Working Paper

July 2007

Imperfect Portfolio Diversification as the Cause for the Credit Spread Puzzle

Michael Kuhlen

HSH N Financial Markets Advisory AG

Küterstr. 8-12

D-24103 Kiel

Germany

Abstract

The observed credit spreads for corporate bonds are much larger than what would be required to cover just the expected loss from defaults. The nature of this gap has not yet been resolved. In this paper a heuristic one-parameter model is presented, which is able to describe the observed spreads as a function of the default probability simultaneously for all investment grade bonds. The basic model assumption is that the gap is due to imperfect portfolio diversification, giving rise to an extra risk premium to cover unexpected losses.

Introduction

The spread between corporate and government bonds has presented a puzzle to financial research since Altman (1989) noted that the expected default accounts for a surprisingly small fraction of the spread. Elton, Gruber, Agrawal and Mann (2001) decomposed the spread into several components: expected default loss, a component due to tax preferences, and a risk premium which is coupled to the systematic risk in the equity market. Delianedis and Geske (2001) concluded that the major part of the credit spread is due to recovery, jumps, tax effects, illiquidity, and market factors. Collin-Dufresne, Goldstein and Helwege (2003) argue for a large contagion risk premium. Amato and Remolona (2003) question the implicit assumption of previous studies that corporate bond portfolios are perfectly diversified. They argue that spreads are so wide because they are pricing undiversified credit risk. Hull, Predescu and White (2005) summarise the situation and emphasize that the relative importance of the various factors is still far from being settled. We may add that corporate bond portfolio managers experience year on year a varying number of defaults in their portfolios. For investment grade portfolios most years pass without a single default. Would the portfolios be perfectly diversified, portfolio managers should experience a rather continuous number of defaults, with possible time variations due to dynamic effects. We shall take the fluctuations in the observed number of defaults as a starting point for our investigation.

This paper pursues a phenomenological approach analysing the functional dependence of the corporate bond spreads on the default probabilities, based upon empirical data. It is organised as follows. The first Section discusses the data used in the analysis. In the second Section the characteristic features of the data are explored. The hypothesis is formulated that the conspicuous gap between the expected default loss and the market spread can be explained entirely by a risk premium due to imperfect portfolio diversification. In the third Section a simple model for the risk premium in terms of the degree of diversification in real world corporate bond portfolios is developed. The model parameters are the effective number of independent names in the portfolio, and a price of risk which is taken from the equity market. An excellent description of the data can be obtained with the model by adding a term proportional to the square root of the default probability to the expected default loss. Different risk measures are explored in Section IV. Robustness checks are performed in Section V. The summary is given in Section VI.

I Data Used in the Analysis

The data used in this paper are displayed in Table I. Spread data are taken as option adjusted spreads from the Merrill Lynch Global Large Cap Corporate indices rated AAA – BBB, and from the Merrill Lynch Global High Yield indices rated BB – C. The average durations of the bonds covered by these indices lie between 4.6 and 6.0. In this analysis, we use the average spreads of each rating class from the year 2005.

We use the rating classification, recovery rates and rating migrations for corporate bonds as compiled by Moody’s Investors Services (Hamilton, Varma, Ou and Cantor, 2005). Moody’s derive migration matrices based on observed migrations of rated bonds. A default can be considered as migration to a special rating category “default”. The quoted 5-year default probabilities (PD) are taken from the “default” column of the average 5-year whole letter rating migration matrix covering 1970-2004 (Moody’s exhibit 33). Moody’s calculate the default probabilities from the observed defaults of rated bonds. We calculate the annualised default probability p from Moody’s 5-year default probability p_5 as

$$p = 1 - \sqrt[5]{1 - p_5}\ . \qquad (1)$$

When multiplied with the loss given default (LGD), this value represents the annual spread which is required to cover the expected loss for a corporate bond which matures in 5 years, neglecting discounting effects.

Besides the 5-year migration matrix, Moody's also quote 1-year and 10-year migration matrices. We decided to use the 5-year migration matrix, because it matches rather well the average durations of the corporate bond indices, from which the spread data are taken. In addition, the reported 1-year default probability for bonds rated Aa or better vanishes. Such highly rated bonds never default in the first year. Only after a few years their default probability reaches a measurable size, which makes them accessible for quantitative research.

According to Hull, Predescu and White (2005), a recovery rate of 40 % is a common assumption of market participants. For simplicity, we adopt this assumption and set the LGD to a value of $L = 0.6$ for all rating classes. This value agrees roughly with Moody's data (exhibit 28), who quote on average a recovery rate of about 40 %.

A word of caution is in order. Estimates of default probabilities and recovery rates are based on historic defaults. Given the rather limited size of the data sample, at least for investment grade bonds, the estimation uncertainty on the PD and LGD may be sizeable. For investment grade bonds, Moody's counts 40 defaults in the period 1970-2004, 24 of which happened in the three years 2000-2002. For sub-investment grade bonds, the numbers are 1222 and 421, respectively.

Relating the spread from 2005 with an average default probability measured over the years 1970 to 2004 is perfectly consistent, if the past default probabilities represent a good estimate of future default probabilities, or at least that this is the expectation of market participants. This may not be the case in reality. As an alternative, one could try and compare the spread observed in one year with the default probability measured in the same year, or in a few subsequent years. In practice, however, in particular for investment grade bonds, many years of observation are needed to arrive at a statistically sound default probability. Therefore we stick with the chosen procedure for the time being. In Section V we shall check the robustness of the results against variations of the procedure.

II Data Exploration

As a first step we examine the data. The last row of Table I contains the fraction e of the observed spread, which is explained by the expected loss, $p \cdot L$:

$$e = \frac{p \cdot L}{s} = \frac{p}{s / L} \tag{2}$$

For the best rated bonds, this fraction is around 5 %. It rises towards lower ratings, reaching a level of 85 % for rating category Caa or worse.

For a closer look at the data, we concentrate first on investment grade bonds, where the gap is largest. The gap is the difference between the explained fraction e and 1. We define the spread corrected for the LGD, $\hat{s} = s / L$, which is the spread which would be required for a bond if its LGD were 1 instead of the actual value. For rating classes Aaa to Baa, the LGD corrected spread $\hat{s}$ is shown in Figure 1 as a function of their respective default probabilities p, taken from Table I.

The data lie far above the straight line which is defined by equating the spread $\hat{s}$ with the expected loss p. The area below the straight line is considered a forbidden zone; spreads on corporate bonds should at least compensate for the expected loss. The huge gap between the straight line, where spreads cover just the expected losses, and the data represents the credit spread puzzle.

We already noted that the unexplained gap is largest for the best rated bonds, and almost vanishes for junk bonds. Such behaviour would be expected, if the unexplained portion were due to fluctuations in the unexpected loss: the larger the default probability, the larger the expected number of defaults, the smaller the relative uncertainty in the number of actually observed defaults, the smaller the unexplained fraction. Could it be that the uncertainty on the number of defaults in a given year is governed by the statistical law for a limited sample size, and that this uncertainty requires a premium to be paid which can explain the gap? For example, from a Poisson process the statistical uncertainty on the relative number of "successes", defaults in our case, decreases proportional to $1/\sqrt{n}$, where n is the expected number of successes. In fact, the gap appears to follow a square root law of some kind. Poisson statistics may be sufficient for a rough attempt to grasp the data – the more appropriate binomial statistics will be used later on.

We shall consider the hypothesis that the gap can be explained fully as a risk premium for imperfect diversification in real world corporate bond portfolios. The risk premium covers the uncertainty in the actual number of defaults to be observed in a given time period. In the next section, a simple model is developed to derive the size of the risk premium from the effective number of independent names in a portfolio and the market price of risk. The result is shown as a full line in Figure 1. The data are remarkably well fitted by the model.

III Credit Spread as a Premium for Unexpected Losses

A) The simple model with standard deviation as risk measure

We consider a homogeneous corporate bond portfolio with N independent names, and assume LGD = 1 without loss of generality. The expected rate of return r from this portfolio is composed from the yield y of government bonds, plus the spread, minus the expected loss from defaults:

$$r = y + \hat{s} - p \quad . \tag{3}$$

The actual return earned is subject to the fluctuation of defaults from the portfolio. The expected number of defaults D per year is $\langle D \rangle = N \cdot p$. The standard deviation for the number of defaults is $\sigma_D = \sqrt{N \cdot p \cdot (1-p)}$ from binomial statistics. The standard deviation on the rate of return of the portfolio is therefore

$$\sigma_r = \frac{\sigma_D}{N} = \frac{1}{\sqrt{N}} \sqrt{p \cdot (1-p)} \quad . \tag{4}$$

In order to calculate the premium to be paid for the risk represented by σ_r, we need to know the price of risk. We turn to the stock market, and calculate the price of risk R from the expected excess return r_e over the return of a riskless asset, and the volatility of the stock market, quantified by the standard deviation of its rate of return σ_e:

$$R = \frac{r_e}{\sigma_e} \ . \tag{5}$$

The size of the equity premium r_e is subject to much debate. The European stock market index MSCI Europe has gained on (arithmetic) average 13.4 % p.a. in the years 1970 – 2006, with a volatility given by σ_e = 20.2 %. Dimson, Marsh and Staunton (2003 and 2006) arrive at a historic risk premium of 6.1 % with volatility 16.7 % as world average over the years 1900-2005 on an arithmetic basis. As a forward looking equity risk premium they favour a lower value of 5 % p.a. For the sake of simplicity we assume $r_e = 10\,\%$

and $\sigma_e = 20\,\%$, from which $R = 0.5$ is obtained. If one takes the historic values from Dimson, Marsh and Staunton, the value would be $R = 0.37$ instead.

If we assume that the price to carry risk is the same per unit of risk in the corporate bond market as in the stock market, we get

$$R = \frac{r_e}{\sigma_e} = \frac{r - y}{\sigma_r} = \frac{\hat{s} - p}{\sigma_r} \tag{6}$$

Solving for $\hat{s}$ and inserting eq. 4, we can express the spread corrected for LGD as a function of the number of independent names N in the portfolio, the default probability p, and the equity premium and volatility of the stock market:

$$\hat{s} = p + \frac{r_e}{\sigma_e} \cdot \sigma_r = p + \frac{r_e}{\sigma_e} \cdot \frac{1}{\sqrt{N}} \sqrt{p \cdot (1 - p)} \;. \tag{7}$$

Defining the parameter

$$g = \frac{r_e}{\sigma_e} \cdot \frac{1}{\sqrt{N}} \tag{8}$$

which comprises the price of risk and a diversification measure, we obtain a one-parameter equation for the dependence of the spread on the default probability:

$$\hat{s} = p + g \cdot \sqrt{p \cdot (1 - p)} \;. \tag{9}$$

This equation describes the spread in terms of the expected default, plus a term with a square-root dependence on the default probability which vanishes for $p \to 1$.

B) Comparison with the data for investment grade bonds

The parameterisation of eq. 9 with $g = 0.3$ is shown as full line in Figure 1, and compared with market data for investment grade bonds. The parameter $g = 0.3$ was obtained from an eye ball fit to the data. The remarkable agreement with the data is quite suggestive. It appears that imperfect diversification alone is able to explain the credit spread puzzle. Of course this does not mean that other explanations for the gap do not contribute in reality. It merely means that they are not needed for explanation.

Assuming that imperfect diversification according to our model is the true mechanism determining the observed market spreads, we can determine the parameter N from eq. 8:

$$N = \left(\frac{r_e}{\sigma_e} \cdot \frac{1}{g} \right)^2 . \qquad (10)$$

The quantity N can be considered as the effective number of independent names in the portfolio. We obtain $N = 2.8$ from the assumed stock market parameters and the empirically determined parameter g. At first glance, the value of $N = 2.8$ appears to be unrealistically low. Real portfolios typically consist of 50, but rarely more than 100 different names. However, the observed defaults since 1970 cluster in the three years 2000-2002. Given this observation, it appears possible that the number of truly independent obligors in a corporate bond portfolio is significantly lower than one might expect naively.

C) Extension to sub-investment grade bonds

In Figure 2 the sub-investment grade data are added to the investment grade data. The plot is on a logarithmic scale for both axes in order to cover the three orders of magnitude spanned by the PD data. The data for ratings B and lower are quite close to the dotted line, which stands for a model where the spread covers just the expected loss, $\hat{s} = p$.

The model parameterisation eq. 9 with $g = 0.3$ fails to describe the sub-investment grade data. Instead, the data for ratings B and Caa-C are well described with the same model, but with $g = 0.05$. The data point for Ba lies in between the two parameterisations. For PD approaching unity, all parameterisations converge; if a default becomes certain, the premium required for uncertain returns vanishes.

With the parameter g as low as 0.05, the effective number of independent names becomes sizeable. We obtain $N = 100$ from eq. 10. We note that the parameter g cannot be determined precisely from the data, which carry significant uncertainty in themselves. A parameter g of for example 0.1 would give only a slightly worse description of the data, but would result in a value of $N = 25$.

If one follows the hypothesis that N represents the effective number of independent names in a corporate bond portfolio, one might speculate that sub-investment grade portfolios are better diversified than investment grade portfolios. This conclusion is corroborated by the fact that clustering is more pronounced for investment grade bonds than for sub-investment grade bonds. The ratio between the number of defaults observed in the years 2000-2002 to all defaults between 1970 and 2004 is 34 % for sub-investment grade bonds, and 60 % for investment grade bonds.

IV Alternative Risk Measures

The conclusions we arrived at in the previous Section may depend on the employed risk measure. Standard deviation as a risk measure is simple and convenient, but may be insufficient to capture credit risk. Credit risk is characterised by an asymmetric return distribution. Large negative surprises are more likely than large positive surprises. It is quite likely that the market prices the asymmetry, and considers other risk measures like downside deviation or value at risk as more appropriate than standard deviation to assess the credit risk.

A) Downside Deviation

Downside deviation is defined in analogy to the standard deviation. The standard deviation squared is defined as the sum over all squared deviations of the observation from the average. For the downside deviation, only averse deviations from the mean contribute to the sum. The squared sum of upside and downside deviation equals the squared standard deviation. For continuous symmetric distributions, upside and downside deviation are equal.

Given a default probability p and the number of independent names N, the probability to observe k defaults is given by the binomial distribution

$$\Phi_B(k, N, p) = \frac{N!}{k!(N-k)!} p^k (1-p)^{N-k} \quad , \tag{11}$$

with mean Np and variance $Np(1-p)$. In order to calculate the downside deviation d, a sum is constructed over all observations with more defaults than expected on average:

$$d(N, p) = \sqrt{\sum_{k \geq Np} \Phi_B(k, N, p) \cdot (k - Np)^2} \quad . \tag{12}$$

We return to eq. 7, where we substitute the downside deviation for the standard deviation. Assuming that equity returns are normally distributed, we can set the downside deviation for equities equal to $\sigma_e / \sqrt{2}$. The result is a formula for the spread corrected for LGD, based on downside deviation as risk measure:

$$\hat{s} = p + \frac{r_e}{\sigma_e / \sqrt{2}} \cdot \frac{d(N, p)}{N} \, . \tag{13}$$

The spread predictions from this formula are compared to the market data in Figure 3. The sub-investment grade data are well described with $N = 5$, whereas a good description of the data for ratings B and Caa-C is obtained with $N = 100$. For investment grade bonds, the effective number of independent names is larger when downside deviation is used as risk measure than when standard deviation is used. For speculative grade bonds the difference disappears. This can be expected, because the probability distribution becomes symmetric for large N .

B) Value at Risk

Value at Risk is a frequently used risk measure when asymmetric return distributions or return distributions with fat tails need to be considered. For investment grade portfolios, the most likely event is that there is no default, due to the small default probability. The next likely event is that there is one default. Events with more than one default are suppressed with a power of the number of defaults. This makes the return distribution quite asymmetric. It may be possible that the mechanism at work in the market is based on the value at risk concept. In our case, some effort will be necessary to apply the value at risk concept to the discrete return distribution defined by the binomial distribution.

The probability $P_{D\geq 1}$ to observe one or more defaults can be expressed in terms of the probability $P_{D=0}$ of no default:

$$P_{D\geq 1} = 1 - P_{D=0} = 1 - \Phi_B(0, N, p) = 1 - (1-p)^N . \tag{14}$$

In other words, the $P_{D\geq 1}$- quantile of the return distribution is reached for $D = 1$ defaults. The value at risk, relative to the expectation value Np , at the confidence level $CL = 1 - P_{D\geq 1}$ is thus given by $D - Np$ with $D = 1$:

$$VaR_B\left(CL = (1-p)^N\right) = \frac{1}{N}(1 - Np) . \tag{15}$$

The factor $1/N$ has been introduced in order to normalise the value at risk to the total exposure. Due to the discrete character of the binomial distribution it is not possible to determine the value at risk for any given confidence level.

Again, we calculate the risk premium for the equity market and apply it to credit risk in the corporate bond market. Consider a normal distribution with mean μ and standard deviation σ. The quantile matching the probability x is given by the inverse of the cumulative normal distribution, $\Phi_N^{-1}(x,\mu,\sigma)$. For normally distributed returns with mean r_e and standard deviation σ_e the value at risk at the confidence level CL relative to the expectation value r_e is thus

$$VaR_N(CL) = \left|\Phi_N^{-1}(1-CL, r_e, \sigma_e) - r_e\right|. \tag{16}$$

Using the price of risk from the equity market, $R = r_e / VaR_N$, we can calculate the spread which is necessary to cover both the expected loss p and the unexpected loss of 1 or more defaults at the confidence level $CL = 1 - P_{D\geq 1}$:

$$\begin{aligned}\hat{s} &= p + R \cdot VaR_B\left(CL = (1-p)^N\right) \\ &= p + \frac{r_e}{\left|\Phi_N^{-1}(1-CL, r_e, \sigma_e) - r_e\right|} \cdot \left(\frac{1}{N} - p\right) \\ &= p + \frac{r_e}{\left|\Phi_N^{-1}(1-(1-p)^N, r_e, \sigma_e) - r_e\right|} \cdot \left(\frac{1}{N} - p\right)\end{aligned} \tag{17}$$

With eq. 17 we have derived the spread corrected for LGD $\hat{s}$ as a function of the default probability p and the effective number of names in the portfolio N, using value at risk as risk measure. Note however, that in our approach the confidence level defining the value at risk is also a function of p and N.

In Table II we give the effective number of names which are needed to explain the observed spreads from the quoted default probabilities, if eq. 17 were used to price credit risk. Also given are the resulting confidence levels. We obtain values between $N = 43$ for Aaa rated bonds and $N = 18$ for bonds rated Baa. The corresponding confidence levels run from 99.1 % to 93.9 %.

We have restricted ourselves to investment grade bonds, for which the unexplained gap $\hat{s} - p$ is largest. For sub-investment grade portfolios the probability to observe one or more defaults would be so large that it does not make sense to apply the value at risk concept according to eq. 15. Assuming $N = 20$, the confidence level for which the value at risk would have to be calculated would be as low as 70 % for rating Ba.

To conclude, the gap between observed spreads and quoted default probabilities can also be explained with a risk premium for unexpected losses, when value at risk is used as risk measure. Roughly speaking, the effective number of names needed to explain the gap is between 20 and 40 for investment grade bonds, significantly larger than in the case of standard deviation or downside deviation as risk measure. However, contrary to the other risk measures, it is not possible to describe the data with a unique number of effective names for all rating categories within investment grade.

V Robustness of the Analysis

In this Section we shall examine how robust the results are against alterations in the procedure to analyse the data and in the data selection. In Section II we argued that it may not be completely satisfactory to compare the default data from an extended period of time to spread data from one particular year. Hence we now analyse the spread for each rating class averaged over the available time series from 1997-2006.

One may also argue that the assumption of the same LGD for all rating classes is too simple. We shall now compare the spread data not to the PD data, but rather to the loss rate which comprises both the PD and the LGD. When PD and LGD are known, the loss rate can be calculated as PD times LGD. In this analysis, the loss rate is extracted from the average cumulative credit loss rates compiled by Moody's Investors Services (Hamilton, Ou, Kim and Cantor, 2007, exhibit 8), covering the period 1982-2006. This period is smaller than the period analysed previously, but still larger than the period covered by the spread data. The annualized loss rate l is calculated from the 5-year loss rate in analogy to eq. 1. The spread data and the loss rates used in this section are given in Table III.

Equation 9 can be transformed into a relation between the spread and the loss rate:

$$s = pL + g \cdot \sqrt{pL \cdot (1-p)L} = l + g\sqrt{l \cdot (1-p)L} \; . \qquad (18)$$

For sufficiently small PD, we can approximate

$$s \approx l + g\sqrt{l} \; . \qquad (19)$$

In Figure 4 the average market spread is plotted against Moody's loss rate for the different rating classes. The data are compared to the model prediction eq. 19, with the same parameters $g = 0.3$ and $g = 0.05$ as previously. Also shown is the line representing the naïve parameterisation $s = l$. Qualitatively, the data exhibit the same features as discussed in Section II. Using the same parameters as obtained in Section III, we still obtain a satisfactory description of the data, with a slight loss in precision. We conclude that the results obtained above are fairly insensitive to details on the data selection and analysis method.

VI Summary

This research addresses the credit spread puzzle, the nature of the large gap between observed credit spreads and the premium required to cover expected default losses. The hypothesis is examined that the gap is entirely due to imperfect portfolio diversification. A simple model is developed in which the gap is regarded as a risk premium covering the uncertainty in the number of defaults to be expected. The hypothesis is tested by comparing market spread data and observed default and loss rates with the model predictions. The following results are obtained.

For investment grade bonds, where the gap is largest, the dependence of the corporate bond spreads on the default rates can be remarkably well described with a one-parameter ansatz, adding a term proportional to the square root of the default probability to the default probability. The functional dependence can be explained with a model in which the gap is due to a risk premium required for imperfect portfolio diversification, leading to fluctuations in the number of actual defaults. The fluctuations are calculated using binomial statistics, and risk is measured as standard deviation. With a price of risk analogous to the stock market, the effective number of independent names in an investment grade portfolio would be around three.

The ansatz works also for speculative grade bonds, but with a larger number of independent names, around 100. Following our model, the data indicate that portfolio diversification is much better in speculative grade portfolios than in investment grade portfolios. A possible economic explanation would be that contagion reduces the effective number of names in realistic corporate bond portfolios significantly, and that investment grade bonds are more subject to contagion than sub-investment grade bonds.

Downside deviation and value at risk have been studied as alternative risk measures, possibly better suited to capture the asymmetry of credit risk. With downside deviation as risk measure, similar results are obtained as in the case of standard deviation, but with a somewhat larger number of independent names. Value at risk applied to investment grade bonds can also explain the gap, but not with a unique number of independent names for all rating classes. Rather, the number varies between 20 and 40.

It has been demonstrated that the credit spread puzzle can be resolved entirely with a model in which the unexplained gap represents a risk premium to cover unexpected losses due to a lack of diversification in corporate bond portfolios. This does not mean that other effects which have been brought forward for explanation do not exist in reality. Rather, their existence cannot be concluded with certainty, because the model presented here can explain the data without such assumptions.

In summary, we have obtained quantitative support for the hypothesis of Amato and Remolona (2003), that the credit spread puzzle can be explained to large extent as a risk premium for imperfect portfolio diversification. It appears that portfolio diversification is much better for sub-investment grade portfolios than for investment grade portfolios. If contagion, the mechanism proposed by Collin-Dufresne, Goldstein and Helwege (2003), is responsible for imperfect portfolio diversification one is lead to conclude that contagion is much more severe for investment grade bonds than for sub-investment grade bonds.

References

Altman, Edward I., 1989, Measuring Corporate Bond Mortality and Performance, *Journal of Finance,* 44, 902-922.

Amato, Jeffery D., and Eli M. Remolona, 2003, The Credit Spread Puzzle, *Bank for International Settlement Quarterly Review*, December 2003, 51-63.

Collin-Dufresne, Pierre, Robert S. Goldstein, and Jean Helwege, 2003, Is Credit Event Risk Priced? Modelling Contagion via the Updating of Beliefs, *Working Paper, Carnegie Mellon University*.

Delianedis, Gordon, and Robert Geske, 2001, The Components of Credit Spreads: Default, Recovery, Tax, Jumps, Liquidity, and Market Factors, *UCLA Anderson Working Paper,* No. 22-01.

Dimson, Elroy, Paul Marsh, and Mike Staunton, 2003, Global Evidence on the Equity Risk Premium, *Journal of Applied Corporate Finance*, 15.

Dimson, Elroy, Paul Marsh, and Mike Staunton, 2006, The Worldwide Equity Premium: A Smaller Puzzle, *London Business School*

Elton, Edwin J, Martin J. Gruber, Deepak Agrawal, and Christopher Mann, 2001, Explaining the Rate Spread on Corporate Bonds, *Journal of Finance,* 56, 247-278.

Hamilton, David T., Praveen Varma, Sharoun Ou and Richard Cantor, 2005, *Default and Recovery Rates of Corporate Bond Issues, 1920-2004*, Special Comment, Moody's Investors Service.

Hamilton, David T., Sharoun Ou, Frank Kim and Richard Cantor, 2007, *Corporate Default and Recovery Rates, 1920-2006*, Special Comment, Moody's Investors Service.

Hull, John, Mirela Predescu, and Alan White, 2005, Bond Prices, Default Probabilities and Risk Premiums, *Journal of Credit Risk,* Vol. 1, No. 2, 53-60.

Tables

Table I: *Spread and default data for corporate bonds*. For each of Moody's rating categories the 5-year probability of default (PD) and the annualised PD is given as well as the loss given default (LGD), as they are used in this paper. The PDs are extracted from the rating migration matrix based upon Moody's analysis of data covering 1970-2004. Also shown is the spread over government bonds (average from 2005) required for corporate bonds in the respective rating categories, as well as the spread corrected for LGD (Spread/LGD). The last row shows the fraction e of the spread which can be explained by the expected loss, PD x LGD / Spread.

	Investment grade				Speculative grade		
Rating	Aaa	Aa	A	Baa	Ba	B	Caa-C
5-year PD: p_5 (%)	0.11	0.21	0.43	1.72	8.12	20.58	42.85
Annualised PD: p (%)	0.02	0.04	0.09	0.35	1.68	4.50	10.59
LGD: L (%)	60	60	60	60	60	60	60
Spread: s (%)	0.30	0.43	0.61	1.24	2.45	3.32	7.46
Spread/LGD: $\hat{s}$ (%)	0.50	0.72	1.02	2.07	4.08	5.53	12.43
Explained fraction: e (%)	4.4	5.9	8.5	16.8	41	81	85

Table II: *Effective number of names in a corporate bond portfolio for the value at risk ansatz.* For each of Moody's investment grade rating categories the effective number of names is shown, which would explain the gap between observed spread and quoted default probabilities using a value at risk ansatz to calculate a risk premium for unexpected losses. Also shown is the corresponding confidence level.

	Investment grade			
Rating	Aaa	Aa	A	Baa
Effective number of names: N	43	32	26	18
Confidence level: CL (%)	99.1	98.7	97.7	93.9

Table III: *Spread and loss rate for corporate bonds*. For each of Moody's rating categories the 5-year loss rate and the annualised loss rate is given The loss rates are extracted from Moody's cumulative credit loss rates, covering 1982-2006. Also shown is the spread over government bonds required for corporate bonds in the respective rating categories, averaged over 1997-2006.

	Investment grade				Speculative grade		
Rating	Aaa	Aa	A	Baa	Ba	B	Caa-C
5-year loss rate (%)	0.034	0.106	0.264	1.17	6.37	15.74	34.60
Annualized loss rate: l (%)	0.0068	0.0212	0.053	0.234	1.31	3.37	8.14
Spread: s (%)	0.38	0.50	0.80	1.46	3.16	5.07	11.66

Figures

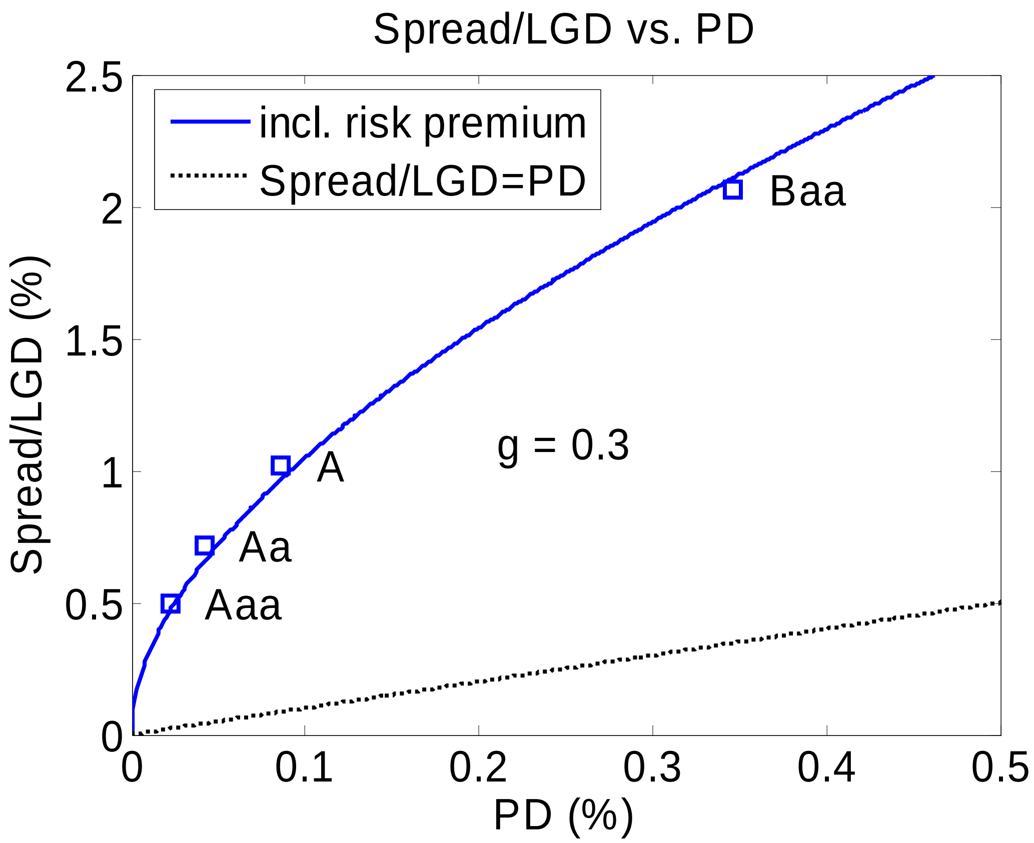

Figure 1: *Spread/LGD as a function of PD for investment grade corporate bonds*. The spread corrected for LGD, $\hat{s}$, is shown as a function of the default probability p for corporate bonds according to Moody's rating categories. The straight line is defined by Spread/LGD = PD. The full line includes a risk premium for unexpected losses, derived from the binomial default model explained in the text, where the parameter g is set to 0.3. Spread data are averages from 2005. PD data are extracted from Moody's data covering 1970-2004.

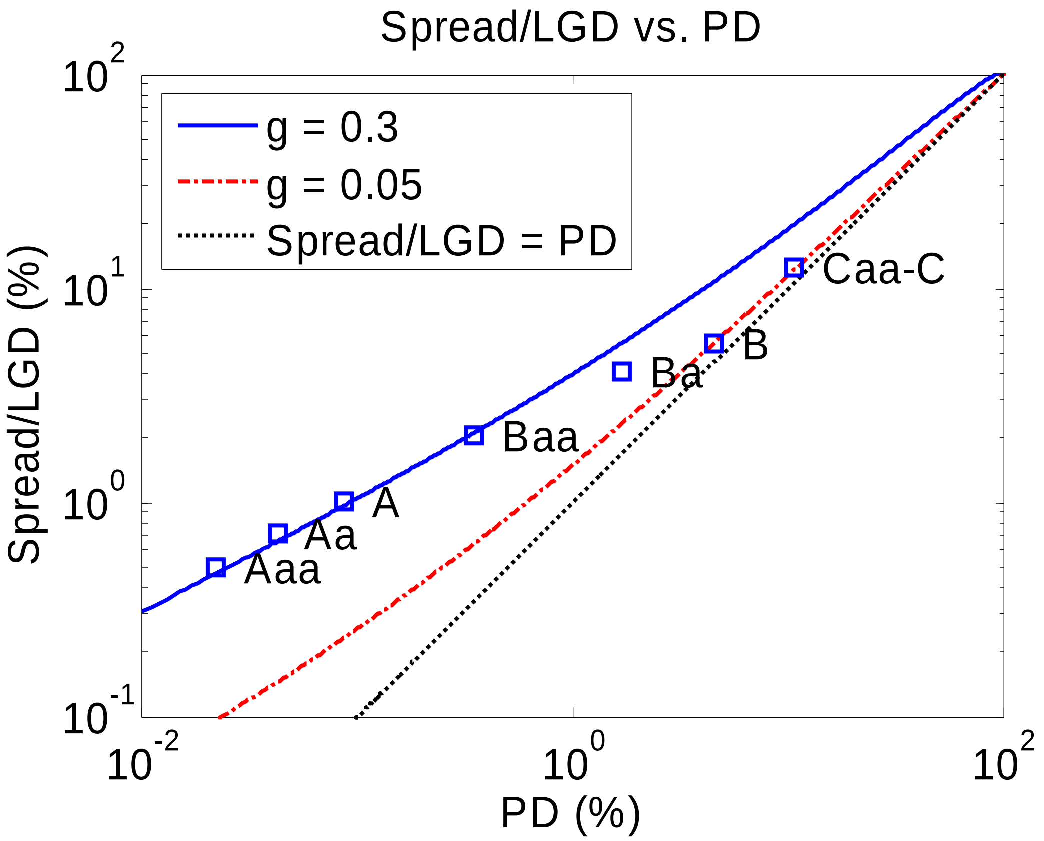

Figure 2: *Spread/LGD as a function of PD for ratings Aaa to C on a log-log scale*. The spread corrected for LGD, $\hat{s}$, is shown as a function of the default probability p for corporate bonds according to Moody's rating categories. The dotted line is defined by Spread/LGD = PD. The full line includes a risk premium for unexpected losses, derived from the binomial default model explained in the text, where the parameter g is set to 0.3. The dash-dotted line represents the same model, but with the parameter g set to 0.05. Spread data are averages from 2005. PD data are extracted from Moody's data covering 1970-2004.

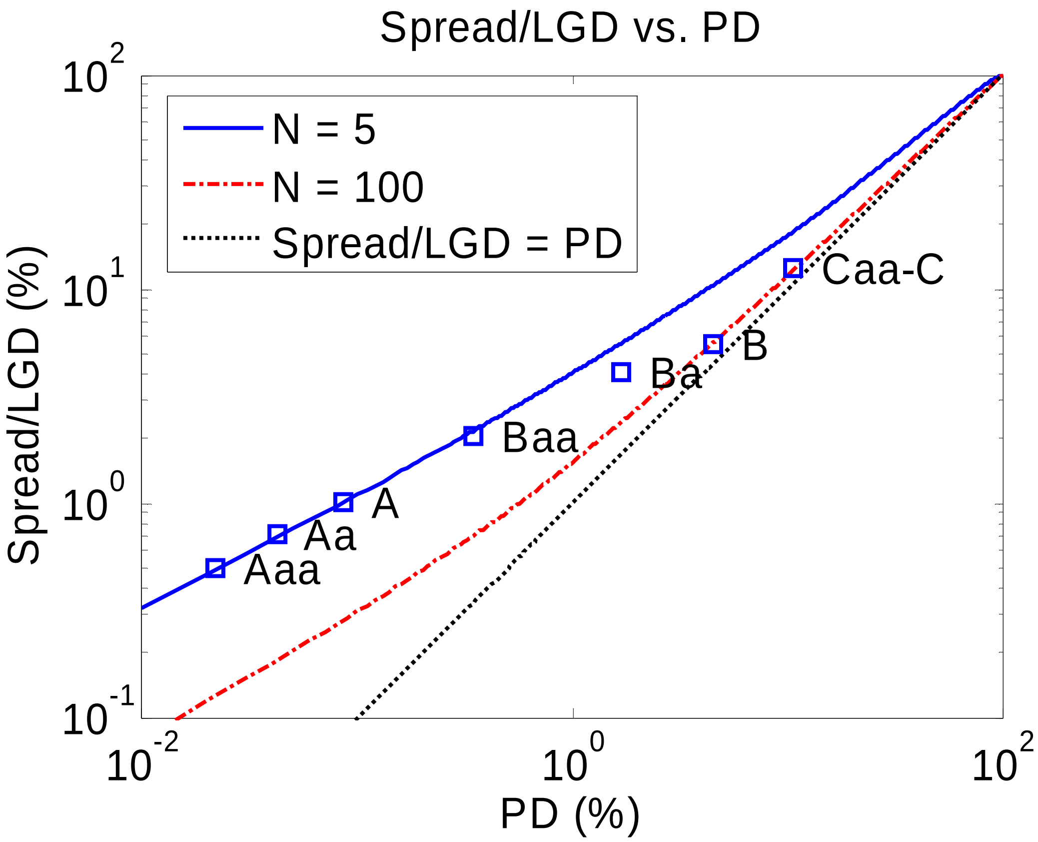

Figure 3: *Spread/LGD as a function of PD, compared to a model based on downside deviation as risk measure*. The spread corrected for LGD, $\hat{s}$, is shown as a function of the default probability p for corporate bonds according to Moody's rating categories. The dotted line is defined by Spread/LGD = PD. The full line includes a risk premium for unexpected losses, derived from the binomial default model explained in the text with downside deviation as risk measure and the parameter N set to 5. The dash-dotted line represents the same model, but with the parameter N set to 100. Spread data are averages from 2005. PD data are extracted from Moody's data covering 1970-2004.

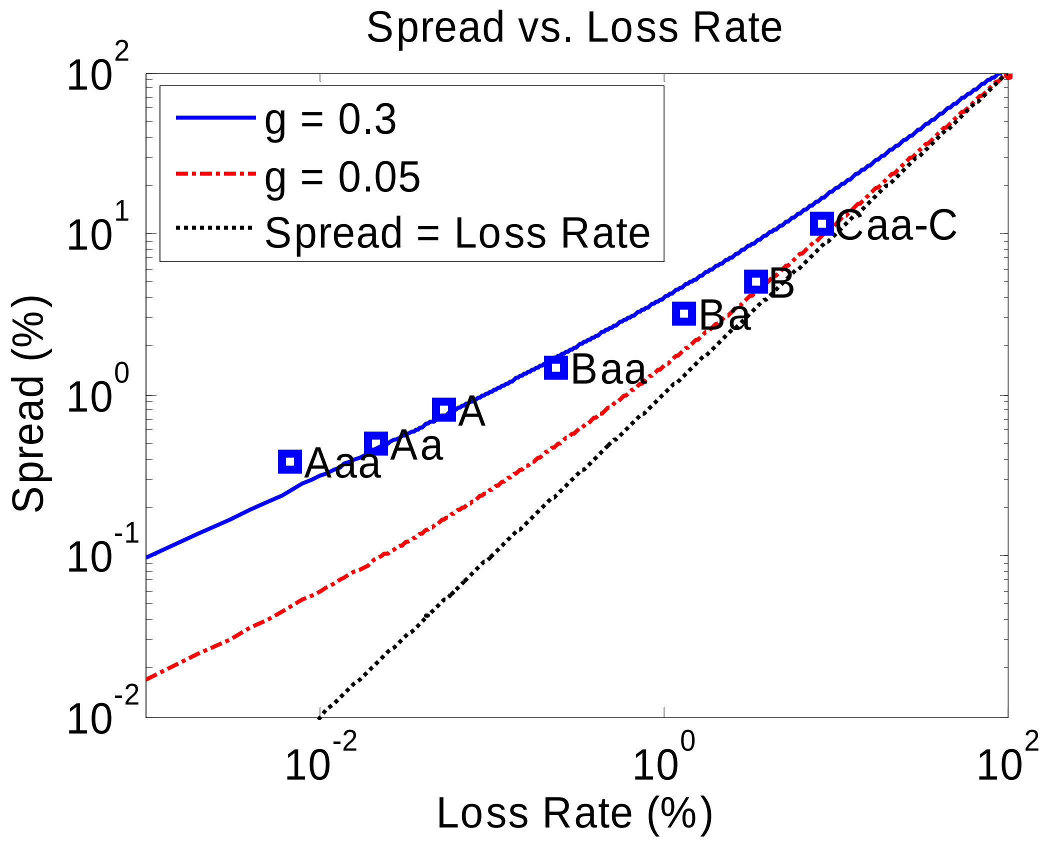

Figure 4: *Spread as a function of the loss rate for ratings Aaa to C.* The spread s is shown as a function of the loss rate l for corporate bonds according to Moody's rating categories. The dotted line is defined by spread = loss rate. The full line includes a risk premium for unexpected losses, derived from the binomial default model explained in the text, where the parameter g is set to 0.3. The dash-dotted line represents the same model, but with the parameter g set to 0.05. Spread data are average spreads from 1997-2006. The loss rates are from Moody's data covering 1982-2006.