AN EMPIRICAL INVESTIGATION IN CREDIT SPREAD INDICES

PRIGENT J.-L.[a], O. RENAULT[b], and O. SCAILLET[c,d] [1]

[a] THEMA, Université de Cergy-Pontoise, 33 bd du Port, 95011 Cergy-Pontoise, France, prigent@u-cergy.fr

[b] Financial Markets Group, London School of Economics, Houghton Street, London WC2A 2AE, UK, o.m.renault@lse.ac.uk

[c] IRES, Département des Sciences Economiques, Université Catholique de Louvain, 3 Place Montesquieu, 1348 Louvain-la-Neuve, Belgium, scaillet@ires.ucl.ac.be

[d] Institut d'Administration et de Gestion, Université Catholique de Louvain.

September 2000

We study the dynamics of the spread between U.S. corporate and Treasury bonds. We focus on *Aaa* and *Baa* corporate yield indices and estimate nonparametrically the dynamics of the spreads assuming that they follow a univariate diffusion process. Using techniques developed for interest rate processes we try to infer from the data what acceptable process can be used to model aggregate credit spreads for option pricing or risk management purposes. We find that there is significant evidence of mean reversion especially for higher rated spreads and that the volatility of *Aaa* spreads exhibit a U-shape while the volatility of *Baa* spreads is monotonically increasing in the level of spreads. Based on these observations and on the evidence of jumps in the series, we propose a new model for credit spread indices (an Ornstein-Uhlenbeck with jumps) and estimate it by maximum likelihood.

Keywords: credit spread, risk management, jump diffusion, volatility, nonparametric.

JEL : C14, C22, E40, G20.

[1]We thank R. Anderson, P. Jakubenas and W. Perraudin for helpful comments. The third author gratefully acknowledges financial support from the Belgian program on interuniversity poles of attraction (PAI nb. P4/01).

1 Introduction

Credit risk has been a very active area of research over the past few years. Much progress has been made in improving the classic firm-value based model of Merton (1974) where default is modelled as the first time the value of the firm crosses some lower boundary (see e.g. Anderson and Sundaresan (1996) or Ericsson and Reneby (1998) for recent advances in structural models). Several new approaches including hazard rate models and Markov chain models have also been proposed to circumvent the empirical and practical difficulties encountered in structural models[2].

Hazard rate models (see for example Jarrow and Turnbull (1995) or Duffie and Singleton (1999)) do not model the value of the issuing firm explicitly but assume that default occurs by surprise, as the first jump of a Cox process. Although somehow less intuitive than firm-value based models, these models enable to calibrate the spreads more easily and are useful tools to price credit derivatives[3].

Markov chain models (see Jarrow, Lando and Turnbull (1997)) model credit risk in a credit class framework. They use transition matrices either internal to banks or published by rating agencies to infer the probability that, given that the bond is in class i at date t, it will end up in class j at date $t+1$. One of the classes is default and it is thus possible to calculate the probability of drifting down to default over a given time period.

Some versions of these three approaches are implemented by many banks under the joint pressure of regulators and shareholders. The three leading providers of credit risk packages propose models which each fit into one of the above three approaches : the KMV corporation sells a model closely inspired by Merton (1974), while the Credit Risk+ product of Credit Suisse Financial Products is a hazard rate model and JP Morgan's CreditMetrics is a rating's based model. A survey of credit risk approaches currently used by practitioners can be found in Crouhy, Galai and Mark (2000) and connections

[2]We will not review the credit risk literature in details here. We refer the interested reader to Lando (1998) and the survey by Jeanblanc and Rutkowski (1999) for an overview of the literature.

[3]Duffie and Lando (1997) and Madan and Unal (2000) propose models where the value of the firm enters the hazard rate process, thus bridging the gap between the two streams of literature.

between the three approaches are reported in Gordy (2000).

The empirical literature has lagged theoretical papers by a few years. Anderson and Sundaresan (2000) test to what extent aggregate spreads can be explained by structural credit risk models. They find that these models track historical spreads very well during some periods but fail to do so during others. Duffee (1999) estimates a one-parameter intensity model of credit risk on US corporate bonds. The author claims to be "reasonably successful at fitting corporate bond yields". Unfortunately, as noted in Duffie and Singleton (1999), the retained process (square root diffusion), does not allow to generate the observed negative correlation of spreads and riskfree rates (see Duffee (1998)) while maintaining non negative hazard rates. Little empirical research has been carried out on rating's based models but some articles including Kiesel, Perraudin and Taylor (1999) have shown that most risk stems from spreads changes within a credit risk class rather than from class changes, especially for highly rated bonds. This casts some doubts on the ability of Markov chain models to replicate individual prices accurately.

Two empirical articles are directly related to this paper. Pedrosa and Roll (1998) study a variety of corporate credit spread indices (pooled by ratings and sectors) over an 18 month period. Some of their series are similar to those we analyze in this paper. They propose to use a Gaussian mixture to model credit spread changes. Duellmann and Windfuhr (2000) focus on sovereign spreads (the difference between Italian and German bond yields) and test whether Ornstein-Uhlenbeck and square-root diffusion models are appropriate to capture the dynamics of these series. Their results indicate that the two specifications are not able to capture all the shapes of the term structure of spreads but that no single model outperforms the other although pricing errors in the Ornstein-Uhlenbeck model are more stable.

In this paper, we focus on the dynamics of corporate bond spread indices. These are average spreads calculated by Moody's for bonds in a given rating class. There are two main reasons for studying these series. First, they serve as an indicator of the level of credit spreads for many investors : they reflect the average yield spread on a well diversified corporate bond portfolio with long maturity. Investors holding such a portfolio may find it convenient to protect themselves against moves in the general level of corporate spreads in a given class rather than hedge each individual issue.

Second, they can be used as underlying for credit spread options. These instruments are now traded by many banks: a recent survey carried out by a consulting firm (KATES (1999)), showed that a quarter of U.S. large banks traded these instruments and the credit derivatives markets is growing rapidly. Another study (Risk, April 1999) estimated that the nominal amount of credit spread options would reach $100bn by the end of 2000. These options are over-the-counter contracts whose payoff depends on the terminal value and/or the path taken by the yield spread of an instrument over the yield on a riskfree bond.

In theory, these contracts can be written on any traded bond whose yield is observable on the market. In practice however, illiquid bonds and small issues should be excluded as market manipulations could distort option payoffs. This is why most trading of credit spread options takes place on a few large liquid issues.

The literature is still fairly thin on the pricing of these options and has been mainly published in practitioners' journals. An early example is FLESAKER et al. (1994) who provide an introduction to their use and basic pricing equations. LONGSTAFF and SCHWARTZ (1995) deal with options on spread when the logarithm of the spread follows an Ornstein-Uhlenbeck process. Interestingly they show that because of the mean reverting dynamics, European option prices on spreads can be below their intrinsic value : if the spread is far above its long term mean, the intrinsic value of a call on spread will be high but mean reversion will tend to push spreads down in the future. Therefore prices may be lower than the payoff the investor would obtain should the option be exercised immediately. MOUGEOT (1999) provides further results on spread options in a three-factor model (one for the riskless rate and one for each yield) where all factors are assumed to follow Ornstein-Uhlenbeck dynamics. DUFFIE and SINGLETON (1999) show how these options can be priced within a hazard rate framework of credit risk.

The vast majority of credit spread options are currently written on sovereign issues rather than corporate bond yield spreads. In a survey published in mid 1996 CIBC Wood

Gundy estimated that over 95% of the market for credit spread options consisted of options on sovereign spreads. One of the main reasons for this is that clearly identified benchmarks exist in sovereign bonds but not in the corporate debt market.

Yield spread indices would be natural candidates to fill this gap. First, they are released by a "neutral entity", the rating agency, on a daily basis : data reliability and availability is thus ensured. Then, as aggregate measures of credit risk, they are not as easy to manipulate because it would require large positions in many corporate bonds to push the index in one direction. Furthermore they have a constant maturity and therefore enable to write long term options while not worrying about pull-to-par or other maturity-related phenomenons.

An alternative solution may be preferred to the direct modelling of spreads we retain in this paper. One could construct a model of riskfree rates and a model of yields on defaultable securities and derive spreads as the difference between the two rates. However errors in both models would add up and may lead to an imprecise description of the spread process. Furthermore spreads observed in the markets are not only due to credit risk but also reflect the relative liquidity of corporate and Treasury bonds. Modelling spreads directly thus enables us to capture both components while most models of credit risk would only capture the credit component of spreads. We will see later that liquidity has indeed a very large impact on spreads.

Throughout this paper, we will assume that the dynamics for the spreads are given by the stochastic differential equation

$$dS_t^i = \mu_i \left(S_t^i \right) dt + \sigma_i \left(S_t^i \right) dW_t^i, \tag{1}$$

where $i = Aaa, Baa$ and W_t^i is a standard Brownian motion. Hazard rate models of credit risk typically assume such a process for the instantaneous spread which has close ties with the instantaneous probability of default.

We restrict ourselves to the class of time-homogenous processes. Other one-factor models used in interest rate modelling allow the parameters to be time-dependent (for example the continuous version of Black, Derman and Toy (1990)). This is typically to enable the model to calibrate exactly the term structure of interest rates and thus allow to price options on bonds consistently with observed bond prices. Our purpose is however different. One major drawback with term structure-consistent models such as those mentioned above is that they need to be constantly re-estimated and their parameters tend to vary a lot within short time intervals. Here, we are not trying to calibrate the term structure of credit spreads but indices with constant time to maturity. We will thus look for a flexible and robust specification which describes the particular series appropriately while allowing to find simple forms for pricing options.

This paper is structured as follows. In section 2 we describe the series and present some statistics and graphs, while section 3 and 4 report results obtained using respectively a nonparametric and a parametric approach to estimating μ and σ. In section 5, we propose to model the dynamics of (the logarithm of) spreads as an Ornstein-Uhlenbeck process with jumps: the solution to the stochastic differential equation is provided and the parameters of the process are estimated by maximum likelihood. Section 6 concludes.

2 The series

We focus on credit spread indices. The data were obtained from the Federal Reserve Board and consist of 3561 daily observations from January 1986 to the end of March 2000. We collected Moody's indices for seasoned corporate bond yields with a *Aaa* rating or *Baa* rating and the 10 year (constant maturity) Treasury yield constructed by the Federal Reserve. Spreads are calculated as the difference between corporate and Treasury yields.

The *Aaa* rating is the highest in Moody's rankings and corresponds to the class of bonds with lowest estimated probability of default while the *Baa* rating is at the boundary between investment grade and speculative grade issues. Historical average default frequencies over a ten-year horizon have respectively been 0.64% and 4.41% for bonds rated initially *Aaa* and *Baa* (see Keenan, Shtogrin and Sobehart (1999) (KSS))

Let us now introduce some notation. We denote the spread series as S_t^{Aaa} and S_t^{Baa} respectively while Y_t^{Aaa}, Y_t^{Baa} and Y_t^T stand for the yield on corporate bonds and on the

Treasury index at date t. We thus have $S_t^{Aaa} = Y_t^{Aaa} - Y_t^T$ and $S_t^{Baa} = Y_t^{Baa} - Y_t^T$.

Before turning to the modelling of spread dynamics, we will present some statistics and graphs of the series. Table 1 reports the basic statistics for the two spread series.

Table 1 : Summary statistics

Statistic	S_t^{Aaa}	S_t^{Baa}
Mean	1.04%	1.91%
Standard dev.	0.28%	0.38%
Minimum	0.31%	1.16%
Maximum	1.96%	3.16%
Skewness	0.363	0.751
Kurtosis	2.719	3.007
DF stat	-4.392	-3.470

The mean of the *Baa* series is about double that of the *Aaa* index. The volatility is also higher but less so in proportion. This indicates that relative changes (i.e. ln (S_t/S_{t-1})) in *Aaa* spreads are on average larger than those of *Baa* spreads. Both S_t^{Aaa} and S_t^{Baa} exhibit positive skewness. The minima for both series were achieved in early 1989 after two years of very low default statistics (KSS p. 25) while the maximum *Aaa* spread was reached at the peek of the 1998 crisis which started with Russia's default and then spread to Asia. *Baa* spreads reached their maximum in our sample in 1986. This year had the highest default rate for *Baa* bonds for the period 1970-1998 (KSS, Exhibit 29) with 1.33%.

The last row of Table 1 reports the values of the Dickey-Fuller unit root test on $S_t^i = \alpha + \rho S_{t-1}^i + u_t^i$. The hypothesis of a unit root is rejected at the 1% level for *Aaa* spreads and at the 5% level for *Baa* spreads. This is important because the econometric theory underlying the tests in the following sections require the stationarity of the series. Here, we cannot reject that these series are stationary when confronted to a unit root hypothesis.

Figure 1 plots the two series. One can check that they are indeed very correlated and that they never intersect. Although individual bonds may exhibit inconsistencies[4] between

[4]For examples of inconsistencies, see Perraudin and Taylor (1999).

spreads and ratings (for liquidity reasons or because ratings adjust slowly to changes in firm value), on average ratings tend to be good indicators of credit quality.

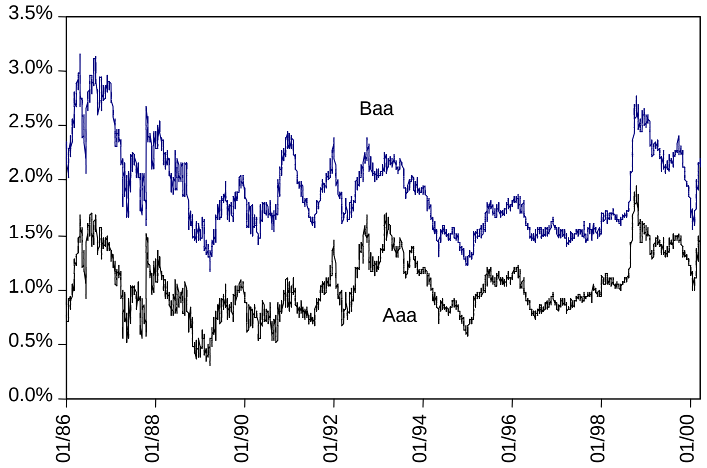

Figure 1: Baa and Aaa Spreads

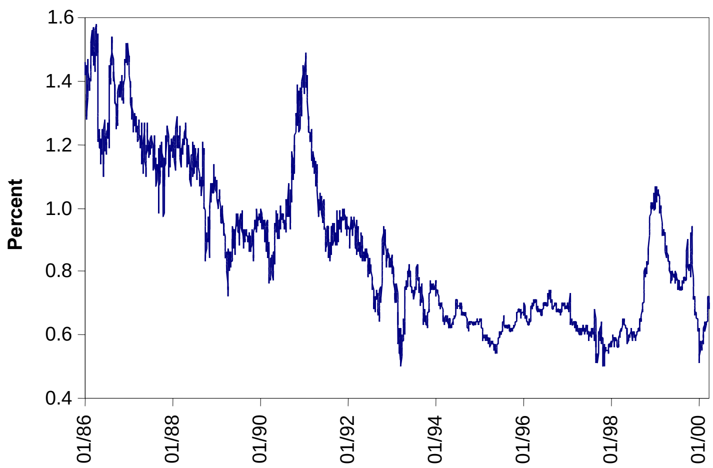

Figure 2: Baa minus Aaa spreads

The spread between *Baa* and *Aaa* yields has followed a downward trend from the beginning of our sample until about 1994 where it stabilized around 60 basis points (bp)(Figure 2). This reflected mainly the improvement in liquidity in the speculative bond market. The trend has been broken on two major occasions. The latter is due to the Russian and Asian crises which started in August 1998 while the former coincides with the run-up to the Gulf war and the war period itself.

These crises were however very different in nature. The Gulf war crisis was a true credit event which affected many US corporates. Uncertainty about oil prices and the prospects for growth created volatility in the value of the firms' assets. Risky bonds can be seen as riskless securities minus a put option on the value of the firm stricken at the nominal value of the debt (see Merton (1974)). The put options in *Aaa* bonds are very deep out-of-the-money and an increase in asset volatility thus has a smaller impact (lower Vega) on these bonds than on more risky securities. This is clearly visible on Figure 1 where *Aaa* spreads only mildly increased while *Baa* spreads soared by over 80 bp.

We believe that, in 1998, the crisis was mainly a liquidity shock rather than a real credit event. One way to validate this hypothesis is to look at spreads within the Treasury bond market. After the Russian default, the US long bond (30 year benchmark) was trading at a 35 bp premium versus the second longest bond (with just a few months less to maturity). The average spread in the past was about 7 bp (see Poole (1998)). Clearly default is not an issue here and the spread is generated by the extremely strong demand for liquidity. We compared spreads within the Treasury market in the Gulf war period and no such widening is observable, thus confirming that the 1991 crisis was a credit event.

Looking again at Figure 1, one can observe a surge in both *Aaa* and *Baa* spreads and most of the increase occurred in August after Russia's default. The economic fundamentals of US corporates were little affected as their exposure to Russia was very small and the domestic growth prospects were sound. Liquidity crises tend to affect lower rated securities more heavily (in bp, not in proportion of total spreads, see Ericsson and Renault (2000)) as being able to sell a position quickly is more valuable when the position may be subject to default. One should thus be careful when interpreting spreads as they are not only due to default probability but also to changes in liquidity and risk aversion.

Turning to Figure 3 we see that heteroskedasticity is present in our series (a similar picture, not reported here, is obtained for *Aaa* spreads). Relative changes in spreads were markedly more volatile at the beginning of our observation period and although recent crises have shaken the corporate bond market, the volatility has never reached the levels attained during the 1980s. This may reflect the greater maturity and liquidity of the

US corporate bond market. Finally, note the spillover effect of stock market volatility[5] on corporate bond spreads since the beginning of 2000. We will see another spectacular example of such a spillover later.

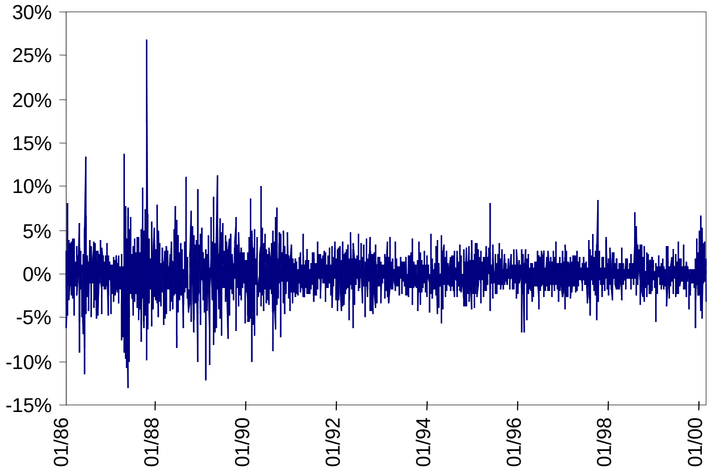

Figure 3: Daily Relative Changes in Baa Spreads

3 Nonparametric estimation

Before imposing some structure on the spread dynamics it is useful to apply nonparametric techniques to extract some information about the possible specifications for the drift and diffusion. Stanton (1997) proposes first-, second- and third-order approximations to μ and σ when the interest rate process follows a SDE as equation (1).

3.1 Density Estimation

The first step is to estimate the densities of the spread processes. We use a Gaussian kernel estimator of the spread density for the bond class i :

$$\widehat{f}_i\left(x\right)=\frac{1}{nh_i}\sum_{t=1}^{n}\varphi\left(\frac{x-S_t^i}{h_i}\right), \tag{2}$$

where $\varphi\left(.\right)$ is the standard normal density function, n is the number of observations and the window width is given by :

$$h_i=c\widehat{\sigma}_i n^{-\frac{1}{5}},$$

[5]In the first quarter of 2000, the S&P 500 volatility was 24.5%, compared to a five year average of 16%.

where c is a constant, and $\widehat{\sigma}_i$ is the empirical standard deviation of spreads in class i.

The choice of c depends mainly on the level of smoothness one is willing to achieve for the density. In this paper we have chosen $c = 3$ which is close to the value used in Stanton (1997). A discussion of the choice of kernel and bandwidth can be found in Scott (1992) and Pagan and Ullah (1999). The estimated densities are shown on Figure 4 and 5. There is evidence of positive skewness especially in the *Baa* series.

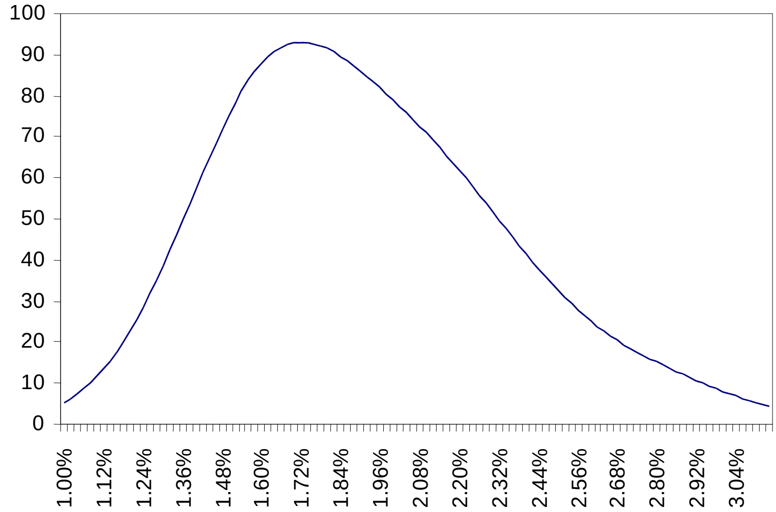

Figure 4: Nonparametric Density of Baa Spreads

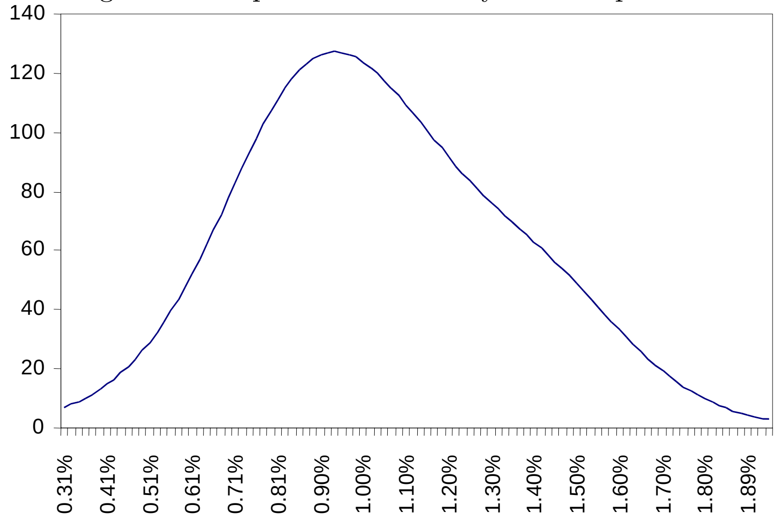

Figure 5: Nonparametric density of Aaa spread

3.2 Drift Estimation

The drift term at level x can be estimated using[6]:

$$\widehat{\mu}_1^i(x) = 250 \times \frac{\sum_{t=1}^{n-1}\left(S_{t+1}^i - S_t^i\right)\varphi\left(\frac{x-S_t^i}{h_i}\right)}{\sum_{t=1}^{n-1}\varphi\left(\frac{x-S_t^i}{h_i}\right)}. \tag{3}$$

Similarly an estimator of the diffusion is given by:

$$\widehat{\sigma}_1^i(x) = \left(250 \times \frac{\sum_{t=1}^{n-1}\left(S_{t+1}^i - S_t^i - \widehat{\mu}_1^i(x)\right)^2\varphi\left(\frac{x-S_t^i}{h_i}\right)}{\sum_{t=1}^{n-1}\varphi\left(\frac{x-S_t^i}{h_i}\right)}\right)^{\frac{1}{2}}. \tag{4}$$

These are estimators corresponding to a first order approximation. Higher order estimators can also be derived (see Stanton (1997), estimators corresponding to second-order approximation given in appendix A).

Figure 6 and 7 plot the estimated drift term as a function of the level of the spreads. We only report results for first- and second-order approximations (The line *Par* will be discussed in the next section). Third-order approximations were also analyzed and lead to very similar results[7].

The drift is clearly not constant in the level of spreads but tends to decrease with spread levels. This is evidence of mean reversion : we can indeed check that the value of the drift is close to zero for values of the spread around their mean. The drift is positive for values of the spread below the mean and negative for values above the mean. Notice also that although the drift is not strictly linear, it does not show the kind of negative exponential shape found in Stanton (1997) for the short rate drift. This latter result may not be robust as shown in Chapman and Pearson (2000) : the nonparametric estimators used in Ait Sahalia (1996b), Stanton (1997) and in this paper may suggest nonlinearities for the drift at high values of the spread although the true drift is linear. The curved shape of the *Baa* drift for S_t^{Baa} above 2.5% should therefore be interpreted with care since distortions may appear at boundaries.

[6] The second-order approximation formulae are given in Appendix A.

[7] These results are available on request.

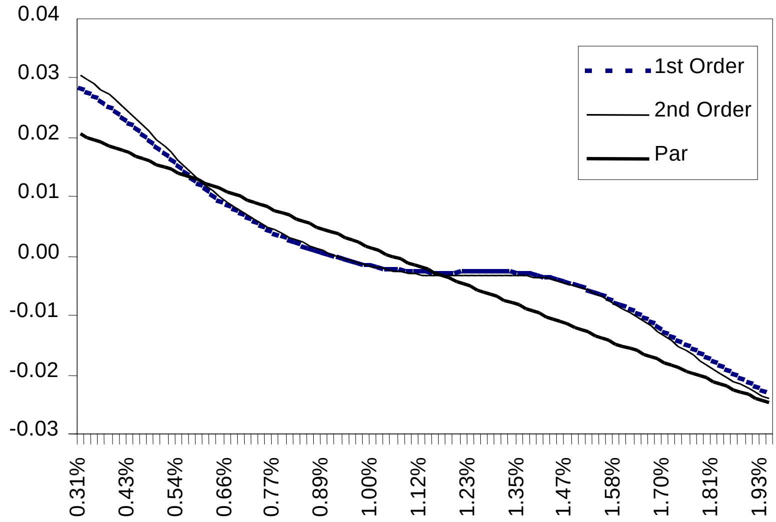

Figure 6: Aaa spread drift as a function of spread level

Comparing the two figures, one can notice that the curve is steeper for *Aaa* spreads than for *Baa* spreads, indicating faster mean reversion in *Aaa* bonds yields than in lower rated bonds. Longstaff and Schwartz (1995) also report a similar finding over a different observation period. Some practitioners have also acknowledged this fact and recommended asset allocation strategies based on it : for example when expecting the end of a crisis where spreads are far above their long-term mean, it may be appropriate to invest first in *Aaa* bonds (which recover faster) and then progressively move to more speculative securities (see Goldman (1998)).

Recall that spreads are not only due to credit risk but also to liquidity premia. Changes in credit risk fluctuate with real economic variables such as the business cycle and are therefore rather longlasting. On the other hand changes in liquidity premia are very volatile and depend a lot on market sentiment. Liquidity crises are usually a matter of months. It is therefore reasonable to expect that mean reversion in the liquidity component of spreads should be much higher than that of the credit component. *Aaa* spreads are explained in a greater proportion by liquidity and should intuitively revert more quickly to their long term average than *Baa* spreads.

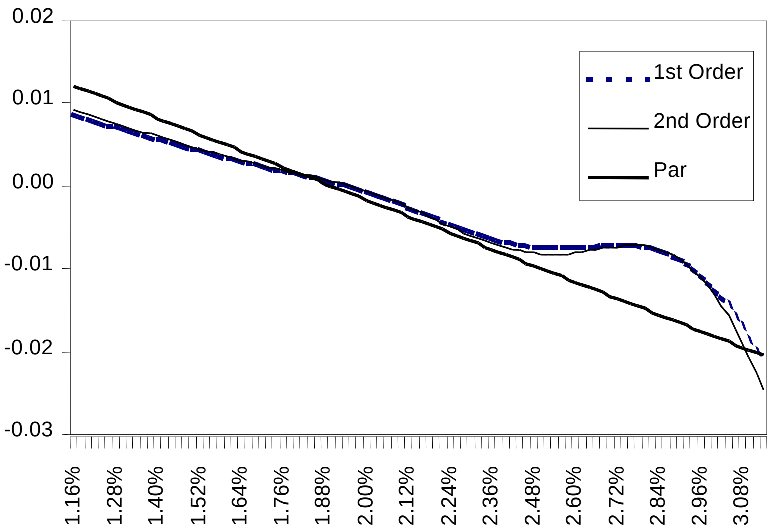

Figure 7: Baa spread drift as a function of spread level

3.3 Volatility Estimation

We now consider the diffusion component. Looking at Figures 8 and 9, which plot the daily changes $S_t^i - S_{t-1}^i$ versus the level of spreads S_{t-1}^i, there is no obvious sign of a link between the level of spreads and the volatility. This contrasts with short term interest rates for example where the volatility is clearly positively correlated with the levels (see e.g. Stanton (1997)) and with order greater than 1, i.e. $\sigma(x) = f(x^\gamma)$, $\gamma > 1$. If this observation is confirmed statistically in the parametric section, one may be able to choose a simple form for the volatility term.

Choosing a simple model may be useful in the pricing of path dependent credit spread options where not only the terminal distribution is required but also other laws such as the first passage time of the process to a boundary. Closed-form expressions for first passage time densities are rare : the law for the Brownian motion is well known (see e.g. Harrison (1985)), that for the Ornstein-Uhlenbeck process is given in Leblanc, Renault and Scaillet (2000). Leblanc and Scaillet (1998) and Goeing and Yor (1999) provide an explicit expression of the Laplace transform of first hitting time by a square root (CIR) process.

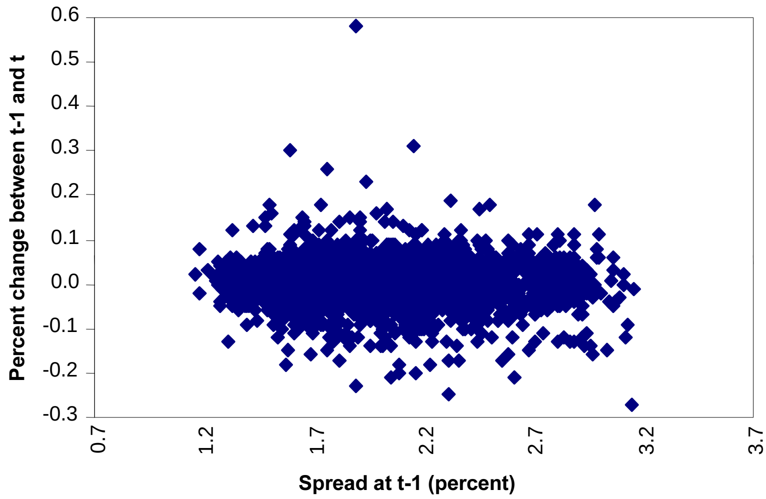

Figure 8: Daily changes in Baa spreads

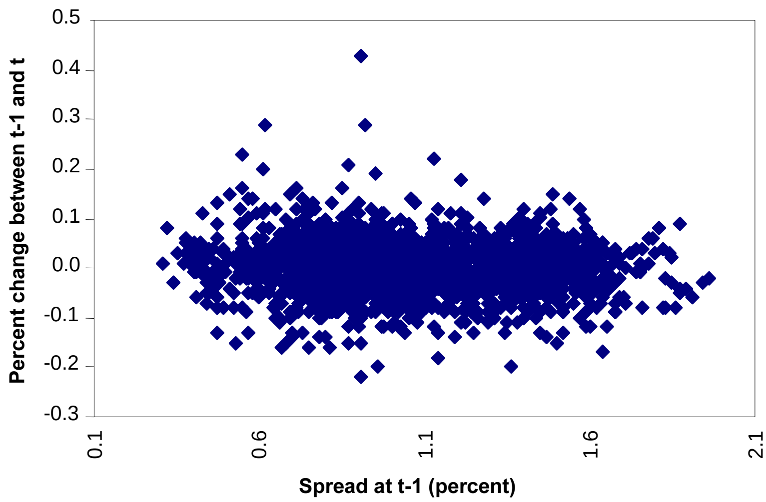

Figure 9: Daily changes in Aaa spreads

Notice the extreme point at the top of each graph. This corresponds to 19-20th October 1987 and is yet another very dramatic sign of a connection between the equity and the corporate bond markets. On this day, *Baa* spreads increased by nearly 60 basis points while *Aaa* spreads increased by over 40 basis points. On the previous trading day the increases were already 30 and 20 basis points respectively.

Our kernel estimator shows two very distinct shapes for the diffusion terms (Figures 10 and 11). The *Baa* spread seems to have a monotonically increasing volatility as a function of the level of spreads while the *Aaa* spread exhibits a U-shaped volatility curve. One may again try to explain these findings in the light of the different proportions of liquidity and credit components in the spreads. As we argued earlier, *Aaa* spread changes are explained to a large extent by changes in liquidity and risk aversion. During quiet periods, the *Aaa* spread will oscillate around its mean of about 1.04%.

Two types of abnormal periods can then occur. First, a flight to quality (such as the second semester of 1998) can quickly dry out the liquidity in the corporate bond market thus pushing spreads very high and increasing their volatility dramatically. This explains the right part of Figure 10.

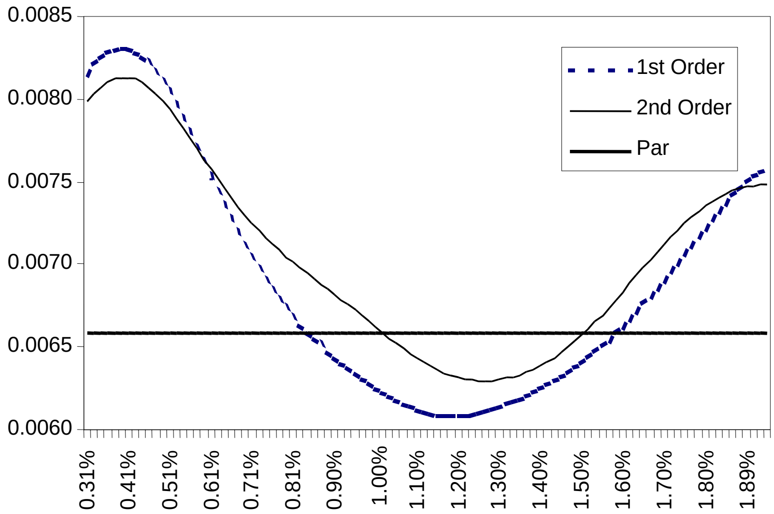

Figure 10: Aaa spread volatility as a function of spread level

Second, a long bull market (such as those of 1988 and 1993-94) may trigger massive liquidity inflows in credit risky markets (both corporate and sovereign). This pushes spreads to levels as low as 35 basis points and implies an extremely low risk premium on corporate bonds. Then an external event such as a crisis in an emerging market or the collapse of a large corporate issuer would immediately lead investors to reassess their credit risk exposure and to cut their share of funds invested in credit risky instruments. This would bring back risk premia to more “normal” levels. Thus very low levels of spreads are also associated with high instability and the left part of Figure 10 thus also makes sense. Again the line *Par* will be discussed in the next section.

The estimation for the *Baa* spread diffusion is slightly surprising when looking at Figure 8. The explanation is however more straightforward than for the *Aaa* spread. An increasing curve would be what one should expect if the spreads were pure credit spreads. Low credit spreads are associated with low risk and low default frequency and thus to quiet periods in the corporate bond market. Conversely, when the economy is in a downturn and defaults become more and more frequent, spreads are high and tend to react dramatically to every new default in the market. Thus high credit spread levels are naturally associated with high volatility. Given that *Baa* spreads contain a substantial part of credit risk component, it is not surprising to find an increasing curve for the diffusion term.

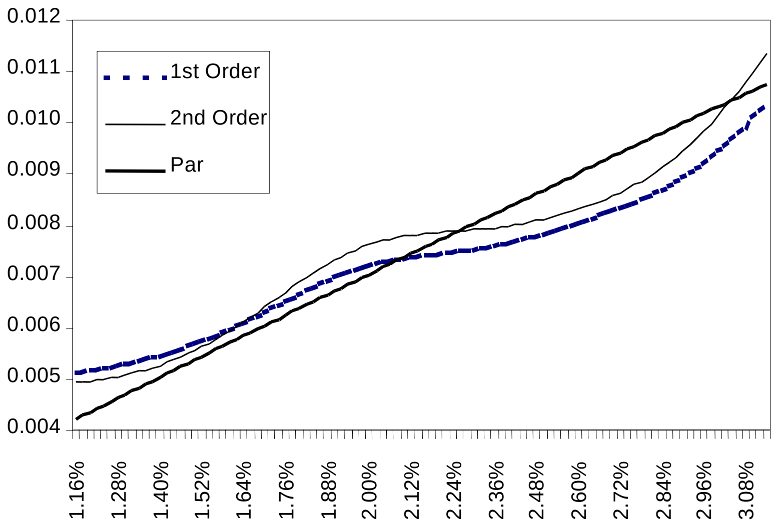

Figure 11: Baa spread volatility as a function of spread level

We have now obtained nonparametric estimates for the drift and diffusion components of *Aaa* and *Baa* spreads. We saw that both drifts are nearly linear (decreasing) in the level of spreads and that a linear specification is also reasonable for the *Baa* spread diffusion. The *Aaa* spread diffusion is however not at all linear but is higher for values of the spread far away from its mean and lower for spread values close to the mean.

4 Parametric estimation

We now move on to parametric estimation of corporate credit spread indices. Although nonparametric pricing is possible (see for example Ait-Sahalia (1996a)), it is typically easier to use parametric specifications to obtain closed-form solutions for options or to carry out simulations.

We follow the approach of Chan et al. (1992) who propose to estimate the discrete time model :

$$S_{t+1}^i - S_t^i = \alpha_D + \beta_D S_t^i + \sigma_D \left| S_t^i \right|^{\gamma_D} + \epsilon_{t+1}, \tag{5}$$

where ϵ_{t+1} are assumed to be i.i.d. normal variables. If $\alpha_D > 0$ and $\beta_D < 0$, we obtain the usual mean-reverting term with $-\beta_D$ as speed of mean reversion and $-\alpha_D/\beta_D$ as long term mean. γ_D measures the degree of nonlinearity of the relationship between the spread level and its volatility.

Most one-factor models used for interest rate modelling are nested in this model. Merton (1973) is obtained by setting $\beta_D = 0$ and $\gamma_D = 0$, while $\gamma_D = 0$ yields the Vasicek (1977) model. Dothan (1978) corresponds to $\alpha_D = \beta_D = 0$ and $\gamma_D = 1$, while Brennan and Schwartz (1980) and Courtadon (1982) have the restriction $\gamma_D = 1$ and Cox, Ingersoll and Ross (1985) set $\gamma_D = \frac{1}{2}$.

The Markovian property of the process and the assumption of normality enable us to easily find the transition density and the log-likelihood function:

$$L = -n \ln \left(\sqrt{2\pi} \sigma_D \right) - \sum_{j=1}^{n} \ln \left(\left| S_{j-1}^i \right|^{\gamma_D} \right) - \frac{1}{2} \sum_{j=1}^{n} \left(\frac{S_j^i - \alpha_D + (\beta_D + 1) S_{j-1}^i}{\sigma_D \left| S_{j-1}^i \right|^{\gamma_D}} \right)^2$$

Since maximum likelihood yields asymptotically unbiased estimates with minimum variance, it is preferable to apply maximum likelihood rather than a method of moments whenever possible (see Broze, Scaillet and Zakoian (1995) for a discussion).

Maximum likelihood estimates for $S_{t+1}^i - S_t^i = \alpha_D + \beta_D S_t^i + \sigma_D \left| S_t^i \right|^{\gamma_D} \epsilon_{t+1}$ for $i = Aaa$ and $i = Baa$ are reported in Table 2[8].

[8]Note that we estimated the model on spreads $\times 100$, so that a spread of 1% would enter the estimation as 1 and not 0.01.

Table 2 : Maximum Likelihood Estimates

	S_t^{Aaa}	S_t^{Baa}
$\widehat{\alpha}_D$	0.0116*	0.0133*
$\widehat{\beta}_D$	−0.0110*	−0.0071*
$\widehat{\sigma}_D$	0.0416*	0.0232*
$\widehat{\gamma}_D$	0.0000	0.9395*

*significant at 1% level.

Recall that we are using daily data. So, the estimators for the parameters of the continuous process are respectively given by $\widehat{\alpha} = 250 \times \widehat{\alpha}_D, \widehat{\beta} = 250 \times \widehat{\beta}_D, \widehat{\sigma} = \widehat{\sigma}_D \sqrt{250}$ and $\widehat{\gamma} = \widehat{\gamma}_D$.

4.1 Aaa spread dynamics

For the *Aaa* series, one can notice that the maximum likelihood fit is obtained with a constant variance. This corresponds to drawing a straight horizontal line through Figure 10 at the level $\widehat{\sigma}_D \sqrt{250} = 0.658\%$. Line *Par* in Figure 10 is set at the level $\widehat{\sigma}_D \sqrt{250}$.

There is evidence of mean reversion ($\beta < 0$) as could have been expected from the nonparametric estimation (Line *Par* in Figure 6 is the parametric estimate for the drift term). The estimated dynamics for the *Aaa* spread (mean reverting drift and constant variance) are those of an Ornstein-Uhlenbeck (O-U) process $dS_t^{Aaa} = \left(\alpha + \beta S_t^{Aaa}\right) dt + \sigma dW_t$. This specification was for example assumed by Duellmann and Windfuhr (2000) to model sovereign spreads between Italy and Germany.

It is well known (see BSZ or De Winne (1998) for example) that the Euler scheme (equation 5) introduces a discretization bias and that one should rather use the exact discretization for the Ornstein-Uhlenbeck process :

$$S_{(t+1)\Delta}^{Aaa} - S_{t\Delta}^{Aaa} = \left(-\alpha/\beta - S_{t\Delta}^{Aaa}\right)\left(1 - e^{\beta\Delta}\right) + \sigma e^{\beta(t+1)\Delta} \int_{t\Delta}^{(t+1)\Delta} e^{-\beta s} dW_s, \qquad (6)$$

where Δ is the length of time between two observations, here $\Delta = 1/250$ year.

This exact discretization can be estimated using a single regression. We performed this estimation and obtained the results reported in Table 3. They are compared to the estimates obtained with the Euler discretization (annualized parameters from Table 2).

Table 3 : Estimation of O-U process

Estimate	Exact	Euler
$\widehat{\alpha}$	2.916*	2.902*
$\widehat{\beta}$	−2.768*	−2.754*
$\widehat{\sigma}$	0.662*	0.658*

*significant at 1% level.

The bias is very small both in the drift and the diffusion term, which is consistent with the findings of Schaefer (1980).

The choice of a Gaussian process for the credit spread is admittedly debatable because of the possibility of negative values. However this probability is very low : over a five-year horizon and using the formula given by Leblanc, Renault and Scaillet (2000) we calculated that the probability is only just above 1%, increasing to 2.5% over a ten-year horizon.

4.2 Baa spread dynamics

Results for *Baa* spreads (Table 2) also show evidence of mean reversion. We find than mean reversion is slower for *Baa* spreads than for *Aaa* spreads $\left(\left|\beta^{Aaa}\right| > \left|\beta^{Baa}\right|\right)$ as expected from the nonparametric section (see line *Par* in Figure 7). The diffusion is however not constant. We find that $\widehat{\gamma}_D$ is close to 1 which implies that the volatility of the spread increases almost linearly with its level. Line *Par* in Figure 11 is a plot of the parametric estimate of the diffusion term.

Setting $\gamma = 1$ and estimating the model again, we get $\widehat{\alpha}_D = 0.0134$, $\widehat{\beta}_D = -0.0070$ (the drift term is unaffected) and $\widehat{\sigma}_D = 0.0224$. This model is the discrete version of Brennan and Schwartz (1980), also used in Courtadon (1982). The diffusion term thus obtained is indistinguishable from line *Par* which is reasonably satisfactory at capturing the general shape of volatility.

5 A new process for credit spreads

Following the empirical observations made in the two previous sections, we now propose a new model for credit spread indices, guarantying positive spreads while capturing jumps and mean reversion. We start with its description before discussing its properties and carrying out its estimation on the data.

5.1 The process

Let Y_t denote the logarithm of the credit spread index (either Aaa or Baa). We assume that Y_t follows the dynamics:

$$dY_t = \alpha(\theta - Y_t)\,dt + \sigma dW_t + dN_t, \tag{7}$$

where W_t is a standard Brownian motion and N_t is a compound Poisson process. T_n denotes the n^{th} jump time of the compound Poisson process and X_n its associated mark (jump size). Throughout, we will assume that the distribution of marks is time independent and uncorrelated with the Brownian part of the process.

Let us denote

$$m_1 = E[X_n] = \int_{\Re} xK(dx),$$

and

$$m_2 = E\left[X_n^2\right] = \int_{\Re} x^2 K(dx),$$

where $K(dx)$ is the conditional probability (kernel) that the mark falls in the interval $dx \in \Re$. N_t is such that $(N_t - \lambda t m_1)_t$ is a martingale.

In the sequel, we will assume that jump times follow an exponential distribution with parameter λ and that marks follow a binomial distribution with probability $\frac{1}{2}$ and jump size a. We thus have $dx \in [-a, +a]$, $K(+a) = K(-a) = \frac{1}{2}$, $m_1 = 0, m_2 = a^2$.

Defining the martingale $M_t = \sigma W_t + (N_t - \lambda t m_1)_t$, the solution to the stochastic differential equation (7) takes an explicit form:

$$Y_t = Y_0 \exp(-\alpha t) + \left(\theta + \frac{\lambda m_1}{\alpha}\right)(1 - \exp(-\alpha t)) + \int_0^t \exp(-\alpha(t-s))\,dM_s, \tag{8}$$

and the conditional expectation and variance are respectively:

$$\begin{aligned}
E_0\left[Y_t\right] &= Y_0 \exp\left(-\alpha t\right) + \left(\theta + \frac{\lambda m_1}{\alpha}\right)\left(1 - \exp\left(-\alpha t\right)\right), \\
Var_0\left[Y_t\right] &= \left(1 - \exp\left(-2\alpha t\right)\right)\frac{\left(\lambda m_2 + \sigma^2\right)}{2\alpha}.
\end{aligned}$$

The proofs are detailed in Appendix B.

5.2 Properties of the process

We can now discuss the properties of our process. α (we assume $\alpha > 0$) is the traditional speed of mean reversion as in the Ornstein-Uhlenbeck specification. Our process is therefore able to capture this feature of the data.

θ is however *not* the long term mean of the process in the general case. We can indeed calculate the expectation of the stationary distribution by taking the limit of the conditional expectation as time increases to infinity :

$$m = \lim_{t \to \infty} E_0\left[Y_t\right] = \left(\theta + \frac{\lambda m_1}{\alpha}\right).$$

In the specific case we retain in this paper, $m_1 = 0$ and we therefore fall back on the intuitive case where θ takes the interpretation of the long run mean. The formula for half life (the expected time the process will take to reach the middle level between its current value and its long term mean) is also unaffected when $m_1 = 0$.

Similarly, taking the limit, we find that the stationary variance is:

$$V = \lim_{t \to \infty} Var_0\left[Y_t\right] = \frac{\left(\lambda m_2 + \sigma^2\right)}{2\alpha}.$$

Notice that the stationary variance is bounded. The stationary distribution has therefore a mean m and variance V[9].

This process enables us to capture jumps which are obviously present in the data (see figure 1). Similar specifications have been proposed by Das (1996) for modelling interest rates. However, using logarithms precludes negative values for spread indices without the need for truncation used by Das (1996).

[9]The characteristic function of the distribution and its derivation are available upon request.

5.3 Estimation

The process described above is conditionally normal, i.e. given that there is an up-jump, a down-jump, or no jump, the distribution is normal with a corresponding mean. It is thus straightforward to decompose the likelihood function in a product of normal distributions weighted by the probability of having a given jump or no jump at all.

We want to estimate the parameters of the process using discrete data spaced by the short time interval Δ, chosen to be 1/250 year (one day) in our estimations. The dataset has been described in section 2. Let x_i denote changes in log spreads over the period $(i-1)\Delta$ to $i\Delta$. We have:

$$\begin{aligned}
\mu_i &= E_{(i-1)\Delta}[x_i] = \left(\theta + \frac{\lambda m_1}{\alpha} - Y_{(i-1)\Delta}\right)(1 - \exp(-\alpha\Delta)), \\
\sigma_i^2 &= Var_{(i-1)\Delta}[x_i] = (1 - \exp(-2\alpha\Delta))\frac{(\lambda m_2 + \sigma^2)}{2\alpha},
\end{aligned}$$

and the log-likelihood function is, save on a constant:

$$L(\underline{x}|\,\underline{\Gamma}) = \sum_{i=1}^{n} \ln\left\{ e^{-\lambda\Delta}\phi\left(x_i; \mu_i, \sigma_i^2\right) + \sum_{j=1}^{\infty}\frac{1}{2}e^{-\lambda\Delta}\frac{(\lambda\Delta)^j}{j!}\left[\phi\left(x_i; \mu_i - ja, \sigma_i^2\right) + \phi\left(x_i; \mu_i + ja, \sigma_i^2\right)\right]\right\},$$

where $\phi(h; k, \Sigma^2)$ is the normal density at point h with mean k and variance $\Sigma^2, \underline{\Gamma} = (\alpha, \theta, \sigma, \lambda, a)$, $\underline{x}$ is the vector of n log spread changes. For practical estimations, we have truncated the infinite sum at $j = 15$.[10] The maximization program used a Newton-Raphson technique and we have checked that the algorithm converges to the same set of optimal parameters when starting from different initial values.

Results for our estimations are reported in table 4. They all are significant at the 1% level. Our results are in line with some observations made in section 2, 3 and 4. First, both series exhibit mean reversion and the *Aaa* spreads revert much faster to their long term average. Second, the lower volatility for the *Baa* series is consistent with our previous observation that relative changes (remember we are working with logs) in *Aaa* spreads are on average larger than those of *Baa* spreads.

[10]This is a very conservative choice. Truncating at $j = 5$ would have yielded a percentage difference in parameters of less than 10^{-4}.

Jumps are found to be significant in both series and a likelihood ratio test of the jump Ornstein-Uhlenbeck versus its diffusion counterpart strongly rejects the assumption of *no jumps* at the 5% level. Note that the size of (relative) jumps in *Baa* spreads is about half that of jumps in *Aaa* spreads. This indicates that average jumps in both series are approximately the same size because the level of *Aaa* spreads oscillated around 1% and *Baa* spreads around 2%.

Table 4 : Parameter estimates

Estimates	Aaa	Baa
$\widehat{\alpha}$	2.828*	1.048*
$\widehat{\theta}$	−4.489*	−3.975*
$\widehat{\sigma}$	0.397*	0.224*
$\widehat{\lambda}$	44.879*	34.337*
$\widehat{a}$	0.0801*	0.0405*

* significant at the 1% level.

We have also considered lifting the assumption of symmetric jumps $[-a, +a]$. We estimated the model without the restriction and found an up-jump size of 0.0806 and down-jump size -0.0751 for the *Aaa* series while corresponding figures for *Baa* indices are 0.0418 and -0.0401. In both cases, likelihood ratio tests at the 5% level cannot reject the hypothesis of equal jump sizes.

6 Conclusion

In this paper, we have used nonparametric and parametric techniques to model *Aaa* and *Baa* corporate bond spread indices. Both series are found to be mean reverting with speed of mean reversion much higher for higher rated bonds. Nonparametric estimates of the drift show an almost linear negative relationship between the level of spreads and the drift terms. On the contrary, the diffusion function is increasing in the level of spreads in the *Baa* series, while it exhibits a U-shape for *Aaa* spreads.

Parametric estimates show that the "best" specification for *Aaa* spreads within a broad class of processes is as an Ornstein-Uhlenbeck process (mean reverting with constant volatility). This is partly due to the inability of the class of process to capture a U-shaped volatility. *Baa* spreads are best described as a Brennan and Schwartz (1980) process.

Based on these empirical findings, we propose a new simple and flexible model for credit spread indices which incorporates both mean reversion and jumps. Parameters of the process are estimated and tests on the restrictions imposed by our model are found to support the assumption of symmetric jumps.

REFERENCES :

- Ait-Sahalia Y. (1996a) "Nonparametric pricing of interest rate derivative securities", *Econometrica*, **64**, 527-560.
- Ait-Sahalia Y. (1996b) "Testing continuous-time models of the spot interest rate", *Review of Financial Studies*, **9**, 385-426.
- Anderson R. and S. Sundaresan (1996) "Design and valuation of debt contracts", *Review of Financial Studies*, **9**, 37-68.
- Anderson R. and S. Sundaresan (2000) "A comparative study of structural models of corporate bond yields: an exploratory investigation", *Journal of Banking and Finance*, **24**, 255-269.
- Berndt E., B. Hall, R. Hall and J. Hausman (1974) "Estimation and inference in nonlinear structural models", *Annals of Economic and Social Measurement*, **3**, 653-665.
- Black F. and J. Cox (1976) "Valuing corporate securities : some effects of bond indenture provisions", *Journal of Finance*, **31**, 351-367.
- Black F., E. Derman and W. Toy (1990) "A one-factor model of interest rates and it's application to Treasury bond options", *Financial Analysts Journal*, Jan-Feb, 33-39.
- Brennan M. and E. Schwartz (1980) "Analysing convertible bonds", *Journal of Financial and Quantitative Analysis*, **15**, 907-929.
- Broyden C. (1970) "The convergence of a class of double-rank minimization algorithms 2, The New Algorithm.", *Journal of the Institute for Mathematics and Applications*, **6**, 222-231.
- Broze L., O. Scaillet and J-M. Zakoian (1995) "Testing for continuous-time models of the short-term interest rate", *Journal of Empirical Finance*, **2**, 199-223.

- Chan K.C., G. Karolyi, F. Longstaff and B. Sanders (1992) "An empirical comparison of alternative models of the short-term interest rate", *Journal of Finance*, **47**, 1209-1227.

- Chapman D. and N. Pearson (2000) "Is the short rate drift actually nonlinear", *Journal of Finance*, **55**, 355-388.

- Courtadon G. (1982) "The pricing of options on default-free bonds", *Journal of Financial and Quantitative Analysis*, **17**, 75-100.

- Cox J., J. Ingersoll and S. Ross (1985) "A theory of the term structure of interest rates", *Econometrica*, **53**, 385-407.

- Crouhy M., D. Galai and R. Mark (2000) "A comparative analysis of current credit risk models", *Journal of Banking and Finance*, **24**, 59-117.

- Das S.R. (1996) "Discrete-time bond and option pricing for jump-diffusion processes", *Review of Derivatives Research*, **1**, 211-243.

- De Winne R. (1998), "Processus de diffusion et biais de discrétisation : une analyse empirique par inférence indirecte appliquée aux taux d'intérêt", *Finance*, **19**, 41-58.

- Dothan U. (1978), "On the term structure of interest rates", *Journal of Financial Economics*, **6**, 59-69.

- Duellmann K. and M. Windfuhr (2000) "Credit spreads between German and Italian sovereign bonds - do affine models work ?", Working Paper University of Mannheim.

- Duffee G. (1998) "The relation between Treasury yields and corporate bond yield spreads", *Journal of Finance*, **53**, 2225-2242.

- Duffee G. (1999) "Estimating the price of default risk", *Review of Financial Studies*, **12**, 197-226.

• Duffie D. and D. Lando (1997) "Term structure of credit spreads with incomplete accounting information", Working Paper, Stanford University, forthcoming in *Econometrica.*

• Duffie D. and K. Singleton (1999) "Modeling term structures of defaultable bonds", *Review of Financial Studies*, **12**, 687-720.

• Ericsson J. and O. Renault (2000) "Liquidity and credit risk", mimeo, Université Catholique de Louvain.

• Ericsson J. and J. Reneby (1998) "A framework for valuing corporate securities", *Applied Mathematical Finance*, **5**, 143-163.

• Flesaker B., L. Hughson, L. Schreiber and L. Sprung (1994) "Taking all the credit", *Risk*, **7**, 104-108.

• Fletcher R. (1970) "A new approach to variable-metric algorithms", *Computer Journal*, **13**, 317-322.

• Goeing A. and M. Yor (1999) "A survey and some generalizations of Bessel processes", Risklab Working Paper, ETH Zurich.

• Goldfarb D. (1970) "A family of variable-metric algorithms derived by variational means", *Mathematics of Computation*, **24**, 23-26.

• Goldman D. (1998), "Corporate bonds : some lessons from past crises and recoveries", *Global Strategy & Asset Allocation*, Credit Suisse First Boston, **October**, 13-16.

• Gordy M. (2000) "A comparative anatomy of credit risk models", *Journal of Banking and Finance*, **24**, 119-149.

• Harrison J. (1985), "Brownian motion and stochastic flow systems", John Wiley and Sons, New York.

- Ho T. and S. Lee (1986) "Term structure movements and pricing interest rate contingent claims", *Journal of Finance*, **41**, 1011-1029.

- Jarrow R., D. Lando and S. Turnbull (1997) "A Markov model for the term structure of credit risk spreads", *Review of Financial Studies*, **10**, 481-523.

- Jarrow R. and S. Turnbull (1995) "Pricing derivatives on financial securities subject to credit risk", *Journal of Finance*, **50**, 53-85.

- Jeanblanc M. and M. Rutkowski (1999) "Modelling of default risk : an overview", Discussion Paper, Université d'Evry.

- Kates G. (1999) "Credit derivatives risks and trends", *Risk Professional*, **1**, 28-31.

- Keenan S., I. Shtogrin and J. Sobehart (1999) "Historical default rates of corporate bond issuers, 1920-1998", Special Comment, Moody's Investors Service.

- Kiesel R., W. Perraudin and A. Taylor (1999) "The structure of credit risk", Working Paper, Birkbeck College.

- Lando D. (1998) "On Cox processes and credit risky securities", *Review of Derivatives Research*, **2**, 99-120.

- Leblanc B., O. Renault and O. Scaillet (2000) "A correction note on the first passage time of an Ornstein-Uhlenbeck process to a boundary", *Finance and Stochastics*, **4**, 109-111.

- Leblanc B. and O. Scaillet (1998) "Path dependent options on yields in the afine term structure model", *Finance and Stochastics*, **2**, 349-367.

- Longstaff F. and E. Schwartz (1995) "Valuing credit derivatives", *Journal of Fixed Income*, **5**, 6-12.

- Madan D. and H. Unal (2000) "A two-factor hazard rate model for pricing risky debt and the term structure of credit spreads", *Journal of Financial and Quantitative Analysis*, **35**, 43-65.

- Merton R. (1973) “Theory of rational option pricing”, *Bell Journal of Economics and Management Science*, **4**, 141-183.
- Merton R. (1974) “On the pricing of corporate debt : the risk structure of interest rates”, *Journal of Finance*, **29**, 449-470.
- Mougeot N. (1999), “Pricing spread options”, Discussion Paper, FAME, HEC Lausanne.
- Pagan A. and A. Ullah (1999) *Nonparametric econometrics*, Cambridge University Press, Cambridge.
- Pedrosa M. and R. Roll (1998) “Systematic risk in corporate bond credit spreads”, *Journal of Fixed Income*, **8**, 7-26.
- Perraudin W. and A. Taylor (1999) “On the consistency of ratings and bond market yields”, Working Paper, Birkbeck College.
- Poole W. (1998) “Whither the U.S. Credit Markets ?”, Presidential Speech, Federal Reserve of St Louis.
- Schaefer S. (1980) Discussion of the article “Conditional predictions of bond prices and returns” by Brennan M. and E Schwartz, *Journal of Finance*, **35**, 417-419.
- Scott D. (1992) *Multivariate density estimation : theory, practice and visualization*, John Wiley, New York.
- Shanno D. (1970) “Conditioning of quasi-Newton methods for function minimization”, *Mathematics of Computation*, **24**, 647-656.
- Stanton R. (1997) “A nonparametric model of term structure dynamics and the market price of interest rate risk”, *Journal of Finance*, **52**, 1973-2002.
- Vasicek O. (1977) “An equilibrium characterization of the term structure”, *Journal of Financial Economics*, **5**, 177-188.

APPENDIX A

Second-order approximation of the drift and variance terms :

$$\begin{aligned}
\widehat{\mu}_2^i(x) &= \frac{250}{2} \times \left\{4E\left[S_{t+1}^i - S_t^i \middle| S_t^i = x\right] - E\left[S_{t+2}^i - S_t^i \middle| S_t^i = x\right]\right\} \qquad \text{(A1)} \\
&\approx 125 \times \left\{4\frac{\sum_{t=1}^{n-1}\left(S_{t+1}^i - S_t^i\right)\varphi\left(\frac{x-S_t^i}{h}\right)}{\sum_{t=1}^{n-1}\varphi\left(\frac{x-S_t^i}{h}\right)} - \frac{\sum_{t=1}^{n-2}\left(S_{t+2}^i - S_t^i\right)\varphi\left(\frac{x-S_t^i}{h}\right)}{\sum_{t=1}^{n-2}\varphi\left(\frac{x-S_t^i}{h}\right)}\right\},
\end{aligned}$$

and

$$\begin{aligned}
\widehat{\sigma}_2^i(x) &= \sqrt{\frac{250}{2} \times \left\{4var\left(S_{t+1}^i \middle| S_t^i = x\right) - var\left(S_{t+2}^i \middle| S_t^i = x\right)\right\}} \qquad \text{(A2)} \\
&\approx \left(\begin{array}{l} 125 \times \frac{4\sum_{t=1}^{n-1}\left(S_{t+1}^i - S_t^i - \widehat{\mu}_1^i(x)\right)^2\varphi\left(\frac{x-S_t^i}{h}\right)}{\sum_{t=1}^{n-1}\varphi\left(x, \frac{x-S_t^i}{h}\right)} \\ -125\frac{\sum_{t=1}^{n-2}\left(S_{t+2}^i - S_t^i - \widehat{\mu}_2^i(x)\right)^2\varphi\left(\frac{x-S_t^i}{h}\right)}{\sum_{t=1}^{n-2}\varphi\left(\frac{x-S_t^i}{h}\right)} \end{array}\right)^{\frac{1}{2}},
\end{aligned}$$

where the constants 4 and -1 come from Taylor's expansion and $\widehat{\mu}_1^i(x)$ is given in (3).

APPENDIX B

In this appendix, we show that the solution to the SDE (7) is given by (8).

We guess a solution of the type $Y_t = \exp(-\alpha t) Z_t$.

Deriving the above expression, we get :

$$dY_t = \exp(-\alpha t)\left(-\alpha Z_t dt + dZ_t\right),$$

and thus:

$$\begin{aligned}
dZ_t &= \exp(\alpha t)\, dY_t + \alpha Z_t dt \\
&= \exp(\alpha t)\left(dY_t + \alpha Y_t dt\right) \\
&= \exp(\alpha t)\left(\alpha(\theta - Y_t)\, dt + \sigma dW_t + dN_t + \alpha Y_t dt\right) \\
&= \exp(\alpha t)\left(\alpha\theta dt + \sigma dW_t + dN_t\right) \\
&= \exp(\alpha t)\left((\alpha\theta + \lambda m_1)\, dt + dM_t\right).
\end{aligned}$$

Finally, by integration the expression for dZ_t :

$$\begin{aligned}
Z_t &= Z_0 + \int_0^t dZ_s, \\
Y_t \exp(\alpha t) &= Y_0 + \int_0^t \exp(\alpha s)\left((\alpha\theta + \lambda m_1)\, ds + dM_s\right), \\
Y_t &= Y_0 \exp(-\alpha t) + (\alpha\theta + \lambda m_1)\int_0^t \exp(-\alpha(t-s))\, ds + \int_0^t \exp(-\alpha(t-s))\, dM_s, \\
Y_t &= Y_0 \exp(-\alpha t) + \left(\theta + \frac{\lambda m_1}{\alpha}\right)(1 - \exp(-\alpha t)) + \int_0^t \exp(-\alpha(t-s))\, dM_s.
\end{aligned}$$

This completes the proof of equation (8).

By definition of M_s, we know that the last term is a martingale and immediately obtain:

$$E_0[Y_t] = Y_0 \exp(-\alpha t) + \left(\theta + \frac{\lambda m_1}{\alpha}\right)(1 - \exp(-\alpha t)).$$

We now proceed to derive the formula for the conditional variance of the process :

$$\begin{aligned}
Var_0[Y_t] &= Var\left[\int_0^t \exp(-\alpha(t-s))\, dM_s\right] \\
&= \exp(-2\alpha t)\, Var\left[\int_0^t \exp(\alpha s)\, dM_s\right] \\
&= \exp(-2\alpha t)\int_0^t \exp(2\alpha s)\, d\langle M, M\rangle_s,
\end{aligned}$$

where $\langle M, M\rangle$ denotes the quadratic variation (angle bracket).

Given that we have assumed that the Brownian motion and the MPP are independent, we have :

$$\begin{aligned}
\langle M, M\rangle_t &= \langle N, N\rangle_t + \sigma^2 \langle W, W\rangle_t \\
&= \lambda m_2 t + \sigma^2 t,
\end{aligned}$$

and thus:

$$\begin{aligned}
Var_0[Y_t] &= \exp(-2\alpha t)\int_0^t \exp(2\alpha s)\left(\lambda m_2 + \sigma^2\right) ds \\
&= (1 - \exp(-2\alpha t))\frac{(\lambda m_2 + \sigma^2)}{2\alpha}.
\end{aligned}$$