



## Default Risk in Corporate Yield Spreads

**Georges Dionne, Geneviève Gauthier, Khemais Hammami,  
Mathieu Maurice, and Jean-Guy Simonato\***

*An important research question examined in the credit risk literature focuses on the proportion of corporate yield spreads attributed to default risk. This topic is reexamined in light of the different issues associated with the computation of default probabilities obtained from historical default data. We find that the estimated default risk proportion in corporate yield spreads is sensitive to the ex ante estimated term structure of default probabilities used as inputs. This proportion can become a large fraction of the spread when sensitivity analyses are made with respect to the period over which the probabilities are estimated and the recovery rates.*

---

An important research question studied in the credit risk literature looks at the proportion of corporate yield spreads explained by default risk (i.e., the part of the spread rewarding the investor for the actuarial expected default loss). This question is not only important for the pricing of bonds and credit derivatives but also for computing banks' optimal economic capital for credit risk (Crouhy, Galai, and Mark, 2000; Gordy, 2000). Elton et al. (2001) have verified that only a small fraction of corporate yield spreads can be attributed to default risk or expected default loss. They obtain this result from a reduced form model and have demonstrated that the expected default loss explains no more than 25% of corporate spot spreads. The remainder is attributed to a tax premium and a risk premium for systematic risk. Huang and Huang (2003) reached a similar conclusion with a structural model. They verified that for investment grade bonds (Baa and higher ratings) only 20% of the spread is explained by default risk.

The part of the spread rewarding the investor for the expected default loss can be seen as the product of two key components: 1) the probability of defaulting and 2) the loss given default ( $1 - \text{the recovery rate}$ ). Such quantities may be inferred from databases on historical default frequencies from Moody's and Standard & Poor's. For example, measuring default probabilities can be done by estimating, in a first step, the transition probabilities between rating classes and then by using these, in a second step, to compute the term structure of the default probability. This is the approach used in Elton et al. (2001). Although this method appears straightforward, obtaining probability estimates with such a procedure is not a trivial exercise. As is the case

---

*The authors acknowledge the financial support of the National Science and Engineering Research Council of Canada (NSERC), of the Fonds québécois de recherche sur la nature et les technologies (FORNT), of the Social Sciences and Humanities Research Council of Canada (SSHRC), of the Institut de Finance Mathématique de Montréal (IFM2), and of the Centre for Research on E-finance (CREF). Previous versions of the paper were presented at the Second International Conference on Credit Risk in Montreal, the 2004 SCSE Annual Conference in Québec, the 2004 C.R.E.D.I.T. Risk Conference in Venice, and at the Finance Department of the University of Toronto. J. Christensen, P. François, E. Hansen, D. Lando, T. Schuermann, and an anonymous referee made very useful comments on previous versions of this manuscript.*

*\*Georges Dionne is a Professor of Finance at HEC Montréal, Montréal, Québec, Canada. Geneviève Gauthier is a Professor of Management Science at HEC Montréal, Montréal, Québec, Canada. Khemais Hammami is a Financial Analyst at Caisse de Dépôt et Placement du Québec, Montreal, Québec, Canada. Mathieu Maurice is a Financial Analyst at Caisse de Dépôt et Placement du Québec, Montréal, Québec, Canada. Jean-Guy Simonato is a Professor of Finance at HEC Montréal, Montréal, Québec, Canada.*

with the measurement of recovery rates, many important issues arise in the process and different choices might lead to different results.

The first issue concerns the period over which the estimation is performed. As illustrated in Bangia et al. (2002), transition matrix estimates are sensitive to the period in which they are computed. Business and credit cycles might have a serious impact on the estimated transition matrices and recovery rates and may lead to highly different estimates for the default risk proportion.

The second issue calling for close attention is the statistical approach. Because defaults and rating transitions are rare events, the typical cohort approach used by Moody's and Standard & Poor's will produce transition probability matrices with many cells equal to zero. This does not mean that the probability of the cell is zero but that its estimate is zero. Such a characteristic could lead to underestimation of the default risk fraction in corporate yield spreads. Lando and Skodeberg (2002) have shown that a continuous-time analysis of rating transitions using generator matrices will improve the estimates of rare transitions even when they are not observed in the data, a result that cannot be obtained with the discrete-time cohort approach of Carty and Fons (1993) and Carty (1997).

The third issue arising in computing default and transition probabilities is the data-filtering process that determines the information considered regarding issuers' movements in the database. For example, one must decide whether to consider issuers that are present at the beginning of the estimation period but leave for reasons other than default (withdrawn rating or right censoring). Another choice is whether to consider issuers entering the database after the starting date of estimation. Again, these choices may have nonnegligible impacts on the final estimates.

The fourth consideration in computations of the default and transition probabilities is the statistical precision with which these quantities are calculated. The statistical uncertainty associated with these estimates should be accounted for and reflected in the form of confidence intervals. It should be noted that this uncertainty is not a part of the default risk proportions. The default risk spread we measure here is associated with the expected default loss. Statistical uncertainty is instead linked to the unexpected loss and should not affect the level of the estimated default spread proportions. It does, however, affect the confidence sets that can be built around the point estimates. Thus, confidence intervals associated with the point estimates should be computed and reported to ensure a meaningful interpretation of the results.

The last important issue concerns the link between recovery rates and the proportions of defaulting firms. In the literature, this link is found to be negative (Hu and Perraudin, 2002; Altman et al., 2005). Such a negative correlation is likely to result in higher default spreads. In periods where firms are more likely to default, a lower recovery is obtained for more defaulting firms. In periods of low default, a higher recovery will be obtained, but for fewer defaulting firms. Therefore, recovery rates linked with the proportions of defaulting firms may affect the estimated default spread.

In this article, we revisit the estimation of default spreads in light of the above considerations. For this purpose, we introduce a simple continuous-time model of corporate zero coupon bonds where default time and default probabilities are characterized by a generator matrix describing the firms' credit rating migrations. This approach is interesting in our context as it allows us to address data-filtering issues when estimating the generator. Such a model can also be conveniently simulated. This enables us to address the inference issue and obtain approximate confidence intervals for the default spread proportions. To use historical databases for assessing the various alternatives associated with the estimation of physical default probabilities, our model is built under an assumption of risk neutrality. Consequently, our estimates and analyses only

account for expected default loss and do not include any of the various risk premia potentially present in corporate spreads.<sup>1</sup>

Our empirical analysis proceeds as follows. First, we look at the issues associated with the choice of estimation periods and statistical approach for estimating transition matrices and computing default probabilities. More specifically, sensitivity to the estimation period is illustrated using a rolling window approach to estimate the ex ante time-varying transition and default probabilities that will then be used as inputs to acquire default spread proportions. This approach considers that the recent history of credit migration and default data is most relevant when assessing the probabilities of defaulting. Comparisons are then made between the estimated proportions calculated with the cohort and the continuous-time generator approach. The results indicate that the average default spread proportion for 10 years to maturity Baa bonds can jump from 35% (Table IV) for the case obtained with a fixed cohort transition matrix to 54% (Table VII, 1987-1996 period) for one obtained with an ex ante time-varying cohort transition matrix and recovery approach. These estimates are also variable over time. For instance, for the first half of our sample (1987-1991 period), the estimated proportion jumps from 31% (Table IV) to 74% (Table VII). These results are confirmed with the more robust generator estimation approach.

We then address the data-filtering issues. Three data-filtering procedures reflecting different types of information are considered: 1) the first excludes issuers entering after the starting date of estimation (entry firms hereafter) and withdrawn-rating observations, 2) the second eliminates only entry firm observations, and 3) the third considers entry firms observations and withdrawn-rating observations. Our results demonstrate that the estimated proportions are sensitive to the choices relative to withdrawn ratings and entry firms. Indeed, for a Baa-rated firm, the estimated proportion will range from 42% to 53% (Table X, 1987-1996 period) depending on the filtration approach.

Finally, we study the statistical inference issues. For this purpose, we use a simulation approach to compute approximate confidence intervals in the spirit of Christensen, Hansen, and Lando (2004). In many cases, the 95% confidence sets are wide, illustrating the precision of the point estimates. The impact of the negative link between the recovery and default rate is also studied in this framework. Using a recovery rate series from Moody's and the expected proportions of defaulting firms obtained from our simulation approach, we obtain a statistical relationship similar to those estimated in the literature. This estimated correlation is used in our simulation framework to assess the impact of stochastic recovery rates on the estimated proportions of credit spreads in yield spreads and their associated confidence sets.

The rest of the paper is organized as follows. In Section I, we describe how the empirical bond-spread curves are estimated. Section II presents the default spread model used to estimate the default proportion of the corporate yield spread for different rating categories and maturities. Section III explains the estimation methodologies. The numerical findings are then presented in Section IV. More precisely, Subsections A and B present the default risk proportions obtained with this model and examine their sensitivity to the sample period and estimation methodology of probabilities. The results from the information considered in the default database and inference are then presented in Subsections C and D. Subsection E examines the effect of time-varying recovery rates. Section V provides our conclusion.

---

<sup>1</sup>Credit spreads are usually thought as being formed of various parts: 1) the expected default loss, 2) a risk premium on changes in default intensity, 3) a jump risk premium on the default event, 4) a risk premium on recovery risk, 5) a tax effect, and 6) a liquidity premium. The estimated credit spread obtained with the model used here will only include the first part, without any risk premia.

## I. Empirical Bond-Spread Curves

Our bond price data come from the Lehman Brothers Fixed Income Database (Warga, 1998). We chose these data to facilitate comparisons with other articles in the literature using the same information. The data contain information on monthly prices (quote and matrix), accrued interest, coupons, ratings, callability, and returns on all investment-grade corporate and government bonds for the period from January 1987 to December 1996. All bonds with matrix prices and options were eliminated; bonds not included in Lehman Brothers' bond indexes and bonds with an odd frequency of coupon payments were also dropped. A detailed description of the bond-filtering procedure and of the treatment of accrued interest is available upon request.

Month-end estimates of the yield-spread curves on zero coupon bonds for each rating class are needed to implement the models. These yield-spread curves are computed from zero coupon yield curves obtained with the Nelson and Siegel (1987) approach on government and corporate bonds grouped in three categories: Aa, A, and Baa. When estimating the zero coupon yield curves of corporate bonds in a first pass, we remove all bonds with a pricing error higher than \$5. We then repeat the Nelson and Siegel (1987) calibration procedure and data removal process until all bonds with a pricing error higher than \$5 have been eliminated. Using this procedure, 776 bonds were eliminated (1 Aa, 90 A, and 695 Baa) out of a total of 33,401 bonds found in the industrial sector, which is the focus of this study. Our results are coherent in that all of our estimated empirical bond-spread curves, defined as the difference in yield to maturity of corporate and government zero coupon bonds, are positive. Moreover, the bond-spread curves between a high rating class and a lower rating class are also positive.

Table I reports the average corporate yield spreads for two to 10 years of maturity. The results are very close to those presented in Table I of Elton et al. (2001) for the industrial sector. The small discrepancies might be explained by differences in data filtration and estimation algorithms. In the first panel, the results cover the entire 10-year period, while the second and third panels refer to two subperiods of five years. Finally, Table II compares the average root mean squared errors of the difference between theoretical bond prices computed using the Nelson and Siegel (1987) model and the actual bond prices for treasury and industrial corporate bonds. Again, our results are similar to those reported in the literature.

## II. Default Spread Model

Here we define the corporate yield spread as the difference between the yield curves of the risky zero coupon bond and the risk-free, zero coupon bond. To characterize corporate yield spreads, one needs only to model the values of a risk-free and a corporate zero coupon bond. The model developed here, unlike that of Elton et al. (2001), avoids specifying a coupon rate that might absorb effects unrelated to default risk. Thus, the model we propose focuses on zero coupon bonds and assumes that a corporate yield spread may be totally explained by the recovery rate and the possibility of default. The model will be used to measure how much of the observed corporate yield spread is explained by these two components.

Our model relies on a constant recovery rate  $\rho$  and the intensity  $\{\lambda_t : t \geq 0\}$  associated with the distribution of  $\tau$ , the default time. The risk-free discount factor for the time interval  $(t, T]$  is  $\beta(t, T) = \exp(-\int_t^T r(s) ds)$  where  $r(s)$  denotes the instantaneous continuously compounded risk-free rate. In the following, it is assumed that

**Table I. Measured Corporate Yield Spreads**

This table reports the average corporate spot yield spreads for industrial Aa, A, and Baa corporate bonds for maturities from 2 to 10 years. Corporate bond spreads are calculated as the difference between the corporate spot rates and Treasury spot rates for a given maturity. Spot rates were computed using the Nelson and Siegel (1987) model. The first panel contains the average Treasury spot rates and corporate yield spreads over the entire 10-year period of our sample. The second panel illustrates the averages for the first five years of our sample while the third panel displays the averages for the second five years.

<b>Years to Maturity</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>	<b>8</b>	<b>9</b>	<b>10</b>
<i>1987-1996</i>									
Average Treasury yields (%)	6.454	6.709	6.920	7.090	7.226	7.337	7.426	7.500	7.562
Average Aa yield spread (%)	0.413	0.416	0.447	0.477	0.502	0.526	0.548	0.569	0.590
Average A yield spread (%)	0.612	0.672	0.722	0.752	0.769	0.776	0.779	0.778	0.776
Average Baa yield spread (%)	1.180	1.206	1.229	1.237	1.234	1.224	1.210	1.193	1.174
<i>1987-1991</i>									
Average Treasury yields (%)	7.601	7.775	7.928	8.054	8.157	8.241	8.309	8.364	8.410
Average Aa yield spread (%)	0.512	0.495	0.515	0.545	0.579	0.617	0.656	0.697	0.738
Average A yield spread (%)	0.737	0.802	0.845	0.869	0.882	0.889	0.893	0.895	0.896
Average Baa yield spread (%)	1.421	1.400	1.402	1.400	1.391	1.379	1.363	1.346	1.328
<i>1992-1996</i>									
Average Treasury yields (%)	5.306	5.643	5.912	6.126	6.296	6.433	6.544	6.636	6.713
Average Aa yield spread (%)	0.315	0.336	0.379	0.409	0.425	0.434	0.439	0.441	0.442
Average A yield spread (%)	0.487	0.543	0.599	0.635	0.655	0.664	0.665	0.661	0.655
Average Baa yield spread (%)	0.939	1.012	1.056	1.074	1.077	1.070	1.057	1.040	1.019

**Table II. Average Root Mean Squared Errors**

This table presents the average root mean squared errors obtained from the difference between theoretical bond prices computed using the Nelson and Siegel (1987) model and the actual bond prices for Treasuries and industrial Aa, A, and Baa corporate bonds. The estimation procedure is described in Section II. Root mean squared error is measured in cents per dollar. For a given class of bonds, the root mean squared error is calculated once per period (month). The number reported is the average of all root mean squared errors within a given class over the months of the corresponding period.

	<b>Treasuries</b>	<b>Aa</b>	<b>A</b>	<b>Baa</b>
1987-1996	0.220	0.525	0.812	1.458
1987-1991	0.304	0.555	0.876	1.387
1992-1996	0.136	0.496	0.748	1.529

1. There exists a martingale measure  $Q$  under which the discounted value of any risk-free, zero coupon bond is a martingale.
2. In case of default, a constant fraction  $\rho$  of the market value of an equivalent risky bond is recovered at the default time.
3. Under the martingale measure  $Q$ , the default time intensity is driven by a time-homogeneous Markov process  $X$  describing the credit rating migrations of the firms. This Markov process  $X$  is characterized by the generator matrix  $\Lambda$  and we assume that  $\Lambda$  is diagonal.

In this context, Appendix A illustrates that the intensity can be written as

$$\lambda_t = \frac{\sum_{k=1}^m a_k d_k \exp(d_k t)}{1 - \sum_{k=1}^m a_k \exp(d_k t)}, \quad (1)$$

where the constants  $d_1, \dots, d_m$  are the eigenvalues associated with the generator matrix  $\Lambda$  and the constants  $a_1, \dots, a_m$  are functions of the components of the eigenvectors of  $\Lambda$  and are described explicitly in Appendix A.

4. Investors are risk neutral with respect to default risk.

Assumption 1 is necessary to price a bond as its expected discounted payoff. Assumption 2 is as in Duffie and Singleton (1999). Assumption 3 links the default intensity to the credit rating migration's generator. Therefore, the default time of high-rated bonds reflects the downgrade risk that is the main source of risk for this type of bond. Finally, Assumption 4, which implies that the distribution of the default time  $\tau$  will remain the same under the empirical probability measure  $P$  and the martingale measure  $Q$ , is required to allow the use of databases containing information about default probabilities in our empirical analysis. Under these assumptions, the time  $t$  value of a corporate zero coupon bond paying one dollar at time  $T$  is:

$$\tilde{P}(t, T) = P(t, T) \exp\left(- (1 - \rho) \int_t^T \lambda_s ds\right), \quad (2)$$

where  $P(t, T)$  is the price of a risk-free, zero coupon bond. This result is a special case of the Duffie and Singleton (1999) approach. A derivation of the bond price equation can be found in Appendix B. Given this pricing equation, the corporate yield spread curve at time  $t$  is given by

$$S(t, T) = \frac{\ln P(t, T)}{T - t} - \frac{\ln \tilde{P}(t, T)}{T - t} = \frac{1 - \rho}{T - t} \int_t^T \lambda_s ds. \quad (3)$$

The spreads can then be computed using the following discrete approximation of Equation (3)

$$\frac{1 - \rho}{T - t} \int_t^T \lambda_s ds \cong \frac{1 - \rho}{n} \sum_{j=1}^n \hat{\lambda}_j \Delta t, \quad (4)$$

where  $\Delta t = (T - t)/n = 10^{-6}$  and  $\hat{\lambda}_j$  is the estimated default intensity process.

### III. Generator Estimation

The corporate yield spread's model proposed in the previous section requires the estimation of a generator since such a quantity appears in the construction of Intensity (1). This section describes the different methodologies that may be used to obtain such estimates.

One approach to the estimation of default probabilities imposes little structure on the data; it consists of forming a cohort at one point in time and counting the defaults after one period, two periods, and so on. The drawback of such an approach stems from the large standard errors associated with the estimates. Generating accurate estimates requires the observation of many defaults, an unlikely possibility when working with investment grade bonds. For such a case, many estimated probabilities would simply be zero. This approach would also make it difficult to include the information provided by new firms entering the database and would not capture the downgrade risk. See Appendix C for the data description used for transition matrix estimation.

Another approach found in the literature uses estimates of periodic transition matrices available from Moody's or Standard & Poor's via the cohort method of Carty and Fons (1993) and Carty (1997). The transitions from one credit rating class to another are counted and estimates of transition probabilities are calculated using the number of bonds in the cohort at the beginning of the period. Probabilities of defaulting for more than one period can then be conveniently computed from this transition matrix using simple matrix multiplications. This convenience comes at the cost of imposing a Markovian structure on the data and it is not clear whether such a structure will hold. As with the preceding approach, there are also several drawbacks associated with such estimates of default probabilities. Defaults and rating transitions are rare events, and these transition matrices still contain many cells with estimated probabilities equal to zero. This may lead to an underestimation of the default spread. Again, as with the preceding approach, if one builds confidence intervals around these estimates, the results turn out to be unsatisfactory. With a small sample size, the default spread could be misestimated due to large sampling errors.

Lando and Skodeberg (2002) have suggested estimating a Markov-process generator rather than a one-year transition matrix. Such a generator can then be used to compute transition matrices for any desired horizon. As with the cohort approach, this method also imposes a Markovian structure. Lando and Skodeberg (2002) have demonstrated that this continuous-time analysis of rating transitions using generator matrices improves the estimates of rare transitions even when they are not observed in the data, a result that cannot be obtained with the discrete-time analysis of Carty and Fons (1993) and Carty (1997). A continuous-time analysis of defaults permits estimates of default probabilities even for cells that have no defaults. This is possible because the approach draws on information in the transition from one class to another to infer better estimates of the default probabilities. Finally, as shown in Christensen, Hansen, and Lando (2004), inference in such a framework is informative and can be conveniently computed.<sup>2</sup>

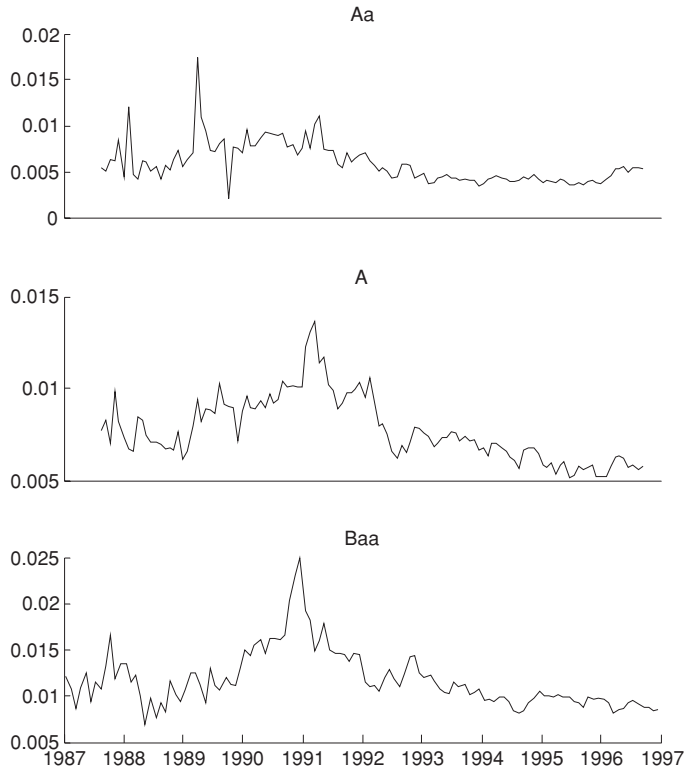
As just discussed, the generator may be estimated using raw data on the timing of credit migration. We use this approach herein under the label of "continuous-time generator." However, for the sake of comparison with the widely used cohort approach, we must construct a generator estimate from transition probability matrices obtained with the cohort estimation approach. As shown in Israel, Rosenthal, and Wei (2001), the existence of such a generator for a given transition probability matrix is not guaranteed. However, as proposed by these authors, a solution to this problem is to obtain a generator that will produce a transition matrix close to the original transition matrix. Therefore, we use the procedure suggested in Israel, Rosenthal, and Wei (2001) to verify the existence and obtain the underlying generators for the transition matrices that will be used in our empirical analysis. Using these generators, we then compute the intensities with Equation (1). We label this approach "cohort."

---

<sup>2</sup>Other recent references regarding estimating transition matrices and the resulting inferences issues are Jafry and Schuermann (2004) and Hanson and Schuermann (2006).

### Figure 1. Ten-Years-to-Maturity Corporate Yield Spread Levels

This figure illustrates the 10-years-to-maturity zero coupon yield spread levels estimated with the Nelson and Siegel (1987) approach on samples of corporate coupon bonds from January 1987 to December 1996.



## IV. Empirical Findings

### A. Sample Period

A first key issue associated with estimating transition and default probabilities is the choice of the estimation period. As discussed in Moeller and Molina (2003), there can be substantial variations in default probability through time. Although we do not observe default probabilities, we can observe substantial variations in spreads over time. These variations can be ascribed to changes in expected recovery rates, liquidity, or risk premia, but also to changes in default probabilities. Figure 1 plots the times series of empirical yield spreads for Aa, A, and Baa industrial bonds with 10 years to maturity. Given the wide variation in the spread level over time, it is not clear that using a long history of past data to assess the probabilities of defaulting is the best approach for our purposes. With the model described in Section II, a constant term structure of default probabilities will generate a constant credit spread. A long history of default data updated regularly will most likely produce term structures of default probabilities and credit spreads that will be fairly constant over time. This would be at odds with the substantial time

**Table III. Correlations between the Yield Spreads and Estimated Default Spreads for Varying Sampling Window Lengths**

This table reports the correlations between the 10-years-to-maturity estimated default spread, Equation (3), and the corresponding Nelson and Siegel (1987) yield spread. The generators have been estimated using the cohort method and several sampling window lengths. Recovery rates are 59.59% for Aa, 60.63% for A, and 49.42% for Baa. These rates are from Altman and Kishore (1998) and were used in Elton et al. (2001).

Window	1 Year	2 Years	3 Years	5 Years	7 Years	10 Years
<i>Industrial Aa Bonds</i>						
1987-1996	0.12	0.18	0.08	-0.23	-0.67	-0.75
1987	-0.94	-0.94	-0.94	-0.94	-0.94	-0.94
1988-1996	0.32	0.32	0.12	-0.20	-0.72	-0.82
<i>Industrial A Bonds</i>						
1987-1996	0.36	0.39	0.34	0.27	-0.45	-0.70
1987	-0.99	-0.99	-0.99	-0.99	-0.99	-0.99
1988-1996	0.69	0.48	0.36	0.30	-0.50	-0.78
<i>Industrial Baa Bonds</i>						
1987-1996	0.34	0.14	-0.04	-0.30	-0.62	-0.59
1987	-0.97	-0.97	-0.97	-0.97	-0.97	-0.97
1988-1996	0.50	0.18	-0.02	-0.28	-0.69	-0.64

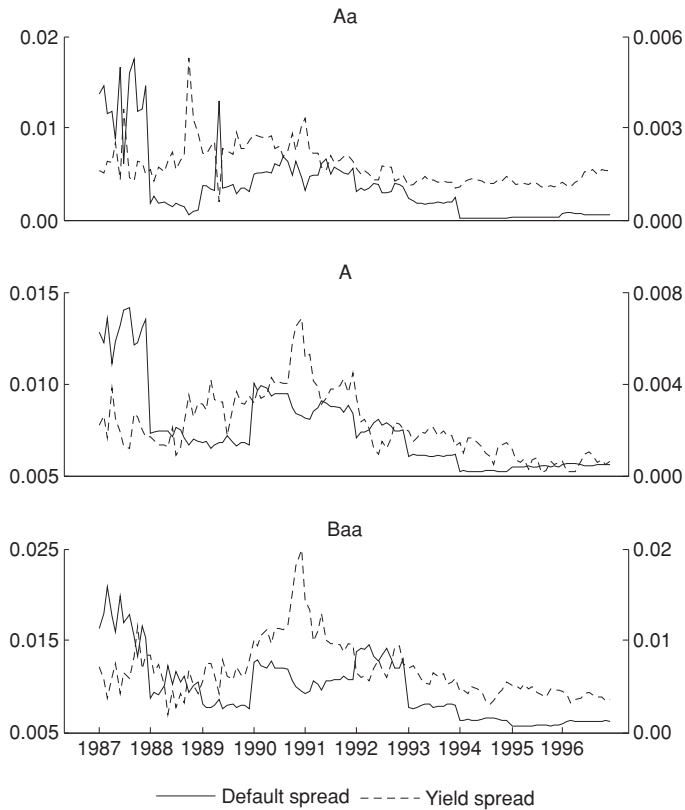
variations observed in spreads. Here, we adopt the view that the most recent ex ante credit migration and default history is perhaps a more valid indicator of the subjective probability of defaulting used by investors to determine the proper yield for bonds in the various credit classes. Thus, we assume that, at a given year, economic agents use the most recent rating transition data to form their anticipation regarding survival and default probabilities for various horizons. The default probabilities are estimated using a rolling window approach to frame new transition and default probabilities each year. For example, with a one-year window, the default proportions for each month in 1987 would be assessed with default data from January 1986 to December 1986.

With such an approach, the length of the window is an important consideration. To provide some guidance about what a proper length should be, Table III displays the sample correlations between yield spreads and estimated default spreads obtained with our model and various window lengths. In this table, the time series of estimated default spreads are computed with transition matrices estimated with the cohort approach. For the whole sample, we see that short window lengths are associated with positive but modest correlations. A detailed look of the data indicates that these low correlations are primarily due to the high negative correlation in the first year of the sample. Removing these first 12 observations realizes correlations of 0.32, 0.69, and 0.50 for Aa, A, and Baa bonds with a one-year window. These correlations then decrease as the window length is increased and actually become negative with longer window lengths. Figures 2 and 3 plot the time series of yield spreads and estimated default spreads on a two scale graph for one-year and 10-year window length cases. As illustrated in these graphs, a short-window length appears to be in better agreement with the yield spreads than a long one. Although the one-year window length obtains higher correlations, it is still important to look at how different window lengths affect the estimated default proportions. Thus, we will analyze the results with window lengths of one, two, and three years.

To assess how different treatment of default data impacts the estimated proportions of credit spreads, a benchmark case is required. Table IV demonstrates the estimated proportions for such

**Figure 2. One-Year Window Length**

This figure displays on two scale graphs the yield spreads and the estimated default spreads for 10-years-to-maturity zero coupon bonds from January 1987 to December 1996. For each year, the estimated spreads are estimated with default probabilities based on a one-year window of out of sample default data and the cohort approach.

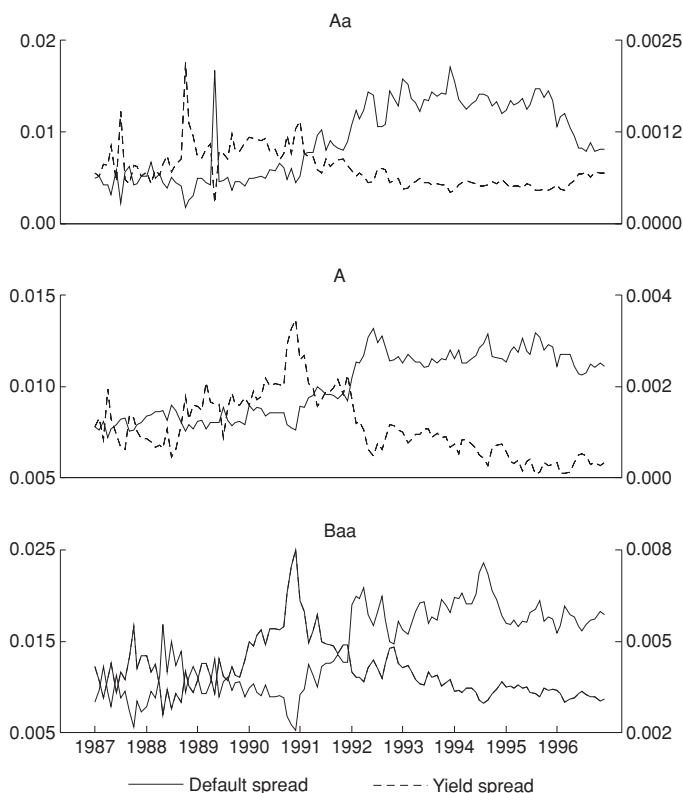


a case, computed with a constant transition probability matrix and recovery rates as in Elton et al. (2001). This transition matrix is the one used in their analysis and it was estimated by applying the cohort approach to Moody's default data for the period 1970-1993. Although their pricing model is different from ours since it deals with coupon bonds and a different theoretical recovery assumption, the results are almost identical. The estimated proportions with our model are 5%, 12%, and 35% for 10-years-to-maturity Aa, A, and Baa bonds whereas the proportions in Elton et al. (2001) model are 5%, 12%, and 37%. This suggests that the results presented next cannot be attributed to differences in our modeling approach or recovery assumptions.

Table V presents the results obtained with the time-varying probabilities' term structure computed with the window approach described previously in this section. As with the previous table, the transition matrices are estimated using a cohort approach. Window lengths of one, two, and three years are considered. For two-years-to-maturity bonds, the proportions are roughly doubled for the 1987-1991 period for all credit classes and window lengths. For the 10-years-to-maturity

**Figure 3. Ten-Year Window Length**

This figure demonstrates on two scale graphs the yield spreads and the estimated default spreads for 10-years-to-maturity zero coupon bonds from January 1987 to December 1996. For each year, the estimated spreads are estimated with default probabilities based on a 10-year window of out of sample default data and the cohort approach.



case, the proportions are also roughly doubled except for the Baa case that progresses from 35% in Table IV to numbers around 47% for this case. Results are also presented for the first and second halves of our sample (the 1987-1991 and 1992-1996 subperiods). As seen in the table, the proportions vary substantially across subsamples and window lengths. For the first portion of the sample, a shorter window length produces higher estimated default proportions, while the reverse situation occurs for the second part of the sample. This can be explained by looking at the estimated term structure of default probabilities shown in Table VI. Comparing the estimates for different window lengths confirms that, for the second half of the sample, a longer window length tends to include years with many defaults, producing high estimates of default probabilities. If investors give greater weight to the information provided by the more recent default history when forming their expectations, it is not clear that the results obtained with a longer window length, such as three years, will be relevant. Nevertheless, they are indicative of the sensitivity of the estimated proportions to different sample periods for the default data.

**Table IV. Average Default Spread Proportions: Constant Moody's Cohort Transition Matrix**

This table reports the average of 2- and 10-years-to-maturity default spread proportions obtained with the Moody's matrix reported in Elton et al. (2001) and used over the entire 10-year period. This matrix is estimated with Moody's (2005) data over the 1970-1993 period. Recovery rates, from Altman and Kishore (1998), are 59.59% for Aa, 60.63% for A, and 49.42% for Baa. Averages are computed for the entire 10-year period and the two 5-year subperiods of our sample.

	<b>2 Years</b>	<b>10 Years</b>
	<i>Proportions (%) Aa</i>	
1987-1996	0.55	4.97
1987-1991	0.43	3.90
1992-1996	0.66	6.03
	<i>Proportions (%) A</i>	
1987-1996	2.27	11.79
1987-1991	1.82	10.04
1992-1996	2.72	13.54
	<i>Proportions (%) Baa</i>	
1987-1996	9.68	34.78
1987-1991	7.37	31.03
1992-1996	12.00	38.53

Another of our model's input, the recovery rate (which was assumed to be constant), is seen to vary greatly across time. Figure 4 plots the average recovery rates obtained from Moody's (2005). These rates are defined as the ratio of the defaulted bond's market price to its face value, as observed 30 days after the default date, for all bonds irrespective of their rating. The average recovery rates vary significantly across time. They range from a high of 62% to a low of 28%. The average recovery rate during the 1987-1991 subperiod is equal to 40.8%, while that of the 1992-1996 subperiod is equal to 45.8%. It is also documented in Moody's (2005) that the recovery rates are even lower for industrial bonds. Since these recovery rates are for all bond ratings, they can be interpreted as the recovery rates of bonds with an average risk. Thus, they should be a good approximation of the expected recovery rates of Baa ratings, a category falling between the high-quality investment grade bonds (like Aaa, Aa, and A) and the speculative grades (like Ba, B, and Caa-C). Table VII displays the average proportions obtained for Baa bonds using these time-varying recovery rates. Again, these rates are used ex ante. Thus, to derive the 1987 average default proportion, the average recovery rate estimated for 1986 was used. Using these time-varying recovery rates does affect the results. For example, in the case of the one-year window with 10 years to maturity, the proportions are 47%, 64%, and 29% (Table V) for the whole sample and while the two subsamples increase to 53%, 74%, and 33% (Table VII). The effect is similar but less pronounced for the two years to maturity case.

## B. Generator's Estimation

As mentioned in the introduction, estimating the transition matrices, generators, and default probabilities used to measure the proportion of the spread from default data involves choosing a particular statistical approach. It is not clear that the results for default proportions are invariant to these different approaches. We have already used the cohort approach in the previous subsection.

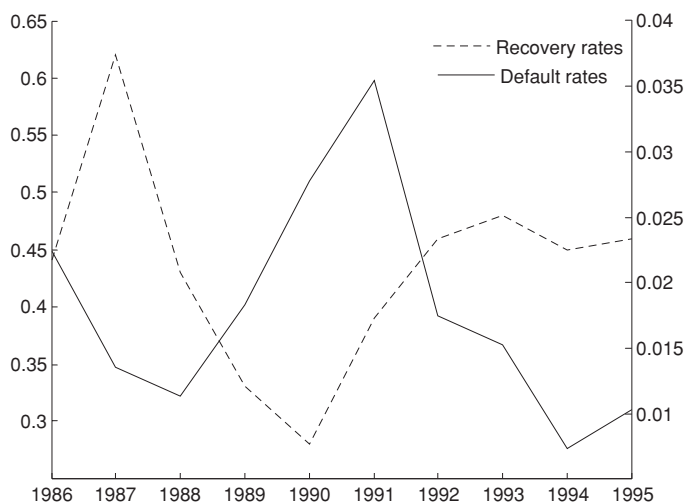
**Table V. Average Default Spread Proportions: Moody's Time-Varying Cohort Transition Matrices**

This table reports the average of 2- and 10-years-to-maturity default spread proportions for generators estimated with the cohort approach and a rolling window of 1-, 2-, and 3-year lengths of ex ante default data. Recovery rates, from Altman and Kishore (1998), are 59.59% for Aa, 60.63% for A, and 49.42% for Baa. Averages are computed for the entire 10-year period and the two 5-year subperiods of our sample.

Window	2 Years to Maturity			10 Years to Maturity		
	1 Year	2 Years	3 Years	1 Year	2 Years	3 Years
<i>Proportions in %: Industrial Aa Bonds</i>						
1987-1996	1.45	1.44	1.48	10.88	11.14	11.49
1987-1991	2.61	2.41	2.16	17.76	16.47	14.60
1992-1996	0.29	0.46	0.81	3.99	5.80	8.38
<i>Proportions in %: Industrial A Bonds</i>						
1987-1996	5.42	4.93	4.77	20.02	20.40	20.64
1987-1991	9.43	7.86	6.72	31.83	29.43	26.00
1992-1996	1.42	1.99	2.81	8.21	11.37	15.29
<i>Proportions in %: Industrial Baa Bonds</i>						
1987-1996	16.89	16.39	18.09	46.56	47.68	48.93
1987-1991	23.53	21.83	21.79	64.23	62.22	55.70
1992-1996	10.26	10.95	14.39	28.90	33.14	42.16

**Figure 4. Annual Defaulted Bond Recovery Rates and Simulated Average Default Rates**

This figure illustrates the time series of average recovery rates obtained from Moody's (2005) and defined as the ratio of the defaulted bond's market price observed 30 days after its default date to its face value (par) for all bonds. Also shown are the default rates obtained from our simulation procedure (Section E) used to assess confidence intervals around our estimated default proportions.



**Table VI. Estimated Term Structure of Default Probabilities: Moody's Time-Varying Cohort Transition Matrices**

Estimated term structure of default probabilities for horizons of 1 to 10 years and window lengths of 1, 2, and 3 years of ex ante default data.

	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996
<i>Default Probabilities in %: 1-Year Window</i>										
1	1.55	0.15	0.07	0.59	0.23	0.59	0.27	0.02	0.01	0.03
2	3.38	0.56	0.29	1.65	0.85	1.53	0.63	0.09	0.05	0.10
3	5.46	1.20	0.66	3.13	1.81	2.80	1.08	0.21	0.13	0.22
4	7.76	2.03	1.17	4.96	3.07	4.39	1.61	0.39	0.23	0.38
5	10.23	3.02	1.80	7.08	4.60	6.24	2.21	0.62	0.37	0.57
6	12.84	4.14	2.54	9.43	6.37	8.32	2.88	0.91	0.54	0.80
7	15.53	5.37	3.38	11.96	8.33	10.57	3.61	1.24	0.74	1.05
8	18.27	6.68	4.29	14.61	10.45	12.95	4.38	1.60	0.97	1.32
9	21.04	8.06	5.28	17.34	12.69	15.42	5.18	2.00	1.22	1.62
10	23.81	9.50	6.32	20.12	15.02	17.94	6.02	2.42	1.50	1.94
<i>Default Probabilities in %: 2-Year Window</i>										
1	0.81	0.77	0.10	0.30	0.29	0.30	0.41	0.13	0.02	0.02
2	1.87	1.82	0.41	0.90	1.07	0.93	1.00	0.33	0.07	0.07
3	3.16	3.11	0.89	1.77	2.30	1.90	1.77	0.60	0.17	0.15
4	4.64	4.60	1.54	2.88	3.91	3.18	2.71	0.94	0.31	0.27
5	6.29	6.26	2.33	4.20	5.85	4.75	3.81	1.34	0.49	0.42
6	8.07	8.04	3.24	5.69	8.07	6.56	5.03	1.80	0.72	0.60
7	9.95	9.93	4.25	7.31	10.49	8.57	6.37	2.31	0.98	0.81
8	11.91	11.89	5.34	9.05	13.07	10.74	7.80	2.85	1.28	1.06
9	13.93	13.91	6.51	10.87	15.77	13.03	9.29	3.43	1.61	1.32
10	15.99	15.96	7.73	12.76	18.53	15.41	10.84	4.04	1.97	1.62
<i>Default Probabilities in %: 3-Year Window</i>										
1	0.66	0.53	0.51	0.21	0.20	0.38	0.28	0.27	0.09	0.02
2	1.49	1.32	1.25	0.70	0.75	1.18	0.80	0.66	0.22	0.08
3	2.50	2.34	2.18	1.46	1.64	2.38	1.55	1.18	0.41	0.18
4	3.66	3.55	3.30	2.44	2.83	3.93	2.52	1.83	0.66	0.31
5	4.95	4.92	4.57	3.61	4.27	5.79	3.70	2.59	0.96	0.49
6	6.36	6.42	5.97	4.95	5.93	7.90	5.05	3.45	1.31	0.71
7	7.87	8.03	7.47	6.42	7.77	10.22	6.55	4.39	1.70	0.96
8	9.45	9.71	9.06	8.00	9.74	12.69	8.18	5.40	2.13	1.24
9	11.09	11.46	10.71	9.66	11.82	15.27	9.90	6.46	2.60	1.55
10	12.76	13.24	12.40	11.38	13.97	17.92	11.69	7.57	3.09	1.88

The goal here is to see whether continuous-time estimation of the generator produces similar results.

Table VIII presents the results of the time-varying probabilities calculated with the window approach now using generators estimated with the continuous-time approach. When compared with Table V, we find only marginal impact in most cases. Our earlier results, which were obtained by applying the cohort method to small sample sizes, might have inherited the large sampling errors associated with this approach. We find here that the generator approach, which has been found to have better statistical properties, brings similar results and confirms our previous findings.

**Table VII. Average Default Spread Proportions: Moody's Time-Varying Cohort Transition Matrices and Recovery Rates**

This table reports the average of 2- and 10-years-to-maturity default spread proportions for generators estimated with the cohort approach and a rolling window of 1-, 2-, and 3-year lengths of ex ante default data. Time-varying recovery rates, obtained from Moody's (2005) database, have been used here and are defined as the ratio of the defaulted bond's market price observed 30 days after its default date to its face value (par) for all bonds.

Window	2 Years to Maturity			10 Years to Maturity		
	1 Year	2 Years	3 Years	1 Year	2 Years	3 Years
<i>Proportions in %: Industrial Baa Bonds</i>						
1987-1996	19.65	17.84	19.83	53.44	52.88	54.16
1987-1991	27.58	23.58	23.82	73.79	68.56	61.29
1992-1996	11.72	12.09	15.84	33.09	37.20	47.04

**Table VIII. Average Default Spread Proportions: Moody's Time-Varying Generators**

This table reports the average of 2- and 10-years-to-maturity default spread proportions for estimated generators computed using the continuous-time approach and a rolling window of 1-, 2-, and 3-year lengths of ex ante default data. Recovery rates, from Altman and Kishore (1998), are 59.59% for Aa, 60.63% for A, and 49.42% for Baa. Averages are computed for the entire 10-year period and the two 5-year subperiods of our sample.

Window	2 Years to Maturity			10 Years to Maturity		
	1 Year	2 Years	3 Years	1 Year	2 Years	3 Years
<i>Proportions in %: Industrial Aa Bonds</i>						
1987-1996	1.61	1.94	2.22	8.94	10.37	11.18
1987-1991	3.06	3.70	3.40	14.87	16.61	15.13
1992-1996	0.16	0.18	1.03	3.01	4.13	7.24
<i>Proportions in %: Industrial A Bonds</i>						
1987-1996	3.11	3.20	3.37	15.78	17.80	19.12
1987-1991	5.12	5.02	4.67	25.03	26.48	24.30
1992-1996	1.10	1.38	2.07	6.53	9.11	13.95
<i>Proportions in %: Industrial Baa Bonds</i>						
1987-1996	12.46	14.19	16.64	42.76	48.80	52.84
1987-1991	15.20	16.28	16.49	59.50	63.05	58.83
1992-1996	9.73	12.09	16.80	26.02	34.55	46.84

### C. Data Filtering

Here we discuss the impact of the data-filtering process and the information considered when estimating transition matrices and generators. Such an analysis is important for financial institutions that are building their own internal rating system for Basel II and for the regulators who will have to monitor these systems.

When working with default databases, one must deal with issuers' movements in the database. For instance, a decision must be made regarding whether to consider issuers that are present at the beginning of the estimation period but who leave for reasons other than default. Such cases

**Table IX. Sensitivity to Data Filtering**

This table reports the number of firms, transitions, and defaults used to estimate the continuous-time transition matrices. There are three cases: 1) exclusion of withdrawals and exclusion of entry firms, 2) inclusion of withdrawals and inclusion of entry firms, and 3) inclusion of withdrawals and exclusion of entry firms.

	Excluding Withdrawals and Entry			Including Withdrawals and Entry			Including Withdrawals and Excluding Entry		
	87-96	87-91	92-96	87-96	87-91	92-96	87-96	87-91	92-96
Issuers	1,239	1,539	1,432	3,879	2,656	3,090	1,977	1,977	1,867
Rating observations	2,731	2,672	2,236	7,652	4,690	4,829	4,590	3,667	3,213
Defaults	250	196	92	399	267	132	250	196	92
Defaults/ issuers	20.18%	12.74%	6.42%	10.29%	10.05%	4.27%	12.65%	9.91%	4.93%

will be referred to here as withdrawn ratings (or right censoring). Another factor is whether or not to consider issuers that enter the database after the starting date of estimation. These cases will be referred to as entry firms. Excluding withdrawn ratings and entry firms is more in the spirit of Moody's (2005) standard cohort analysis that also produces statistics including withdrawals (right censoring).

To demonstrate the impact of these decisions on the data set used to estimate a generator, Table IX examines the data composition with respect to the three filtering alternatives. First, we exclude entry firms and withdrawn rating data. Second, we include entry firms and right censored data. Finally, we exclude entry firm data, but include withdrawals. The analysis was done for the period 1987-1996 and for the 1987-1991 and 1992-1996 subperiods. We observe, from Table IX, that the proportions of default issuers (defaults/issuers) vary substantially when the filtering approach is varied. For example, when compared with the case of entry and withdrawal exclusion, this proportion decreases when including withdrawals and entry firm data. These differences in proportions may affect the estimated generators and default probabilities. Thus, a sensitivity analysis of the impact of these issues on corporate default proportions is done here.

Table X presents a sensitivity analysis of the data-filtering procedure. As the results indicate, important differences are observed. For Baa-rated firms, the 10-years-to-maturity default spread proportion goes from 42% to 53%. The case of excluded withdrawn ratings and entry data report higher default proportions. A detailed examination of the results also concludes that these are essentially caused by higher estimates of default probabilities. We observe, from Table XI, that the number of defaults is the same in the first and third cases, while the numbers of issuers and rating observations are higher in the third case. Inclusion of the withdrawals reduces default probabilities and default risk proportions in yield spreads. The same conclusion is obtained when entry firms are added. Default risk proportions and implied default probabilities are even lower.

#### D. Inference

As argued in the introduction, inference is another important issue associated with computing default proportions. Because defaults are rare events, default probabilities and recovery rates are typically estimated with much uncertainty. Therefore, statistical confidence intervals should be reported to allow meaningful interpretations of the point estimates. We propose a procedure based

**Table X. Average Default Spread Proportions: Sensitivity to Data Filtering**

This table reports the average of 2- and 10-years-to-maturity default spread proportions for generators estimated with the continuous-time approach and a rolling window. Three cases are examined for each maturity; 1) wre & efe: withdrawn ratings exclusion and entry firms exclusion; 2) wri & efe: withdrawn ratings inclusion and entry firms exclusion; and 3) wri & efi: withdrawn ratings inclusion and entry firms inclusion. Recovery rates, from Altman and Kishore (1998), are 59.59% for Aa, 60.63% for A, and 49.42% for Baa.

	2 Years to Maturity			10 Years to Maturity		
	wre & efe	wri & efe	wri & efi	wre & efe	wri & efe	wri & efi
<b>Window Length: 1 Year</b>						
	<i>Proportions (in %): Industrial Aa Bonds</i>					
1987-1996	1.61	1.55	1.37	8.94	8.94	8.01
1987-1991	3.06	2.94	2.59	14.87	14.93	13.34
1992-1996	0.16	0.16	0.15	3.01	2.95	2.68
	<i>Proportions (in %): Industrial A Bonds</i>					
1987-1996	3.11	3.09	2.71	15.78	15.80	14.25
1987-1991	5.12	5.11	4.45	25.03	25.37	22.75
1992-1996	1.10	1.06	0.97	6.53	6.22	5.75
	<i>Proportions (in %): Industrial Baa Bonds</i>					
1987-1996	12.46	12.42	11.51	42.76	42.35	38.81
1987-1991	15.20	15.44	14.11	59.50	59.89	54.24
1992-1996	9.73	9.39	8.91	26.02	24.81	23.38
<b>Window Length: 2 Years</b>						
	<i>Proportions (in %): Industrial Aa Bonds</i>					
1987-1996	1.94	1.81	1.52	10.37	10.17	8.76
1987-1991	3.70	3.45	2.88	16.61	16.40	13.92
1992-1996	0.18	0.18	0.16	4.13	3.95	3.59
	<i>Proportions (in %): Industrial A Bonds</i>					
1987-1996	3.20	3.12	2.64	17.80	17.45	15.09
1987-1991	5.02	4.95	4.15	26.48	26.37	22.42
1992-1996	1.38	1.29	1.13	9.11	8.52	7.75
	<i>Proportions (in %): industrial Baa bonds</i>					
1987-1996	14.19	13.65	12.06	48.80	46.74	41.17
1987-1991	16.28	15.83	13.48	63.05	61.06	52.10
1992-1996	12.09	11.47	10.63	34.55	32.43	30.24
<b>Window Length: 3 Years</b>						
	<i>Proportions (in %): Industrial Aa Bonds</i>					
1987-1996	2.22	1.97	1.72	11.18	10.65	9.37
1987-1991	3.40	3.07	2.68	15.13	14.77	12.89
1992-1996	0.87	0.87	0.77	7.24	6.54	5.84
	<i>Proportions (in %): Industrial A Bonds</i>					
1987-1996	3.37	3.17	2.62	19.12	18.08	15.46
1987-1991	4.67	4.49	3.65	24.30	23.62	19.74
1992-1996	2.07	1.85	1.60	13.95	12.54	11.17
	<i>Proportions (in %): Industrial Baa Bonds</i>					
1987-1996	16.64	15.43	13.30	52.84	48.74	42.48
1987-1991	16.49	15.45	12.61	58.83	55.13	46.26
1992-1996	16.80	15.40	13.98	46.84	42.36	38.69

**Table XI. Rating Distributions**

This table reports the distribution of issuers, by rating, at the starting date of the estimation period used to construct the confidence sets of average default spread proportions.

	<b>Aaa</b>	<b>Aa</b>	<b>A</b>	<b>Baa</b>	<b>Ba</b>	<b>B</b>	<b>Caa-C</b>
1986	64	253	537	326	351	174	7
1987	65	253	511	328	445	254	11
1988	74	235	511	337	491	308	9
1989	73	214	520	348	492	341	12
1990	75	205	504	338	476	356	14
1991	71	198	495	361	402	316	19
1992	67	176	484	374	357	264	19
1993	56	164	491	393	365	265	19
1994	54	166	526	416	391	313	35
1995	50	176	553	409	386	407	54

on the parametric bootstrap simulation approach described in Christensen, Hansen, and Lando (2004) to compute approximate confidence intervals for the estimated default proportions. The procedure is as follows:

- In a first step, using the one-year cohort transition matrices with which the default spread proportions were assessed in Tables V and VII, we compute the 10 associated generators using the procedure in Israel, Rosenthal, and Wei (2001). These estimated generators are considered the *true* generators governing the data-generating process.
- The second step then repeats, for each year of our sample period, the following procedure. Using the estimated generator derived in the first step and the distribution of issuers at the beginning of the year, we simulate one year of rating history for each issuer (see Appendix D for details about this simulation procedure). A generator is then estimated with the yearly rating histories of all issuers. A term structure of credit spread can then be computed using Equation (3) and the proportions of default in the spread for this year can be assessed. Using the estimated proportions for each of the 10 years, we then compute an average proportion for the whole 10-year period and the two subperiods of 5 years.

The second step is repeated 10,000 times to generate 10,000 estimates of average default proportions in yield spreads. We then compute different statistics (mean, median, and percentiles 2.5 and 97.5 used as our approximate confidence intervals) of the average default proportion for each rating and maturity. Table XI reports the distribution of issuers by rating at the starting date of the years over which our transition matrices were estimated. Two different sets of results are looked at. The first set does not account for the sampling variability of the recovery rates that are considered constant. A second set of results then looks at cases for which variability is introduced for these rates.

Table XII reports the results for the approximate confidence intervals obtained with the simulation approach for the whole sample and the two subperiods of our data set for the constant recovery rate case. The results from this table should be compared to those of Table V for the one-year window circumstance. As should be expected, the averages of the mean default spread proportions are very close to those reported in Table V. The 95% confidence intervals, reported

**Table XII. Approximate Confidence Intervals for Default Spread Proportions with Fixed Recovery Rates**

This table reports the mean, standard error (Std), and percentiles 2.5 and 97.5 (columns lb and ub) of average default spread proportions for 2- and 10-years-to-maturity zero coupon bonds obtained from Monte Carlo simulations using the time-varying cohort matrices from the rolling window approach with a one-year length of ex ante default data. The recovery rates, from Altman and Kishore (1998), are 59.59% for Aa, 60.63% for A, and 49.42% for Baa.

	2 Years to Maturity				10 Years to Maturity			
	Mean	Std	lb	ub	Mean	Std	lb	ub
<i>Proportions in %: Industrial Aa Bonds</i>								
1987-1996	1.44	0.75	0.78	2.74	10.68	1.40	8.18	13.63
1987-1991	2.59	1.15	1.35	4.61	17.44	2.59	12.78	22.93
1992-1996	0.29	0.95	0.08	0.44	3.92	1.03	2.34	6.15
<i>Proportions in %: Industrial A Bonds</i>								
1987-1996	5.44	0.90	3.84	7.37	19.82	1.92	16.17	23.75
1987-1991	9.46	1.72	6.40	13.14	31.51	3.48	25.00	38.72
1992-1996	1.42	0.48	0.61	2.48	8.12	1.56	5.35	11.48
<i>Proportions in %: Industrial Baa Bonds</i>								
1987-1996	16.88	2.70	11.99	22.49	46.15	3.83	38.77	53.88
1987-1991	23.52	4.38	15.87	32.75	63.63	6.30	51.78	76.56
1992-1996	10.23	3.15	4.97	17.14	28.67	4.28	20.74	37.52

under the columns labeled *ub* and *lb* (upper and lower bound), are wide in many cases, especially for the 1987-1991 period for the 10-years-to-maturity bond case.

## E. Recovery Rates

A caveat concerning these measures is the use of constant recovery proportions that are at odds with the wide variations in recovery rates observed over time (see Section A). The results from Table XII might understate the estimated proportions and sampling variability due to the fixed recovery used in the procedure. Indeed, empirical evidence reported in Hu and Perraudin (2002) and Altman et al. (2005) finds that recovery rates are negatively correlated with the proportion of defaulting firms. This negative correlation is likely to result in higher default spreads. In periods in which firms are more likely to default, a smaller recovery will be obtained for a greater number of defaulting firms. In periods where firms are less likely to default, a higher recovery will be obtained, but for a smaller number of defaulting firms. Another way to interpret the effect of this negative correlation is to realize that the expected default loss can be seen as the product of the probability of default and the loss given default (one minus the recovery rate). The expected value of this product of random variables implies a positive covariance that is zero when the recovery is considered nonrandom.

Table XIII reports the results for cases with 1) nonrandom but time-varying recovery rates and 2) random recovery rates. The first panel of this table details the results of the simulation procedure amended with nonrandom, time-varying recovery rates. These rates come from the yearly recovery time series used in Table VII. Again, as argued in Section A, since these rates are measured without rating distinctions, they can be interpreted as the recovery rates of bonds with an average risk and should approximate well the expected recovery rates of Baa ratings, a category falling between the investment and noninvestment grade bonds. The results from this panel should

**Table XIII. Approximate Confidence Intervals for Default Spread Proportions with Random Recovery Rates**

This table reports the mean, standard error (Std), and percentiles 2.5 and 97.5 (columns lb and ub) of average default spread proportions for 2- and 10-years-to-maturity zero coupon bonds obtained from Monte Carlo simulations using the time-varying cohort matrices from the rolling window approach with a one-year length of ex ante default data. The time-varying, nonrandom recovery rates are from Moody's (2005) and are defined as the ratio of the defaulted bond's market price observed 30 days after its default date to its face value (par) for all bonds. The random time-varying recovery rates are computed, for each simulated history of ratings, with  $\ln(\text{Recovery}) = -1.797 - 0.222 \times \ln(\text{DefProp}_t) + u_t$  where  $\text{DefProp}_t$  is the proportion of default in the simulated history for year  $t$  and  $u_t$  is a random  $N(0, 0.2^2)$  variate.

	2 Years to Maturity				10 Years to Maturity			
	Mean	Std	lb	ub	Mean	Std	lb	ub
<i>Prop. in % Baa Bonds: Nonrandom Time-Varying Recovery</i>								
1987-1996	19.02	3.05	13.51	25.34	52.01	4.32	43.70	60.72
1987-1991	26.51	4.94	17.88	36.91	71.71	7.10	58.35	86.28
1992-1996	11.53	3.54	5.60	19.32	32.31	4.82	23.38	42.28
<i>Proportions in % Baa Bonds: Random Recovery</i>								
1987-1996	20.39	3.57	13.93	27.95	55.75	5.72	44.78	67.19
1987-1991	28.17	5.87	18.05	40.76	76.15	9.66	58.05	96.06
1992-1996	12.60	4.05	5.92	21.60	35.35	6.05	24.40	47.98

be compared with those from the one-year window case of Table VII. Again, averages for the mean default spreads are very close to those of Table VII. The recovery rates generating the above results capture some of the negative correlation effects between recovery and default rates. In our simulation framework, this can be measured by looking at the correlation of the recovery rate series with the average proportions of defaulting firms obtained from simulating 10,000 rating histories. These average proportions of defaulting firms obtained for years 1986-1995 are (in %) 2.24, 1.36, 1.13, 1.84, 2.78, 3.55, 1.75, 1.53, 0.73, and 1.03. As illustrated in Figure 4, there is a negative association between the recovery rates and our simulated average proportions of defaulting firms. The correlation of these two series is estimated to be  $-0.51$ .

Although these rates capture some of the negative correlation between the recovery and defaulting firm proportions, they may understate the estimated portion of the spread rewarding the investor for default. Indeed, they are not statistically linked with the proportions of defaulting firms in each of our 10,000 rating histories. Thus, the second panel of Table XIII reports the results of our procedure amended for random recovery rates generated with the following specification:

$$\ln(\text{Recovery}_t) = -1.797 - 0.222 \times \ln(\text{DefProp}_t) + u_t,$$

where  $\text{DefProp}_t$  is the proportion of default firms in the simulated history for year  $t$  and  $u_t$  is a random  $N(0, 0.2^2)$  variate. The coefficients from this relationship are derived from a least square regression of the log recovery series on the above log average default rate series. In Altman et al. (2005), such a specification was found to be best for describing the empirical relationship between the recovery rates and the proportion of defaulting firms. Interestingly, our estimated coefficients are close to those reported in Altman et al. (2005), which are estimated with different data series and sample periods. Their estimates are  $-1.983$  for the intercept and  $-0.293$  for the

slope while their sample period is 1982-2001. Our system is generating average defaulting firm proportions with properties close to the observed defaulting firm proportions used in Altman et al. (2005).<sup>3</sup>

With such a specification, our estimated default spread proportions have increased to 56%, 76%, and 35% for the whole sample and the two subperiods. The confidence intervals are also now wider due to the additional uncertainty brought in by the random recovery rates.

## V. Conclusion

We have revisited the estimation of default risk proportions in corporate yield spreads. Past studies have found that only a small proportion of the spreads can be attributed to default risk. These results do not hold for all periods of our 1987-1996 sample when sensitivity analyses are made with respect to the sample period used to estimate ex ante default and transition probabilities. We find here that the 1987-1991 period corresponds to a high default period, while the 1992-1996 period relates to a low default period. The estimated proportions can reach 76% of the estimated spread for maturities of 10 years for Baa bonds during the 1987-1991 period. We also find that the estimated proportion of default in credit spread is sensitive to changes in recovery rates and to the data filtration approach used to estimate default probabilities. Finally, the sampling variability is estimated to be large in many cases.

These conclusions are important for financial institutions planning to use internal rating systems and for the regulators that will have to monitor these systems. For example, the Basel II accord on banking regulations recommends measuring the required capital for credit risk based on three parameters: 1) exposure at default, 2) probability of default, and 3) loss given default.<sup>4</sup> As our study confirms, the use of probabilities estimated from long histories of default data produces default spreads that are at odds with the observed credit spreads and might underestimate the required capital for default in some periods. Furthermore, in Basel II, the given loss default for a specific risk pool is discussed mostly as a constant parameter over the business cycles. Our results indicate that default spread estimates are sensitive to the use of time-varying loss given defaults (recovery rates) and their correlation with default proportions. This association increases the default component of credit risk and should also be accounted for when computing required capital. Basel II also emphasizes the documentation of the credit risk rating system as related to the internal ratings based approach. This approach requires accurate statistical models based on appropriate data. A participating bank must document that the data are representative of the bank's risks. Our results confirm that the estimated default risk proportion in credit spreads is a function of data filtration for estimating default risk. Therefore, the bank should document its data filtration approach and the appropriate regulatory authority should be able to monitor this approach.

Our study could be extended in several directions by relaxing some of the restrictive assumptions used here. First, the assumption of risk neutrality could be relaxed. Risk-neutral probabilities different from the default probabilities obtained under the objective measure could then be computed. Building confidence intervals around such estimates may produce results that leave

---

<sup>3</sup>A specification in levels instead of logarithms was also estimated and examined in our simulation framework. Again, the coefficients were close to those found in Altman et al. (2005). The estimated default spread proportions and confidence intervals were similar to those from the log specification.

<sup>4</sup>Basel II is mainly concerned with default risk when discussing credit risk capital requirements.

little room for taxes once liquidity premia are taken into account. This would produce results consistent with the vast and successful literature on derivative securities in which the inclusion of taxes has been found to be of little help.

Finally, it should be noted that we have observed substantial increases in the estimated proportion in the first half of our sample only. The results in the 1991-1996 low-default period confirm that a small proportion of the spread is attributable to the default risk. ■

## Appendix A: Intensity under Assumption (iii)

If the generator matrix  $\Lambda$  is diagonalizable, then one can write  $\Lambda = \mathbf{P}\mathbf{D}\mathbf{P}^{-1}$  where the columns of the matrix  $\mathbf{P}$  contain the eigenvectors of  $\Lambda$ , and  $\mathbf{D} = (d_i)$  is a diagonal matrix filled with the eigenvalues of  $\Lambda$ . Let  $\mathbf{Q}_t = (Q[X_t = j | X_0 = i])_{i,j=1,\dots,m}$  denote the transition matrix of the Markov process  $X$ . Then:

$$\begin{aligned} \mathbf{Q}_t &= \exp(\Lambda t) = \sum_{k=1}^{\infty} \frac{(\Lambda t)^k}{k!} = \sum_{k=1}^{\infty} \frac{\mathbf{P}\mathbf{D}^k \mathbf{P}^{-1} t^k}{k!} = \mathbf{P} \exp(\mathbf{D}t) \mathbf{P}^{-1} \\ &= \left( \sum_{k=1}^m p_{ik} \exp(d_k t) p_{kj}^{-1} \right)_{i,j=1,\dots,m}, \end{aligned}$$

where  $p_{ij}$  are the components of  $\mathbf{P}$ ,  $p_{ij}^{-1}$  are the components of  $\mathbf{P}^{-1}$ , and the first equality is justified by the definition of the generator of a time-homogenous Markov process. Let  $\tau_i$  be the default time of a firm initially rated  $i$  and note that the default state corresponds to state  $m$ . The cumulative distribution of  $\tau_i$  is  $Q[\tau_i \leq t] = Q[X_t = m | X_0 = i]$ . Therefore, the intensity associated with  $\tau_i$  is

$$\begin{aligned} \lambda_{i,t} &= \frac{\frac{\partial}{\partial t} Q[X_t = \text{default} | X_0 = i]}{1 - Q[X_t = \text{default} | X_0 = i]} = \frac{\sum_{k=1}^m p_{ik} p_{km}^{-1} d_k \exp(d_k t)}{1 - \sum_{k=1}^m p_{ik} p_{km}^{-1} \exp(d_k t)} \\ &= \frac{\sum_{k=1}^m a_k d_k \exp(d_k t)}{1 - \sum_{k=1}^m a_k \exp(d_k t)}. \end{aligned}$$

## Appendix B: Default Spread Model Derivation

In case of default, the bondholder recovers, at time  $\tau$ , a fraction of the market value of an equivalent bond. The value of the corporate zero coupon bond is expressed as the expectation, under the martingale measure  $Q$ , of its discounted payoff

$$\begin{aligned}
\tilde{P}(t, T) &= E_t^Q[\beta(t, T)1_{\tau>T} + \beta(t, \tau)\rho\tilde{P}(\tau, T)1_{\tau\leq T}] \\
&= E_t^Q\left[\exp\left(-\int_t^T [r(s) + (1 - \rho)\lambda_s] ds\right)\right] \\
&= E_t^Q\left[\exp\left(-\int_t^T r(s) ds\right)\right]\exp\left(-(1 - \rho)\int_t^T \lambda_s ds\right) \\
&= P(t, T)\exp\left(-(1 - \rho)\int_t^T \lambda_s ds\right),
\end{aligned}$$

where the second line is obtained using results from Duffie and Singleton (1999).

## Appendix C: Data Description for Transition Matrix Estimation

The rating transition histories used to estimate the generator are taken from Moody's Corporate Bond Default Database (January 9, 2002). We consider only issuers domiciled in the United States and having at least one senior unsecured estimated rating. We started with 5,719 issuers (in all industry groups) with 46,305 registered debt issues and 23,666 ratings observations. For each issuer, we checked the number of default dates in the Master Default Table (Moody's, January 9, 2002). We obtained 1,041 default dates for 943 issuers during the period 1970-2001. Some issuers (91) had more than one default date. In the rating transition histories, there are 728 withdrawn ratings that are not the last observation of the issuer. These irrelevant withdrawals were eliminated. Ultimately, we obtained 22,938 ratings observations.

The most important and difficult task is to acquire a proper definition of default. In order to compare our results with recent studies, we treat default dates as do Christensen, Hansen, and Lando (2004). First, all of the nonwithdrawn rating observations up to the date of default have typically been unchanged. However, the ratings that occur within a week before the default date were eliminated. Rating changes observed after the date of default were eliminated unless the new rating reached the B3 level (which is a subcategory of the B rating) or higher. The new ratings were related to debt issued after the date of default. In these cases, we treated these ratings as related to a new issuer. It is important to emphasize that the first rating date of the new issuer is the latest date between the date of the first issue after default and the first date we observe an issuer rating higher than or equal to B3. The same treatment is applied for the case of two and three default dates. Finally, a few issuers have a registered default date that falls before the first rating observation in the Senior Unsecured Estimated Rating Table (Moody's, January 9, 2002). In these cases, we considered there to be no default. Using this procedure, we derived 5,821 issuers with 965 default dates. We aggregated all rating notches and ended up with the nine usual ratings Aaa, Aa, A, Baa, Ba, B, Caa-C, Default, and NR (not rated) with 15,564 rating observations.

## Appendix D: Simulation Procedure for the Rating History

The procedure for simulating a rating history of  $T$  years for an issuer is as follows. With the initial rating  $i$ , we simulate  $t_1$ , the waiting time for leaving this state. This waiting time is simulated with an exponential distribution with mean  $\frac{1}{|\lambda_{ii}|}$ , where  $\lambda_{ii}$  are the elements of the generator matrix for  $j = i$ . If  $t_1 > T$ , the issuer stays in its current rating for the whole period

and the simulation ends. If the waiting time is shorter than  $T$  years, we determine the rating class to which it migrates. To do this, the interval  $[0, 1]$  is divided into subintervals with the migration intensities  $\frac{\lambda_{ij}}{|\lambda_{ii}|}$  for all  $j \neq i$ . A uniform random variate between zero and one is then drawn to determine the new rating. If the new rating is not a default (in which case the simulation ends), we then determine whether the company stays with the new rating or migrates (i.e., we draw a waiting time  $t_2$ ) and verify if  $t_1 + t_2 > T$ . If  $t_1 + t_2 > T$ , the simulation ends. If  $t_1 + t_2 < T$ , we have to determine the new rating. The procedure is repeated until  $\sum t_k > T$  or the firm migrates to the default state.

## References

- Altman, E., B. Brady, A. Resti, and A. Sironi, 2005, "The Link between Default and Recovery Rates: Theory, Empirical Evidence and Implications," *Journal of Business* 78, 2203-2228.
- Altman, E. and V. Kishore, 1998, "Defaults and Returns on High Yield Bonds: Analysis through 1997," New York University Salomon Center Working Paper.
- Bangia, A., F. Diebold, A. Kronimus, C. Schagen, and T. Schuermann, 2002, "Ratings Migration and the Business Cycle with Application to Credit Portfolio Stress Testing," *Journal of Banking and Finance* 26, 445-474.
- Carty, L., 1997, "Moody's Rating Migration and Credit Quality Correlation, 1920-1996," Special Comment, Moody's Investors Service, New York, NY.
- Carty, L. and J. Fons, 1993, "Measuring Changes in Credit Quality," *Journal of Fixed Incomes* 4, 27-41.
- Christensen, J., E. Hansen, and D. Lando, 2004, "Confidence Sets for Continuous-Time Rating Transition Probabilities," *Journal of Banking and Finance* 28, 2575-2602.
- Crouhy, M., D. Galai, and R. Mark, 2000, "A Comparative Analysis of Current Credit Risk Models," *Journal of Banking and Finance* 24, 59-117.
- Duffie, D. and K. Singleton, 1999, "Modeling Term Structures of Defaultable Bonds," *Review of Financial Studies* 12, 687-720.
- Elton E., M. Gruber, D. Agrawal, and C. Mann, 2001, "Explaining the Rate Spread on Corporate Bonds," *Journal of Finance* 56, 247-277.
- Gordy, M., 2000, "A Comparative Anatomy of Credit Risk Models," *Journal of Banking and Finance* 24, 119-149.
- Hanson, S. and T. Schuermann, 2006, "Confidence Intervals for Probabilities of Default," *Journal of Banking and Finance* 30, 2281-2301.
- Hu, Y.T. and W. Perraudin, 2002, "The Dependence of Recovery Rates and Defaults," Birkbeck College Working Paper.
- Huang, J. and M. Huang, 2003, "How Much of the Corporate-Treasury Yield Spread Is Due to Credit Risk?" Stanford University Graduate School of Business Working Paper.
- Israel, R., J. Rosenthal, and J. Wei, 2001, "Finding Generators for Markov Chains via Empirical Transition Matrices with Applications to Credit Ratings," *Mathematical Finance* 2, 245-265.
- Jafry, Y. and T. Schuermann, 2004, "Measurement, Estimation and Comparison of Credit Migration Matrices," *Journal of Banking and Finance* 28, 2603-2639.

- Lando, D. and T. Skodeberg, 2002, "Analyzing Rating Transitions and Rating Drift with Continuous Observations," *Journal of Banking and Finance* 26, 423-444.
- Moeller, T. and C. Molina, 2003, "Survival and Default of Original Issue High-Yield Bonds," *Financial Management* 32, 83-107.
- Moody's, 2005, "Default and Recovery Rates of Corporate Bond Issuers," 1920-2004, Moody's Investor Services, Global Credit Research, New York.
- Nelson, R. and F. Siegel, 1987, "Parsimonious Modeling of Yield Curves," *Journal of Business* 60, 473-489.
- Warga, A., 1998, "Fixed Income Database," University of Houston, Houston, TX.