest some general observations. For ratings AAA through BBB, exceedance frequencies statistically greater than 1% are concentrated in the years 1998 and 2001. For ratings BB and lower, statistically significant exceedances are more widespread.

The actual exceedance frequencies also tend to be greater for the lower credit quality test portfolios. Compared to index test portfolios, the exceedance frequencies for these test portfolios are greater. The difference between the two seems to increase with declining credit quality.

Generally similar characteristics are seen for the average exceedance in Exhibit 10.

Performance of the variants. For all practical purposes, variants *V0, V1,* and *V2* deliver essentially identical results: 1) virtually identical patterns of statistically significant exceedance; and 2) exceedance frequencies of similar magnitude. Particularly interesting is that, for test portfolios constructed from individual bonds, variant *V2* fails to deliver the improvement noted for rating-level index portfolios. An intuitively plausible explanation is that the returns of these test portfolios are partly driven by an idiosyncratic risk factor, relative to the rating-level index risk factor already included.

Strong support for this explanation comes from the markedly better performance of variant *V3*. Exceedance frequencies are substantially closer to the null value of 1%, and there are far fewer statistically significant exceedances. The improvement is especially striking where the other variants fare most poorly, e.g., in 2001.[22]

Similar patterns are evident using the average exceedance metric. Namely, average exceedances for variants *V0, V1,* and *V2* are very similar to each other, while the average exceedance under variant *V3* is generally noticeably smaller.

Robustness: Larger portfolios. An immediate question is the performance of larger portfolios. To check the robustness of the patterns, I repeat the analysis for larger test portfolios—namely, test portfolios consisting of all available bonds in rating-year group as shown in Exhibit 5. The conclusions remain unaffected.

Time distribution of exceedance dates. The premise that idiosyncratic variation is responsible for the findings would be further strengthened if different portfolios tended to experience exceedances at different dates.[23] Spot checks for several rating-year combinations show no evidence of

EXHIBIT 9

VaR Performance—Exceedance Frequencies

A. Rating = AAA and AA

		Q01 Variant #				*Q99* Variant #			
Year	Stat.	V0	V1	V2	V3	V0	V1	V2	V3
1997	Median	0.40	NA	0.00	NA	0.00	NA	0.00	NA
	Extreme	0.80	NA	1.20	NA	0.80	NA	0.40	NA
1998	Median	1.20	0.40	0.40	0.00	0.40	0.00	0.80	0.00
	Extreme	1.99	1.59	1.20	0.80	1.19	0.79	1.59	0.40
1999	Median	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Extreme	0.80	0.80	0.80	0.00	0.40	0.00	0.40	0.00
2000	Median	0.60	0.40	1.20	0.40	0.00	0.00	0.00	0.00
	Extreme	1.99	1.99	3.19	1.59	0.80	0.79	0.80	0.40
2001	Median	0.58	0.78	0.39	0.00	0.00	0.00	0.00	0.33
	Extreme	3.10	3.10	2.71	1.94	0.78	0.78	0.78	0.39

B. Rating = A

		Q01 Variant #				*Q99* Variant #			
Year	Stat.	V0	V1	V2	V3	V0	V1	V2	V3
1997	Median	1.60	NA	1.60	NA	1.60	NA	2.80	NA
	Extreme	2.40	NA	2.00	NA	2.00	NA	3.60	NA
1998	Median	4.78	3.98	3.59	1.59	4.38	3.98	4.58	2.39
	Extreme	6.80	5.60	4.40	2.40	5.98	5.18	6.77	3.19
1999	Median	1.79	1.20	0.80	0.00	0.40	0.00	0.00	0.00
	Extreme	3.60	3.20	2.00	0.80	1.19	0.80	0.80	0.00
2000	Median	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Extreme	0.80	0.80	1.20	0.40	0.00	0.00	0.40	0.00
2001	Median	6.61	6.61	4.67	1.95	5.06	5.45	4.28	1.56
	Extreme	8.20	8.20	6.20	4.30	9.70	9.71	6.99	3.11

C. Rating = BBB

		Q01 Variant #				*Q99* Variant #			
Year	Stat.	V0	V1	V2	V3	V0	V1	V2	V3
1997	Median	2.80	NA	2.80	NA	4.40	NA	5.20	NA
	Extreme	3.20	NA	3.20	NA	6.00	NA	6.80	NA
1998	Median	3.19	3.19	1.99	1.59	3.59	3.19	3.19	2.39
	Extreme	4.00	4.40	4.00	2.40	5.18	4.38	5.57	3.19
1999	Median	2.79	2.39	2.79	1.20	2.79	1.20	2.79	0.40
	Extreme	5.60	5.20	4.40	2.40	5.57	3.18	5.58	1.99
2000	Median	0.80	0.80	1.59	0.00	0.80	0.80	1.59	0.00
	Extreme	2.40	2.40	3.20	1.20	2.78	2.39	3.98	1.59
2001	Median	1.56	2.33	1.56	0.39	1.17	1.17	1.56	0.39
	Extreme	4.30	4.70	3.50	1.60	3.50	3.50	3.50	1.95

D. Rating = BB

		Q01 Variant #				*Q99* Variant #			
Year	Stat.	V0	V1	V2	V3	V0	V1	V2	V3
1997	Median	2.40	NA	3.60	NA	7.60	NA	7.60	NA
	Extreme	3.60	NA	4.40	NA	10.80	NA	9.60	NA
1998	Median	3.59	3.19	3.19	0.40	5.58	4.78	5.18	0.00
	Extreme	5.20	4.80	7.60	1.60	8.76	7.17	8.76	1.20
1999	Median	4.98	3.59	7.57	0.40	9.56	5.98	10.76	0.40
	Extreme	8.00	7.20	10.40	2.00	13.54	9.96	15.13	1.59
2000	Median	3.59	3.59	5.18	0.40	4.78	4.38	6.77	0.00
	Extreme	7.60	7.60	11.50	2.00	9.16	8.36	11.14	1.59
2001	Median	8.95	10.51	9.34	2.33	12.84	12.65	14.40	3.11
	Extreme	12.40	14.00	13.20	3.90	17.51	17.51	17.12	5.83

E. Rating = B and C

		Q01 Variant #				*Q99* Variant #			
Year	Stat.	V0	V1	V2	V3	V0	V1	V2	V3
1997	Median	0.40	NA	4.40	NA	0.40	NA	9.20	NA
	Extreme	0.80	NA	6.00	NA	1.60	NA	13.99	NA
1998	Median	2.79	4.78	6.37	0.40	1.99	9.36	6.37	0.40
	Extreme	4.00	7.20	9.60	1.20	3.98	13.55	9.56	1.20
1999	Median	1.99	3.19	8.76	0.40	2.79	4.38	11.55	0.80
	Extreme	4.00	7.20	17.10	1.99	5.98	7.96	15.94	2.39
2000	Median	4.38	6.37	8.37	1.20	1.99	3.19	5.98	0.40
	Extreme	8.40	11.50	12.30	3.58	3.98	6.37	9.16	1.59
2001	Median	4.28	5.45	3.89	1.95	4.67	6.61	5.45	1.95
	Extreme	8.20	8.90	7.80	5.05	8.55	11.66	10.49	3.89

EXHIBIT 10

VaR Performance—Average Exceedance

A. Rating = AAA and AA

		Q01 Variant #				*Q99* Variant #			
Year	Stat.	V0	V1	V2	V3	V0	V1	V2	V3
1997	Median	0.10	NA	0.11	NA	0.04	NA	0.05	NA
	Extreme	0.26	NA	0.29	NA	0.12	NA	0.14	NA
1998	Median	0.27	0.20	0.18	0.16	0.09	0.04	0.16	0.10
	Extreme	0.79	0.71	0.53	0.59	0.31	0.24	0.35	0.33
1999	Median	0.03	0.02	0.01	0.02	0.04	0.02	0.08	0.03
	Extreme	0.10	0.08	0.07	0.05	0.08	0.06	0.12	0.09
2000	Median	1.32	1.32	1.43	1.27	0.10	0.09	0.11	0.11
	Extreme	3.29	3.23	3.78	2.99	0.29	0.28	0.42	0.26
2001	Median	0.58	0.61	0.51	0.27	0.26	0.25	0.33	0.18
	Extreme	4.04	4.08	3.73	3.26	0.99	0.96	1.12	0.73

B. Rating = A

		Q01 Variant #				*Q99* Variant #			
Year	Stat.	V0	V1	V2	V3	V0	V1	V2	V3
1997	Median	0.41	NA	0.38	NA	0.28	NA	0.32	NA
	Extreme	0.54	NA	0.51	NA	0.40	NA	0.50	NA
1998	Median	1.89	1.71	1.34	0.73	1.08	0.90	0.99	0.28
	Extreme	2.46	2.25	1.84	1.13	1.50	1.30	1.40	0.57
1999	Median	0.26	0.18	0.16	0.71	0.02	0.02	0.01	NA
	Extreme	1.63	1.40	1.25	0.82	0.13	0.07	0.04	NA
2000	Median	0.42	0.40	0.09	0.28	NA	NA	0.00	NA
	Extreme	0.65	0.61	0.76	0.45	NA	NA	0.01	NA
2001	Median	2.06	2.08	1.49	0.47	1.31	1.38	1.05	0.48
	Extreme	3.63	3.66	2.78	1.37	2.45	2.59	2.04	1.07

C. Rating = BBB

		Q01 Variant #				*Q99* Variant #			
Year	Stat.	V0	V1	V2	V3	V0	V1	V2	V3
1997	Median	0.96	NA	1.10	NA	1.11	NA	1.56	NA
	Extreme	1.34	NA	1.49	NA	1.60	NA	2.07	NA
1998	Median	1.18	1.25	0.44	0.46	0.67	0.5	0.42	0.23
	Extreme	1.81	1.88	0.85	0.94	1.33	1.1	1.02	0.88
1999	Median	0.79	0.75	0.70	0.22	0.24	0.1	0.22	0.06
	Extreme	1.41	1.33	1.28	0.64	0.68	0.3	0.61	0.21
2000	Median	0.13	0.12	0.16	0.05	0.06	0.0	0.27	0.05
	Extreme	2.05	2.03	2.10	1.62	0.47	0.4	0.87	0.19
2001	Median	0.49	0.66	0.41	0.47	0.17	0.1	0.19	0.07
	Extreme	2.15	2.33	1.96	1.60	0.68	0.6	0.73	0.41

D. Rating = BB

		Q01 Variant #				*Q99* Variant #			
Year	Stat.	V0	V1	V2	V3	V0	V1	V2	V3
1997	Median	0.67	NA	0.89	NA	1.19	NA	1.36	NA
	Extreme	0.95	NA	1.21	NA	1.78	NA	1.87	NA
1998	Median	0.82	0.74	0.64	0.06	0.67	0.46	0.66	0.02
	Extreme	1.52	1.40	1.25	0.38	1.23	0.95	1.24	0.09
1999	Median	1.08	0.83	1.53	0.17	1.11	0.63	1.51	0.05
	Extreme	2.85	2.39	3.42	0.77	1.87	1.23	2.33	0.22
2000	Median	0.70	0.70	1.12	0.19	0.58	0.44	0.83	0.04
	Extreme	4.99	4.99	5.93	3.48	1.35	1.17	1.72	0.23
2001	Median	3.60	4.04	3.72	1.54	2.70	2.62	3.02	0.59
	Extreme	6.66	7.13	6.19	3.82	4.33	4.21	4.53	1.41

E. Rating = B and C

		Q01 Variant #				*Q99* Variant #			
Year	Stat.	V0	V1	V2	V3	V0	V1	V2	V3
1997	Median	0.13	NA	0.69	NA	0.04	NA	1.01	NA
	Extreme	0.24	NA	1.06	NA	0.13	NA	1.66	NA
1998	Median	1.32	2.40	1.62	0.31	0.53	1.9	0.91	0.10
	Extreme	2.63	3.82	2.84	0.95	1.52	3.5	2.03	0.60
1999	Median	1.60	2.03	3.44	0.92	1.02	1.3	2.65	0.40
	Extreme	4.42	4.93	6.95	3.11	2.53	2.8	4.40	1.02
2000	Median	2.95	3.91	3.96	0.71	0.40	0.7	1.13	0.11
	Extreme	8.21	9.68	8.46	3.78	1.62	2.0	2.24	0.40
2001	Median	6.62	7.34	6.09	4.34	2.38	2.8	2.18	0.85
	Extreme	17.77	18.62	16.63	12.82	4.93	5.6	4.71	2.65

extreme clustering. Exceedance dates for different portfolios appear to be quite dispersed in time.

The main conclusion from these results is that a single risk factor derived from the returns on rating-level indexes appears to be inadequate to capture the tail properties of portfolios consisting of individual corporate bonds. The proximate cause is the presence of substantial idiosyncratic risk (beyond the single risk factor) in such portfolios. Adding a generic, easily constructed, idiosyncratic risk factor appears to provide substantial improvement in daily VaR analysis of such portfolios.

VI. CONCLUSIONS

I have examined alternative ways to carry out daily VaR for credit-risky portfolios. Employing historical simulation (HS) as the basic framework, I focus on variants differing in two main respects. First, a variant either employs just a single risk factor, based on the returns of rating-level indexes, or includes an added generic idiosyncratic risk factor. Second, the treatment of the index-based risk factor differs: 1) time series properties are either incorporated or ignored; and 2) either the full sample or just the cumulative history to date is used. The analysis is carried out for rating-level indexes (a special test portfolio, which is large and well-diversified) and test portfolios constructed from individual bonds.

For rating-level indexes, I find that incorporation of the time series properties of the risk factor leads, on balance, to reasonable improvements in VaR performance. This roughly parallels findings elsewhere in the literature. For test portfolios constructed from individual bonds, the performance of variants employing a single risk factor (based on rating-level index returns) leaves substantial room for improvement. Incorporating a second generic idiosyncratic risk factor developed here leads to significantly improved VaR performance.

ENDNOTES

The author thanks members of the Risk Analysis Division for helpful comments. Special thanks to Regina Villasmil for invaluable research assistance.

[1]That is, they may be computationally inexpensive for occasional reestimation, but potentially expensive on a daily basis.

[2]Extreme-value methods could also be considered.

[3]Separate upper and lower VaR thresholds are considered 1) because the forecast distributions are typically not symmetric; and 2) to allow for both long and short (net) positions.

[4]That is, whenever the realized return is algebraically below the lower VaR threshold or above the upper VaR threshold.

[5]This definition of average exceedance differs slightly from the customary definition of expected shortfall, for which the denominator would be the number of exceedances. That is, the customary definition considers the expected excess *conditional* on exceedance, while my definition is an unconditional average. To compare VaR methods yielding different exceedance levels, the latter measure appears to be more suitable, since a common denominator applies.

[6]More involved statistical tests, such as discussed in Christoffersen, Hahn, and Inoue [2001], would be somewhat counter to the spirit and purposes of my research.

[7]Of course, the rating imputation and hence the computation of the weights may occur more often than just once at the beginning of the test period.

[8]The benchmark series X_t is assumed to be strictly stationary, and its dynamics given by $X_t = \mu_t + \sigma_t Z_t$, where the innovation process Z_t is assumed to be strictly white noise. Using lower case symbols for realizations and hats for estimated parameters, we have $z_t = x_t - \hat{\mu}_t)/\hat{\sigma}_t$; I refer to z_t as the residual. See McNeil and Frey [2000] for details.

[9]Phoa [1999], in an analysis of swap spread data, suggests that extreme-value theory may be useful in modeling credit spread behavior.

[10]Other researchers provide empirical evidence of a relationship between credit spreads (more generally, credit quality) and assorted macroeconomic variables, such as default-free rates and explicit measures of economic conditions. These include Duffee [1998], Alessandrini [1999], and Christiansen [2000]. Carty [2000] documents that corporate default rates are influenced by economic conditions.

[11]Each bond is uniquely identified by an eight-character CUSIP code. Bonds that share the first six characters are deemed to originate from the same issuer.

[12]Researchers have an abiding concern about the quality of corporate bond price data, especially between so-called matrix prices versus dealer quotes. For a discussion of these and related issues, see Hong and Warga [2000], Hotchkiss and Ronen [2000], and Schultz [2000].

[13]The following are paraphrased versions of Bloomberg's descriptions of these prices.

> First, Bloomberg creates fair market sector curves. Each curve is composed of bonds having the same issuer type and credit quality. BFV quantifies the value of any embedded options and, depending on the option type, adds to or subtracts that from the value.
>
> To reflect the current market for the selected issue, Bloomberg uses BGN to construct the yield curve. BGN is the average of recently contributed prices that come from high-quality price contributors and that agree within a specified margin. BGN prices are recal-

culated every day when the market closes. Other information that Bloomberg considers relevant may also be used. The goal of the methodology is to produce consensus pricing. If Bloomberg is not comfortable that a consensus price can be assigned, an issue is marked as not priced.

[14]This assigns zero returns to non-trading days. Other algorithms for apportioning returns could be used.

[15]If the absolute return based on the average price over $t-4, \ldots, t-2$ and the average price over $t+1, \ldots, t+5$ is less than 2% (while $|r_t| \geq 10\%$), an error is suspected. Then p_{t-1} is compared against the average price over the previous four days, and p_t is compared to the average price over the following four days. If the absolute return is 5% or more, that price (i.e., p_{t-1} or p_t) is set as missing.

[16]Estimation is carried out using maximum-likelihood, assuming that the residuals are conditionally normally distributed. Unfortunately, the residuals from the best fitted model are non-normal; therefore, the initial model estimation should perhaps be carried out under a non-normal assumption. Such a search, as computationally intensive, however, would be tangential and somewhat contrary to my intent.

[17]As further evidence of differences across rating groups, observe that the smallest *Q01* occurs in 1998 for A, BBB, BB, B, and C, but in 2001 for AAA and AA—but 2001 is also when the C index recorded its largest *Q99*.

[18]Rating groups AAA and AA are combined because their behavior is expected to be very similar. The relatively few C-rated bonds in the sample are combined with B-rated bonds.

[19] First, *n*Est (= 40 for results reported) estimation test portfolios are created. For each day k, the residual for estimation portfolio i is defined to be the portfolio's return minus benchmark 1's return for day k. (Regression analyses offer some support for the implicit assumption of unit β sensitivity.) These residual series are built up for each year-rating combination, providing a historical sample from which draws can be made for HS. Note that by *not* pooling the residuals over time, this method recognizes that the *n*Est residuals for day k may be cross-sectionally correlated; i.e., the method allows for a common factor not present in or captured by the rating-level indexes. Second, in VaR application, for any day k, a random draw from the *n*Est estimation residuals for day k is added to the first (index-based) risk factor.

[20]These are obtained from the exact binomial distribution under the null, rather than the normal approximation—for $p_0 = 0.01$, the normal approximation includes negative values for realizable p within the 3σ interval.

[21]The good news is tempered by the fact that there are also cases of statistically significant exceedance under *V2* but not under *V0*.

[22]The presence of an idiosyncratic factor in these test portfolios would also help to explain the (perhaps counter-intuitive) result that the GARCH-adjusted method (*V2*) performs more poorly than the basic method (*V0*). The reason is this. Both methods (*V0* and *V2*) are based on a single risk factor, tied to the index returns. For the basic method, the VaR bounds are simply based on the tail quantiles of the entire sample history of index returns; therefore, the bounds are invariant with time and, roughly speaking, will tend to be conservative in periods of low index volatility (resulting in low exceedance frequencies) and liberal in periods of high index volatility (resulting in higher exceedance frequencies). Variant *V2*, being more adaptive to current volatility, should attenuate these undesirable effects. So, for a portfolio with little idiosyncratic variability (relative to the single risk factor), *V2* should deliver better VaR performance than *V0*. But for test portfolios with substantial idiosyncratic variability, large idiosyncratic moves during periods of relatively low index volatility would give rise to *greater* exceedance frequencies under *V2* than *V0*.

It should also be mentioned that the visual impression from cross-sectional medians is countered at times by the Wilcoxon signed-rank test on *matched differences* (relative to *V0*). Exceedance frequencies for variants *V1* and *V3* are sometimes statistically smaller, even when the *cross-sectional medians* are virtually identical. The reason is that typically many of the matched differences are zero, so the statistical result is heavily influenced by a few non-zero observations. But, statistical excess relative to the null value of 1% is of greater relevance here.

[23]This is not a necessary condition, because, as already noted, the residuals relative to the index risk factor could well entail common factors not present in the index risk factor.

REFERENCES

Alessandrini, F. "Credit Risk, Interest Rate Risk, and the Business Cycle." *The Journal of Fixed Income*, Vol. 9, No. 2 (1999), pp. 42–53.

Barone Adesi, K.G., and L. Vosper. "Filtering Historical Simulation: Backtest Analysis." Working paper, Universita' della Svizzera Italiana and City University Business School, 2000.

Carty, L. "Corporate Credit Risk Dynamics." *Financial Analyst Journal*, Vol. 56, No. 4 (July/August 2000), pp. 67–81.

Christiansen, C. "Credit Spreads and the Term Structure of Interest Rates." Working paper, Aarhus School of Business, 2000. Available at www.asb.dk/cha/.

Christoffersen, P., J. Hahn, and A. Inoue. "Testing and Comparing Value-at-Risk Measures." *Journal of Empirical Finance*, 8 (2001), pp. 325–342.

Collin-Dufresne, P.R.G., and S. Martin. "Determinants of Credit Spread Changes." Working paper, Carnegie-Mellon University, 2000.

Duffee, G. "The Relationship Between Treasury Yields and Corporate Bond Yield Spreads." *Journal of Finance*, 53 (1998), pp. 2225–2241.

D'Vari, R., and J.C. Sosa. "Value at Risk Estimates for Brady Bond Portfolios." *The Journal of Fixed Income*, Vol. 10, No. 3 (2000), pp. 7–23.

Hong, G., and A. Warga. "An Empirical Study of Bond Market Transactions." *Financial Analysts Journal*, March/April 2000, pp. 32–46.

Hotchkiss, E., and T. Ronen. "The Information Efficiency of the Corporate Bond Market: An Intraday Analysis." Working paper, Boston College, 2000.

Kiesel, R., W. Perraudin, and A. Taylor. "The Structure of Credit Risk." Working paper, Birkbeck College and Bank of England, 1999. Available at www.econ.bbk.ac.uk/perraudin/ insider.econ.bbk.ac.uk/dept/ifr/.

Kim, J. "Hypothesis Test of Default Correlation and Application to Specific Risk." *Risk Metrics Journal*, 1 (2000).

McNeil, A., and R. Frey. "Estimation of Tail-Related Risk Measures for Heteroscedastic Financial Time Series: An Extreme Value Approach." *Journal of Empirical Finance*, Vol. 7 (2000), pp. 271–300.

Pedrosa, M., and R. Roll. "Systematic Risk of Corporate Bond Credit Spreads." *The Journal of Fixed Income*, December 1998, pp. 7–26.

Phoa, W. "Estimating Credit Spread Risk Using Extreme Value Theory." *The Journal of Portfolio Management*, Spring 1999, pp. 69–73.

Schultz, P. "Corporate Bonds Trading Costs and Practices: A Peek Behind the Curtain." Working paper, University of Notre Dame, 2000.

Vlaar, P. "Value at Risk Models for Dutch Bond Portfolios." *Journal of Banking and Finance*, 24 (2000), pp. 1131–1154.