In press, Journal of Banking and Finance

Ratings versus market-based measures of default risk in portfolio governance ✧

Gunter Löffler*

University of Ulm

Abstract

This paper assesses whether ratings or market-based credit risk measures are more suitable for formulating portfolio governance rules. Such rules, which consist of buy and sell restrictions, are commonly used in investment management. Based on data from 1983 to 2002, it is not evident that one of the two measures is superior. The relative power of the two measures in predicting defaults depend on the investor's investment horizon and risk appetite. The results support the agencies' claim that their policy of reducing rating volatility, which builds on the though-the-cycle approach and the avoidance of frequent rating reversals, is beneficial to bond investors. The results also suggest that widely used statistical measures of rating quality may be insufficient to judge the economic value of rating information in specific contexts.

Key words: credit ratings, rating agencies, investment restrictions.

JEL classification: G20, G33

* Gunter Löffler, Department of Finance, University of Ulm, Helmholtzstrasse 18, 89069 Ulm, Germany. Phone: ++49-731-5023597. e-mail: loeffler@mathematik.uni-ulm.de

✧ I thank Richard Cantor, Christopher Mann and a referee for helpful comments. Financial support from Moody's Investors Service is gratefully acknowledged.

1 Introduction

Recently, the quality of credit ratings produced by agencies such as Fitch, Moody's or Standard & Poor's has met a lot of skepticism. Spectacular cases like the default of Enron Inc. seem to confirm the view that rating agencies react too slowly or neglect relevant information. Agencies have responded by reassessing their rating practice, but also by clarifying their rating objectives. Moody's, for example, points out that rating more aggressively would increase rating volatility (Fons, 2002). This could be damaging for long-term bond investors, the traditional clientele of rating information.[1] Many pension or mutual funds, for example, are only allowed to purchase bonds with an investment-grade rating; if bonds are downgraded, they often have to be sold. Clearly, timely and informative ratings increase the effectiveness of governance rules. The more volatile a rating system, however, the more transactions are triggered by portfolio governance rules, making it more costly.

This points to a deficit in the discussion on rating quality. Most contributions claim that ratings can be improved with regard to individual aspects such as timeliness or accuracy. The ultimate question, however, is whether users of ratings would benefit if they switched to alternative rating information, or if agencies rated differently. In this paper, I aim at answering this question from the perspective of investors who invest in funds that are subject to eligibility guidelines. Using a large and comprehensive data set on corporate issuers I compare the performance of agency ratings and market-based measures of default risk that, besides issuer balance-sheet data, use equity prices and equity volatilities.

By simulating bond portfolios that are subject to governance rules, I examine whether investors would be better off if they used market-based measures instead of ratings to define investment restrictions. Results depend on the performance metric, the specification of governance rules, the investment horizon and the magnitude of transaction costs. Though the results are mixed, three conclusions emerge: (i) From the past twenty years analyzed in this study, it is not evident that ratings are generally inferior to market-based rules. While there are market-based guidelines that lead to better performance than ratings, ratings lead to lower losses than *some plausible* market-based guidelines; (ii) Since ratings are less volatile than market-based measures, ratings tend to produce lower trading costs than market-based strategies; combining the latter with filter rules that reduce transactions can reduce, or eliminate, this advantage of ratings; (iii) widely-used statistical measures of rating quality

[1] A lower rating volatility may also benefit debtors because a rating change can entail irreversible costs even if it is reversed after a short period of time.

should be interpreted with care because there may not be a one-to-one relationship between these measures and the economic value of default risk indicators in a specific context.

The present study does not incorporate actual bond prices. Performance analysis rests on losses from default, or on hypothetical prices derived from the market-based measures of default risk. In the former case, it is important to note that an analysis of losses does not explicitly take the risk-return trade-off into account, i.e., if strategy *A* leads to smaller losses than strategy *B*, it does not follow that strategy *A* has a larger expected return, or a better risk-return profile. For the analysis of returns, I derive prices through a structural model which translates actual default probabilities into risk-neutral probabilities, or exploit the empirical relationship between market-based risk measures and credit spreads. While this is a limitation of the analysis, it also has two advantages. First, reliable bond price data are scarce, so that the sample is larger than the ones that could be formed with bond price information. Second, since rating-based governance rules are widely used in practice, they could influence prices. If this is indeed the case, it could be problematic to compare the actual performance of widely followed rating-based strategies with a hypothetical market-based strategy that few, if any, followed in the time period under analysis.[2]

The paper is related to several strands of literature. One strand shows that agency ratings can be predicted or improved using public information. Altman and Kao (1992) and Lando and Skødeberg (2002) document drift in rating changes. Delianedis and Geske (1999) show that rating changes are predictable using a market-based measure of default risk. Galil (2002) finds that issuer characteristics like size or leverage increase the default prediction power of ratings.

Another strand of the literature (Kealhofer, 2003a, and Cantor and Mann, 2003) shows that market-based risk measures may be better predictors of short-term default risk than agency ratings. For several reasons, the reported results are not sufficient to conclude that market-based measures are superior for formulating investment restrictions: The analysis corresponds to examining one-year buy-and-hold policies, so that it is uninformative about the effectiveness of selling rules or purchase restrictions on longer-term bonds; the analysis is done over the entire issuer universe whereas investors are often restricted to a subset of issuers; the one-year horizon often chosen for such an analysis may not conform to an

[2] Of course, ratings can still have repercussions on model prices via stock market reactions (which influence market-based risk measures) to rating changes.

investor's horizon; the analysis does not incorporate transaction costs; finally, the statistical measures of predictive power used by Kealhofer (2003a) and others are difficult to translate in performance metrics relevant for bond investors. The latter point is illustrated in Stein and Jordão (2003) who show that, while more powerful models result in meaningful economic differences in portfolio returns, there is no one-to-one relationship between a commercial bank's economic profitability and the difference in predictive power between two credit scoring systems. Moreover, as Hamilton and Cantor (2004) demonstrate, the predictive power of ratings for default is, in fact, much greater if one takes into account rating outlooks and watchlist designations.

Other studies examine the link between ratings and capital market prices. The main conclusion is that rating changes are largely anticipated in bond prices (for a summary see Hull, Predescu and White, 2004) as well as in credit default swap spreads (Hull, Predescu and White, 2004, and Norden and Weber, 2004). Campbell and Taksler (2002) show that a simple measure such as equity volatility can explain as much cross-sectional variation of bond spreads as can ratings. Taken together, these studies also cast doubt on the informativeness of ratings, but again, the results are not sufficient to establish that investors should not base their decision on ratings. Some studies examine whether market-based measures can be used to design profitable trading strategies (e.g. Vasicek, 2001, Rappoport and Sirinathsingh, 2002). While they suggest that investors should not rely solely on ratings, they do not imply that ratings should not be used for setting limits to asset allocation.

The literature is largely silent on the rationale for governance rules, and the requirements that rating systems should fulfill when used for formulating such rules. Governance rules are used by regulators for controlling the insolvency risks of financial institutions (see Cantor and Packer, 1997, for an overview). They may mitigate agency problems between fund managers and investors as fund managers have less opportunity to follow opportunistic, high-risk strategies. Similar to the use of benchmarks and styles in the equity fund domain, they could also ease the fund manager selection and evaluation problem. It is not obvious whether ratings of specialized agencies or market-based measures are more suitable for formulating such rules. The most simple market measure would be spreads in case of traded bonds. Rating agencies do not claim that ratings are superior because they are ahead of the market. According to rating agencies, one advantage of ratings is that they balance the conflicting aims of timeliness and stability (Cantor, 2001). The timelier a rating system is, the better is the discrimination between high-risk and low-risk issuers; the more stable a rating system is, the fewer transactions that are triggered by governance rules, leading to lower transaction costs. Through simulations, Löffler (2003, 2004) investigates whether rating policies aimed at a reduction of rating volatility could explain the apparent deficiencies of ratings. Specifically, he examines the through-the-cycle approach and the propensity to avoid frequent rating reversals. Löffler does not conduct a cost-benefit analysis of these policies, but he shows that they increase rating stability and decrease default prediction power.

The remainder of the paper is structured as follows. Section 2 describes the data. Section 3 compares the performance of rating-based and market-based governance rules. Section 4 concludes.

2 Data

The data set used in the analysis contains monthly information on Moody's long-term ratings and a market-based measure of default risk, EDFs published by Moody’s KMV (“MKMV,” hereafter). An EDF (expected default frequency) is an estimate of a firm's one-year probability of default that is based on balance sheet data, equity market valuation and equity volatilities.[3] The data cover the years 1980 to 2002 and contain both US and non-US corporate issuers. Each entry of the data set has both rating and EDF. In 1982, Moody's refined its rating system by adding rating modifiers. As this led to rating changes that are unrelated to new information, I decided to start the analysis in 1983. Various consistency checks of the data motivated a few minor modifications summarized in appendix A.

Below, descriptive information is given both for the entire sample, and for groups of issuers that also define admissible investment universes in the ensuing analysis. Each month, issuers are categorized in the following ways:

- issuers with a rating better than Ba1 (i.e., investment-grade issuers)
- issuers with EDF $\leq$ 0.5%
- issuers with EDF $\leq$ 0.75%
- issuers with EDF $\leq$ 2%
- issuers with a rating better than B1

Splitting issuers into investment-grade and non-investment grade issuers is the most commonly used categorization based on agency ratings. The EDF cutoff of 0.5% is motivated by the fact that the average default rate of Baa3 issuers from 1983 to 2002 was 0.49% (Hamilton and Varma, 2003). The 0.75% cutoff is used in addition because the universe with

EDF below 0.5% has fewer issuers and lower default risk (as measured by EDF) than the investment-grade universe. According to MKMV,[4] many investors that use EDFs employ a 2% cutoff, which motivates the next cutoff; in terms of the number of observations, this almost exactly corresponds to a rating boundary set at B1. Since the split into investment-grade and non-investment grade is the dominant one in practice, the analysis in section 3 will focus on the relative performance of the investment-grade cutoff.

Figure 1 plots the evolution of the number of issuers, both for the entire sample and the first three issuer groups defined above. The overall number of issuers rises from 461 in 1983 to 1960 in 2002. Classifying issuers according to EDFs produces larger variation in the number of included issuers than a rating-based classification. Figure 2 shows median EDFs across time. Since the investment-grade universe exhibits the largest swings in median EDFs and prior literature concludes that EDFs are better measures of short-term default risk than ratings, the figure already suggests that a rating-based universe leads to larger time-variation in short-term default risk than an EDF-based universe.

Descriptive statistics are given in Table 1. Number of "observations" means number of issuer-months comprised by a universe. With regard to this number of observations, the investment-grade universe lies between the ones defined with the 0.75% EDF cutoff and the 0.5% EDF cutoff, respectively. The universe with the 0.75% cutoff has the same median EDF (0.17%) as the investment grade universe. The latter, however, has a substantially larger arithmetic mean EDF than the former (0.41% vs. 0.23%). Issuers with an EDF below 2% have almost the same mean EDF as investment-grade issuers. This gives a justification for comparing the 2% cutoff with the B1 rating cutoff, which produces about the same number of observations, and the investment-grade cutoff.

The table also shows the number of defaults that occurred following inclusion in a universe. The fact that the number of defaults is fairly low in the lower-risk universes means that the results of section 3 are heavily influenced by a few events. This will be considered in the interpretation of the results.[5]

I also check whether ratings or EDFs are better predictors of default. Figure 3 shows one-year and five-year cumulative accuracy profiles for ratings and EDFs. An accuracy profile is

---

3 For a description of EDFs, see Crouhy, Galai and Mark (2000) or Kealhofer (2003a).

4 Information from personal communication.

5 Note that the number of defaults cannot directly be related to the number of observations to produce default rates.

obtained by plotting the proportion of defaults that occurred among issuers ranked *x* or worse, against the proportion of issuers ranked *x* or worse (see Sobehart et al., 2000). To construct the curve, all observations were pooled.[6] The more northwestern the curve, the more accurate the relative ranking of default risk achieved by the rating system. When two accuracy curves intersect, it implies that one rating system is more effective at ranking risk at a particular portion of the issuer universe and the other rating system is more effective elsewhere – it is not possible to unambiguously compare their relative predictive power except in reference to a particular portion of the issuer universe.

For a one-year horizon, the EDF curve is uniformly above the rating curve in the high risk region (speculative-grade credits), but the curves intersect repeatedly in the low risk region. Without further analysis, it is unclear whether EDFs or ratings are better at predicting defaults among low-risk issuers; due to investment restrictions, this question is the relevant one for many bond investors.

For a five-year horizon, the curves intersect at a point roughly corresponding to Ba3 rated issuers, suggesting that market-based ratings have done better in discriminating among higher risk defaulters, and ratings have done a better rank ordering among the lower risk defaulters. The shift in relative power away from market-based ratings toward credit ratings as horizon lengthens is consistent with the rating agencies' stated objective that they focus on a long-term horizon.

Together, the results already suggest that it may be difficult to use accuracy profiles for judging the economic value of a rating system in a specific context. Though it may be possible to extract the necessary information from such an analysis, it is not yet clear how this should be done. First, one has to determine how relevant the various regions of the accuracy profiles are. For some creditors, for example, the admissible universe is more or less restricted to low risk issuers. Second, it is not evident which time horizon should be used in the analysis of accuracy profiles—even if the time horizon of a creditor were known.[7] For these reasons, it seems important to employ an evaluation strategy that is closely linked to the actual usage of rating information.

[6] Restricting the analysis to December cohorts yields identical patterns.

[7] Due to defaults or rating withdrawals, for example, an investor having an *x*-year horizon may nevertheless rebalance his portfolio within the horizon, making his effective horizon less than *x* years.

3 Performance of rating-based and market-based strategies

3.1 General setup

I use simulated portfolios to compare the performance of rating-based and EDF-based strategies. Having defined a governance rule for portfolio investments, I randomly select portfolios and track their performance for a pre-specified period (1 to 5 years). Throughout the analysis, the number of issuers per portfolio is set to 50, which is chosen to be representative of fixed income funds. Portfolios are equally-weighted, and selected at month end by drawing without replacement from the admissible issuers.

For one-year holding periods, formation dates span from January 1983 to December 2001; for five-year holding periods, the last formation date is December 1997. I use all monthly cohorts, not just year-end cohorts. Results are based on 996 simulation trials (=83 trials for each month of the year); each trial comprises up to 19 portfolios formed from 1983 to 2001. As an illustration, the following exhibit summarizes the output of *one* simulation trial if the month of the year chosen for formation is November and the holding period is one year:

Formation date	11/83	11/84	11/85	.	.	.	11/99	11/00	11/01
Simulated portfolio return	11/83- 11/84	11/84- 11/85	11/85- 11/86	.	.	.	11/99- 11/00	11/00- 11/01	11/01- 11/02

I start by examining the use of ratings and EDFs for buy-and-hold policies with the following purchase restrictions:

(i) Buy only issuers with rating better than Ba1.

(ii) Buy only issuers with EDF ≤ 0.5%.

(iii) Buy only issuers with EDF ≤ 0.75%.

(iv) Buy only issuers with EDF ≤ 2% .

(v) Buy only issuers with rating better than B1.

With a buy-and-hold policy, it is irrelevant whether an issuer leaves the universe after inclusion in the portfolio. I then go on by combining purchase rules with selling rules: issuers have to be replaced[8] immediately after leaving the universe, after 6 months, or after 12 months. In addition, I examine the following EDF-based rule with a transaction-reducing component: buy only issues with EDF≤ 0.5%, sell if the EDF increases above 0.75%. If issuers re-enter the universe within the tolerance period, the tolerance rules applies anew. The following timeline illustrates the rule (gray shaded cells mark months in which the issuer belongs to the universe):

	Part of the universe ("in") or not ("out")												
	in	in	in	out	in	out	out	out	out	Out	out	out	out
0 month tolerance	buy	hold	hold	sell	-	-	-	-	-	-	-	-	-
6 months tolerance	buy	hold	hold	hold	hold	hold	hold	hold	hold	Hold	sell	-	-
12 months tolerance	buy	hold	hold	hold	hold	hold	hold	hold	hold	Hold	hold	hold	hold

Issuers are also replaced if the observation series ends, or is interrupted; in such cases, I do not charge trading costs. More precisely, issuers are replaced in month *t* if there is no observation in month *t+1*. There are various possible reasons for this to occur: the issuer is no longer rated by Moody's; missing data prevent the computation of EDFs; data errors. Replacing issuers with missing observations could create a selection bias. In particular, it could lead to lower losses if such cases are associated with a higher default probability. To assess whether such a correlation is present in the data, I use a probit analysis to examine whether missing observations or the termination of a time series predict defaults. In doing so, I also control for default risk because, through investment restrictions, the simulated portfolio strategies control for default risk as well. Performing the analysis for non-overlapping one-year horizons starting in December yields (t-statistics in parentheses, constant not reported, number of observations=22059):

$$\text{Prob(Default)} = \underset{(1.89)}{0.131}\cdot\text{MISS}$$

$$\text{Prob(Default)} = \underset{(-0.79)}{-0.059}\cdot\text{MISS} - \underset{(-16.09)}{1.421}\cdot I_{\text{Rating better Ba1}}$$

$$\text{Prob(Default)} = \underset{(0.63)}{0.047}\cdot\text{MISS} - \underset{(-14.03)}{1.538}\cdot I_{\text{EDF}\leq 0.5}$$

For defaulting (non-defaulting) issuers, the dummy variable MISS takes the value 1 if the observation series is interrupted at any time before the default month (during the following 12 months); the other indicator variables take the value 1 if rating or EDF is better than Ba1 or 0.5%, respectively. If MISS is the sole regressor, its predictive power is only significant at the 10% level; both the size of coefficient and its significance are strongly reduced if default risk is controlled for in a way that corresponds to the investment restrictions used in the ensuing

[8] Issuers are replaced though random draws from the issuers contained in the universe, but not in the current portfolio.

analysis. Regressions for other calendar months confirm these results. It thus appears that replacing issuers with missing observations does not lead to a selection bias.

When trading costs are considered, I mostly set round-trip transaction costs to 1%. This is based on the study of Chen and Wei (2001), who estimate median (arithmetic average) round-trip transaction cost for corporate bonds to be 0.59% (1.90%).[9] Previous estimates are smaller (e.g. Schultz, 2001), but they are likely to be downward biased because they exclude illiquid bonds. The base-case value of 1% is set above the median obtained by Chen and Wei because their estimates of transaction costs go down over time from 1.61% in 1990 to 1.23% in 1995 and to 0.11% in 2000. The median of 0.59% is thus largely affected by the late 1990s, which make up only one quarter of the time period examined in the present paper.

To examine how sensitive results are to the size of trading costs, I choose two different routes. I report the size of transaction cost that equates the performance of rating- and EDF-based portfolio strategies. Computationally, this is only feasible for the case in which the performance analysis is restricted to losses from default (section 3.2). If the performance analysis is extended to capture price effects (section 3.3), I also run the analysis with round-trip costs of 0.5%. Transaction costs do not arise if issuers are sold because they leave the universe or have missing observations.

3.2 Losses from default

In this section, the analysis of simulated portfolios is restricted to losses which occur upon default, or through transactions induced by sell restrictions. The price of each bond is either one if the issuer is not in default or 0.5 if the issuer is in default. I report average loss rates as well as 99% and 99.9% loss quantiles. (A 99% quantile of *x* means that in 1% of all trials, losses are greater or equal to *x*.)

[9] The median rating in the sample of Chen and Wei is between A2 and A3.

Buy-and-hold strategies

Table 2 summarizes loss rates under various buy-and-hold policies. Thus, there are no secondary market transactions and transactions costs are therefore not pertinent. Note first that simulation error is negligible. Using a *t*-test for the hypothesis that the simulated average loss rate under the investment-grade rule is equal to the one under an EDF-based strategy produces statistically significant (0.1%) *t*-values for most pairs of strategies.[10] The differences are insignificant only in one case, in which they are small in economic terms (0.297% vs. 0.296%). This suggests that the chosen number of trials is sufficient.

Comparing the investment-grade cutoff with the 0.5% and 0.75% EDF cutoffs, the following patterns emerge: At a one-year horizon, EDF strategies perform better as they lead to lower average loss rates than the rating cutoff, and, in some cases, to lower extreme quantiles. With the three-year horizon, differences in average and extreme losses are small with the rating-based strategy performing somewhat better. At a five-year horizon, finally, the rating-based strategy uniformly leads to lower average and extreme losses than the 0.5% and 0.75% EDF cutoffs. In general, results do not differ much between the total sample period and the 1991-2001 sub-period.

The 2% EDF cutoff leads to loss rates that are typically lower than those for the B1 cut-off. As noted above, the 2% EDF cutoff produces a universe that has roughly the same average EDF as the investment-grade cutoff, justifying a comparison of these two rules. The table shows that loss rates are consistently higher than those for the investment-grade cutoff.

Overall the results suggest that, during the 1983-2003 period, long-term investors who restricted purchases to investment-grade issuers incurred similar, or lower, losses than if they had conditioned their purchases on EDFs; EDFs, on the other hand, appear to produce lower losses if the horizon is short and average risk is relatively high. Note that this is consistent with the patterns observed in section 2 in the analysis of default prediction accuracy.

The caveat is that the analysis so far does not model the relation between prices and risk. It is not evident, however, that an analysis of expected returns would change the entire picture in favor of EDFs. As shown in Table 1, the mean EDF under the 0.5% or 0.75% cutoff is distinctly smaller than the mean EDF of investment-grade issuers, while the 2% cutoff has similar, but slightly larger average EDF. If, as the literature suggests (Kealhofer 2003b), market prices are more closely associated with EDFs, more closer than with ratings, it should be cheaper to set up investment-grade portfolios than portfolios with EDF below 0.75%.

Tests of statistical significance are complicated by the overlapping investment horizons of the simulated portfolios. As a basis for inference, I take means over simulated losses of portfolios whose formation dates lie in a given calendar year. For each strategy, this produces a time series with length varying from 15 (5-year horizon) to 19 (one-year horizon). The question of interest is whether performance differences between rating-based and EDF-based strategies are significant, so I proceed by analyzing differences between pairs of EDF-strategies and the investment-grade strategy. To calculate their standard deviation, I use a Newey and West (1989) estimate of the standard deviation with lag length set equal to the holding period.[11] In some situations, Newey-West estimates have been found to be biased in small samples (Andrews 1991), which is why I also use a non-parametric bootstrap procedure. To capture serial correlation, I first estimate an autoregressive process (AR) of order one, separately for each time series of differences in average losses.[12] I then use a recursive bootstrap to generate repeated samples (see Li und Maddala 1996). To create one bootstrap sample, I randomly draw an observation from the underlying time series which serves as the first observation in the sample. The remaining observations are recursively generated by applying the estimated coefficients of the AR process and sampling with replacement from the residuals of the AR analysis. For each pair of strategies, I simulate 10,000 bootstrap samples. Table 3 reports the results. The t-statistics are based on Newey-West;[13] the bootstrap analysis is summarized by reporting the frequency with which the average bootstrapped difference between rating-based and EDF based strategies was positive.

The Newey-West analysis and the bootstrap are consistent in that a positive (negative) t-statistic is coupled with a high (low) probability of observing positive bootstrapped results.

---

[10] The bases for this test are the 996 simulated average loss rates for each strategy. Note that the test examines *simulation* error, which is not the same as the standard error of an estimate that is discussed below. The simulation error can be brought down to zero by increasing the number of trials to infinity.

[11] This, rather than a lag of holding period minus one, is necessary because the use of within-year formation dates extends the overlap. With a 1-year horizon, for example, a portfolio formed in July 1990 will overlap with a portfolio formed in January 1991.

[12] The median t-statistic of the estimated autocorrelation coefficients is 3.4. Adding a second lag does not improve the fit.

[13] On average, the Newey-West standard deviations computed for Table 6 exceed ordinary standard deviations by a factor of 1.23.

The two procedures assign statistically significant statistics to the differences between the investment-grade cutoff and the 2% EDF cutoff, and insignificant ones for the other two pairs of strategies. The small sample size as well as the extent of overlap indicate that great caution is needed in interpreting the results. The size of the significant statistics do not suggest that the observed performance differences between the 2% EDF strategy and the investment grade strategy are completely driven by chance. Put differently, the evidence does not suggest that one should expect to see entirely different patterns if the "experiment" of 1983 to 2002 were to be repeated.

Strategies with purchase and sell restrictions

In an extension of the previous buy-and-hold portfolios, issuers are now randomly replaced if they no longer meet the criterion that defines the universe. If trading is assumed to be costly, transaction costs are charged for such replacements. Strategies differ in the speed with which replacements occur. Results are shown in Table 4.

In the absence of transaction costs, the performance differences between rating-based and EDF-based strategies can be summarized as follows: The investment-grade rules lead to lower loss rates if they require selling bonds immediately upon a downgrade; if selling is required only after 6 or 12 months, EDF-based rules produce lower loss rates. In the light of the criticism that rating agencies face, the former result may seem surprising. In the data set, however, there is only one defaulting issuer that was rated investment grade on end of the month immediately prior to its default.[14] Differences between the 1991-2001 sub-period and the entire sample are minor; this also holds for the ensuing analyses.

The picture is less clear when 1% round-trip transaction costs are charged for transactions that are triggered by the portfolio rules. Following rating-based strategies typically leads to fewer transactions, so relative performance is strengthened whenever loss rates were lower before transaction costs; this is the case for rules that require the selling of lower-rated bonds immediately. The impact of transaction costs can be so large that it reverses previous results. Simple EDF rules which set a maximum EDF of 0.5% or 0.75% now produce higher losses—except for the variant which rules out any transactions (one-year horizon with selling required after 12 months). The refined EDF rule remains superior in about 60% of those cases where it

[14] Of course, there are issuers who did have an investment-grade rating shortly before default, but were not rated as investment grade at the preceding month-end. Enron, for example, defaulted on 12/02/01 after having been downgraded to B2 on 11/30/01.

led to lower losses without trading costs. Differences are relatively small. For example, the loss over three years with a six-month tolerance window is 0.206% under the rating-based rule, and 0.211% under the refined EDF-rule. The 2% EDF cutoff, finally, mostly leads to larger losses than the investment-grade rule, even if transaction costs are zero. On the other hand, losses are usually lower than with the B1 cutoff.

Without doubt, one could devise EDF rules that perform better in the presence of transaction costs. I do not embark on such a search as data mining concerns would make it difficult to judge the results. The general pattern of results appears to be sufficiently founded: In the 1983-2001 period, rating-based rules lead to similar, and possibly lower losses than *some* plausible EDF-based rules, but there are EDF rules whose performance is similar, and possibly superior to rating-based rules.

Table 5 shows the Newey-West t-statistics for the test of equal losses for both EDF strategy and the investment-grade strategy;. I do not undertake a bootstrap analysis here because it led to similar results in the previous section. The t-values are mostly between 1 and 2.5, with 1% trading costs, but the cases where the investment-grade strategy leads to lower losses are often associated with t-values above 4. Again, the statistics should be interpreted with care, but they do not seem to suggest that confidence intervals are so large as to make the general conclusion meaningless. The results are also relatively robust to the assumptions about trading costs. Table 6 shows that in those cases where rating-based rules are superior transaction costs could be lowered without reverting signs. Critical transaction costs are in the range 0.25% to 0.87%. It thus appears that the general conclusion can be upheld if transaction costs were around 0.5% instead of 1%.

Investors and fund managers will not only be concerned about expected loss rates, but also about the distribution of unexpected losses. Table 7 shows extreme loss quantiles. Again, results are mixed. In many comparisons, investment-grade strategies have lower quantiles, i.e. exhibit lower risk, regardless of whether transaction costs are charged or not. Due to the nature of the quantile calculation, transaction costs can have no effect at all. In most cases, the differences seem economically small. Even for the extreme 99.9% quantile, the maximum advantage of EDF-based strategies is 1 percentage point. It thus seems that the stability of ratings, which reduces transaction costs under a governance rule, does not lead to significant increases in portfolio risk. This is not a formal risk-return analysis, of course, but such an analysis would necessarily be incomplete as long as bond prices are not made dependent on default risk. Contrary to the buy-and-hold policies, it is not enough to argue that some portfolios are cheaper to set up than others, because prices of bonds that have to be sold do also matter.

3.3 Incorporating bond prices into performance analysis

In this section, the analysis is extended by taking differences in bond prices into account. Since I cannot associate the issuers in the data set with actual bond prices, I take the following two routes to determine prices at which bonds in simulated portfolios are purchased and sold. Either I use a structural pricing model to derive risky bond prices from EDFs, or bond prices are set according to the empirical relationship between EDFs and bonds. Alternatively, one could derive bond prices from ratings, but this approach would be inconsistent with empirical findings. Empirically, equity volatility, which is an important driver of EDFs, explains as much of the cross-section of EDFs as do ratings (Campbell and Taksler, 2002).

The analysis in this section is only performed for the one-year horizon because the EDFs contained in the data set are estimates of the one-year default probability. The transformation to *n*-year default probabilities is part of the commercial product offered by MKMV, but it requires additional input not available in the data set.

The simplest way of pricing a risky bond is to assume that it pays off at only one date, and that the price in the event of default is fixed. Today's price then obtains by discounting the expected pay-off under the risk-neutral default probability (see Crouhy, Galai and Mark, 2000). In order to use EDFs, which are estimates of actual default probabilities, they have to be converted into risk-neutral probabilities. This can be done through a structural model of default. Assume that firm i defaults as soon as the value of its assets A_i falls below a critical threshold L_i. If the asset value follows geometric Brownian motion with drift μ and volatility σ, the default probability from t_0 to t obtains as (see e.g. Ingersoll, 1987):

$$\text{Prob(Default)} = \Phi\left(\frac{\ln(A_i / L_i) - \mu(t - t_0)}{\sigma\,(t - t_0)^{1/2}}\right) + \exp\left(\frac{2\mu \ln(A_i / L_i)}{\sigma^2}\right)\Phi\left(\frac{\ln(A_i / L_i) + \mu(t - t_0)}{\sigma\,(t - t_0)^{1/2}}\right) \qquad (1)$$

For each observation i and a given μ and σ I determine $ln(A_i/L_i)$ such that expression (1) amounts to EDF_i for $t - t_0 = 1$. Inserting this $ln(A_i/L_i)$ into (1) and replacing the drift rate through the risk-free interest produces the risk-neutral default probability $EDF^{RN}_{i,t-t_0}$. Assuming that the recovery is 50% of the face value leads to the following price at time t_0:

$$\text{Price}_{i,t-t_0} = \frac{1 - 0.5 \times EDF^{RN}_{i,t-t_0}}{(1+i)^{t-t_0}} \tag{2}$$

Throughout the analysis, I set the continuous-time risk-free interest to 5%, and the expected logarithmic asset return to 10%. Asset volatility is either set to 10%, which makes the Sharpe ratio of the asset return equal 0.5.[15]

Since empirical studies conclude that structural models provide an insufficient description of observed bond prices (see Huang and Huang, 2003), I also derive prices from an empirical study by Bohn (2000). Using a sample of 1261 observations, Bohn fits observed bond spreads to EDFs. I take his parameter estimates for bonds with a maturity of less than 2 years (Exhibit 7 in Bohn, 2000)[16], which allows one to associate each observation *i* in my data set with a spread $s_{t-t_0}(EDF_i)$. The price then obtains as $(\exp(0.05) + s_{t-t_0}(EDF_i))^{-(t-t_0)}$. Figure 4 compares the model spreads derived from the structural model with the ones based on Bohn. As should be expected from previous studies (e.g. Huang and Huang, 2003), model spreads for low-risk bonds are lower than fitted empirical ones.

To assess the performance of simulated portfolios, I compute the average portfolio return, Sharpe ratios (excess simple return over the simple risk-free rate divided by the standard deviation of simple portfolio returns), and a modification of the Sharpe ratio in which the standard deviation is replaced by the square root of the semi-variance. The latter measures the variability of returns below the mean return; its use is motivated by the asymmetry of bond portfolio returns. Mean returns and Sharpe ratios are computed separately for each trial of the simulation, which comprises up to 19 non-overlapping annual returns.

Table 8 reports performance figures, averaged across trials, for the structural model with the asset return Sharpe ratio set to 50%; Table 9 does the same for prices derived from the empirical study of Bohn (2000). The assumed positive asset return Sharpe ratio implies that bonds command a positive risk premium: the higher the default risk, the higher the expected return on a bond. This makes returns on rating-based strategies higher than the ones with a cutoff of 0.75% or below because the rating-based universe has larger mean EDFs (cf. Table

[15] Results from choosing a Sharpe ratio of 25% are qualitatively similar and are therefore not reported.

[16] There are two parameters which are either bond specific, or for which Bohn uses different values. The correlation between individual asset returns and the market return is here set to 0.5; the time-scaling parameter (γ) is set to 0.5, which is the correct value if the structural model used by Bohn is correct. Note that the choice of γ is irrelevant for a maturity of T=1.

1). Portfolio Sharpe ratios are relatively high—higher than the ones one would observe on average if the assumptions of the model pricing model were correct. To compute expected Sharpe ratios for a strategy with no portfolio adjustments (=sell after 12 months), one needs the expected return:

$$\textit{Expected Return} = (1 - 0.5 \times EDF) / \textit{Price} - 1. \qquad (3)$$

To obtain the standard deviation I assume that the asset correlation across issuers is uniform. The standard deviation can then be obtained analytically as shown in appendix B. Resulting Sharpe ratios are also given in the appendix. They are generally below 1.5. The reason for the high Sharpe ratios is the fact that default rates in the data set are considerably lower than corresponding EDFs.[17] Based on the expected Sharpe ratio, one can also assess to what extent the average portfolio EDF affects performance. Rating-based strategies have high mean EDFs, and model spreads are such that this leads to higher expected Sharpe ratios. However, the differences in expected Sharpe ratios that can be seen in appendix B are fairly small compared to the ones in Tables 8 and 9. When comparing simulated Sharpe ratios with the ones that can be achieved in practice, one should also bear in mind that the analysis here abstracts from interest rate and spread risk.

I first discuss the relative performance of EDF strategies with a cutoff of 0.75% or less. In the absence of transaction cost, most of the EDF-based strategies yield higher returns per unit of portfolio risk. Adding transaction costs (0.5% or 1% round-trip) has a large impact on the relative performance if bonds have to be sold immediately after leaving the universe. With a six-month tolerance period, transaction costs do not lead to a reversal of signs. The results are thus less favorable for ratings than the ones from the previous section. However, ratings lead to better net performance than simple EDF rules if immediate adjustment is required. In addition, the advantage often seems small in economic terms. If selling is required after 6 months, trading costs are 1%, and prices are from the structural model, the 0.5% EDF rule leads to a Sharpe ratio of 1.78; the Sharpe ratio of the rating-based strategy is 1.55. If the rating-based strategy is combined with a risk-less investment so that the standard deviation is equal to the 0.5% EDF rule, the mean return is 5.28%. This compares to 5.30% achieved under the EDF rule.

---

[17] Using a simple average over default rates of December cohorts, issuers with EDF $\leq$ 0.5% have an average default rate of 0.065%. The arithmetic average of over arithmetic mean cohort EDFs is 0.19%.

The EDF rule with 2% cutoff produces mean returns that are very close to the ones under the investment-grade cutoff. Whether EDF- or rating-based governance provides better performance depends critically on the method used for deriving spreads, the sample, and the performance metric. For the 1991-2001 sub-period and standard Sharpe ratios, the EDF rule mostly produces better results, while the investment-grade strategy often leads to better Sharpe ratios based on the semi-variance. Strategies with a restriction to sell after 12 months are effectively buy-and-hold strategies, and can thus be compared to the one-year results from Table 4. There, the losses of the investment-grade rule were smaller than under the EDF rule. In Tables 8 and 9, however, the expected return of both strategies is more or less the same, meaning that the difference in expected losses is compensated by differences in purchase prices. (Recall that the 2% cut-off has slightly higher average EDF, leading to lower purchase prices.) This illustrates the limitation of an analysis that is restricted to losses from default.

A look at extreme return realizations (Table 10) again shows that following rating-based governance rules does not lead to larger risks than EDF-based rules. On the contrary, 0.1% and 1% return quantiles of EDF-based strategies are always lower, meaning that the risk of experiencing low returns is higher when using EDFs. This also holds before transaction costs, and regardless of the way in which prices are derived.

I do not report t-statistics here. Most of the differences in mean returns are associated with high t-statistics, but this is not the question of interest as it follows from the assumptions that higher risk leads to higher expected return. Differences in Sharpe ratios, on the other hand, cannot be tested with the methodology from section 3.2. Since the ratios are computed across years for individual trials, one cannot use within-year averages as the basis for statistical inference.

4 Concluding remarks

I have examined whether ratings or market-based measures of default risk are more suitable for formulating governance rules. Such rules, which consist of buy and sell restrictions, are commonly used in investment management. Market-based credit risk measures are a natural benchmark to measure ratings against because they are often advocated as an alternative to ratings, and because they are now readily available to investors. Using a large matched sample of monthly observations of Moody’s credit ratings and MKMV’s EDFs covering the years 1983 to 2002, I simulate portfolio performance under various governance rules. Based on the results, one cannot conclude whether ratings or market-based measures unambiguously lead to better results in terms of risk, average losses, or risk-return measures. Depending on the circumstances and the metrics used to judge performance, the relative performance is either similar, or favors one of the two types of credit risk measure.

Ratings are widely used by investors and regulators, but many studies criticize the accuracy and timeliness of ratings. This apparent contradiction can be resolved. First, prior studies do not quantify the economic consequences of informational inefficiencies for a specific usage. If the inefficiencies were as severe as sometimes suggested, they should impair the effectiveness of rating-based governance rules. Portfolios that are formed in accordance with such rules should exhibit large variation in risk, and the risk of extreme losses should be high. This is not the case. Second, the picture is consistent with the agencies' stated rating policies. While many studies examine the short-term performance of ratings, rating agencies state that some information relevant for short-term default risk may not fully or immediately be incorporated into ratings because they trade off timeliness and accuracy against rating stability (Cantor, 2001, and Fons, 2002). More or less, what we observe is what we should expect to see if rating agencies find a good balance between these conflicting rating objectives.

One possible criticism of the study is that the results, being dependent on relatively few defaults, are driven by chance. There are two counterarguments. First, the statistical tests carried out in the paper do not indicate that confidence intervals are so large as to invalidate the general conclusion. Second, related studies used data sets that are comparable, and sometimes even smaller in size. Note that one could not easily resolve the inconsistency between this study and others by arguing that the power of detecting the weaknesses of ratings is simply too low in the context of this study. Even if one were sure that the weaknesses were there and power were low, the very fact that the power is low could mean that it would not make a great difference whether investors used ratings or another measure for defining governance rules.

Another possible criticism is that the study does not take into account the full information content of rating system because it does not account for rating outlook and Watchlist designation. As recently shown by Hamilton and Cantor (2004), the accuracy of ratings is dramatically improved when ratings are conditioned on these aspects of the rating opinion. However, at present, very few rating-based governance systems condition on these factors.

Another limitation of the study is that it does not use actual bond prices but hypothetical prices derived from EDFs. While this has advantages of its own—using actual prices would reduce the sample size, and prices are possibly affected by the governance rules under scrutiny—it is desirable to complement the study with a performance analysis based on observed prices. Contrary to what is assumed in the paper, EDF-based measures of default risk might lead actual bond prices, increasing their value as a governance tool. As such, this could bias the results in favor of ratings, but it would not necessarily invalidate the conclusion that ratings were a reasonable choice for governance rules. During the great part of the sample period, market-based measures were observed only by a small minority of bond market participants. If the majority of investors had used market-based measures instead of ratings, one should not expect the former to have lead prices to an economically significant extent.

A different question is whether the results carry forward into the future. As shown by Chen and Wei (2001), trading costs for corporate bonds have gone down dramatically since the mid 1990s. The lower transaction costs are, the less important is the stability of a rating system if ratings are used in portfolio governance. Hence, the recent move of rating agencies towards rating somewhat more aggressively, which could appear to be caused by outside pressure, can also be seen as a natural response to changes on bond markets.

Though the focus of the paper was on comparing the performance of ratings and market-based credit risk measures, it should be clear that they need not be seen as substitutes. The fact that both are conceptually different suggests that investors might benefit from using both credit risk measures as complements.[18]

[18] To obtain an indication for the potential value of using both measures as complements I examined a buy-and-hold strategy that restricts purchases to issuers with investment-grade rating *and* EDF smaller than 2%. Regardless of the holding period, the strategy yields smaller losses than each of the strategies analyzed in Table 3. The complementary nature of EDFs and Ratings is also pointed out by Rappoport and Sirinathsingh (2002).

References

Altman, E., Kao, D. L., 1992, Rating drift in high-yield bonds. Journal of Fixed Income, 15-20.

Andrews, D., 1991, Heteroskedasticity and autocorrelation consistent covariance matrix estimation. Econometrica 59, 817-858.

Bohn, J.R., 2000, An empirical assessment of a simple contingent-claims model for the valuation of risky debt. Journal of Risk Finance 1, 55-77.

Campbell, J.Y, Taksler, G.B, 2002, Equity volatility and corporate bond yields. Working paper.

Cantor, R., Packer, F., 1997, Differences of opinion and selection bias in the credit rating industry. Journal of Banking and Finance 21, 1395-1417.

Cantor, R., 2001, Moody’s investors service response to the consultative paper issued by the Basel Committee on Banking Supervision and its implications for the rating agency industry. Journal of Banking and Finance 25, 171-186.

Cantor, R., Mann, C., 2003, Measuring the performance of corporate bond ratings. Special comment, Moody's Investors Service.

Chen, L., Wei, J., 2001, An indirect estimate of transaction costs for corporate bonds. Working paper, University of Michigan/Toronto.

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

Delianedis, G., Geske, R., 1999, Credit risk and risk neutral default probabilities: information about rating migrations and defaults. Working paper, UCLA.

Fons, J., 2002, Understanding Moody’s corporate bond ratings and rating process. Special Comment, Moody’s Investors Service.

Galil, K., 2002, The quality of corporate credit rating: an empirical investigation. Working paper, Tel-Aviv University.

Hamilton, D.T., Cantor, R., 2004, Rating transitions and default rates conditional on outlooks, Watchlists, and rating history. Special Comment, Moody's Investors Service.

Hamilton, D.T., Varma, P., 2003, Default & recovery rates of corporate bond issuers. Special Comment, Moody's Investors Service.

Huang, J., Huang, M., 2003, How much of the corporate-treasury yield spread is due to credit risk. Working paper.

Hull, J., Predescu, M., White, A., 2004, The relationship between credit default swap spreads, bond yields, and credit rating announcements. Journal of Banking and Finance, this issue.

Ingersoll, J., 1987, Theory of financial decision making. Savage, Md..

Kealhofer, S., 2003a, Quantifying Credit Risk I: Default Prediction. Financial Analysts Journal 59 (1), 30-44.

Kealhofer, S., 2003b, Quantifying Credit Risk II: Debt Valuation. Financial Analysts Journal 59 (3).

Lando, D., Skødeberg, T., 2002, Analyzing rating transitions and rating drift with continuous observations. Journal of Banking and Finance 26, 423-444.

Li, H., Maddala, G.S., 1996, Bootstrapping time series models. Econometric reviews 15, 115-158.

Löffler, G., 2004, An anatomy of rating through the cycle. Journal of Banking and Finance 28, 695-720.

Löffler, G., 2003, Avoiding the rating bounce: Why rating agencies are slow to react to new information. Working paper, University of Frankfurt.

Newey, W., West, K., 1987, A simple positive semi-definite, heteroscedasticity and autocorrelation consistent covariance matrix. Econometrica 55, 703-708.

Norden, L., Weber, M., 2004, Informational efficiency of credit default swap and stock markets: The impact of credit rating announcements. Journal of Banking and Finance, this issue.

Rappoport, P. Sirinathsingh, M., 2002. Using equity prices to trade corporate bonds: An evaluation. Broker report, JPMorgan.

Schultz, P., 2001, Corporate bond trading costs: a peek behind the curtain. Journal of Finance 56, 677-698.

Sobehart, J.R., Keenan, S.C., Stein, R.M, 2000, Benchmarking quantitative default risk models: a validation methodology. Moody’s Investors Service.

Stein, R.M., Jordão, F., 2003, What is a more powerful model worth? Technical Report, Moody's KMV.

Vasicek, O.A., 2001, EDF credit measure and corporate bond pricing. Working paper, KMV.

Table 1: Descriptive statistics for the entire sample and universes defined on ratings and EDFs

	Obser-vations	EDF (%)			Rating (1=Aaa, 21=C)			Defaults within *x* years after belonging to universe		
		Mean	Median	σ	Mean	Median	σ	1 year	3 years	5 years
Total sample	297,645	2.08	0.39	4.57	9.53	9.00	4.12	-	-	-
Rating<Ba1	175,318	0.41	0.17	1.00	6.63	7.00	2.33	22	44	47
EDF≤0.5%	166,224	0.17	0.14	0.14	7.30	7.00	3.09	25	132	184
EDF≤0.75%	190,839	0.23	0.17	0.20	7.62	7.00	3.22	39	175	232
EDF≤2%	240,433	0.44	0.25	0.47	8.40	8.00	3.51	114	333	363
Rating<B1	239,857	0.76	0.26	1.77	8.11	8.00	3.16	130	216	216

Table 2: Loss rates of buy-and-hold portfolios

At month-end, 50 issuers are randomly selected from the admissible universe. The investment horizon varies from one to five years, and portfolios are equally weighted. Issuers are randomly replaced if they default or leave the database; there is no portfolio revision if an issuer leaves the universe. Loss rates result from default (assuming a recovery rate of 50%), and transaction costs. Gray shaded cells mark entries where EDF-based strategies lead to lower loss rates than the investment-grade strategy. Figures are based on 996 simulation trials; each trial comprises up to 19 portfolios formed between 1983 to 2001. Note that loss rates are not annualized.

Holding Period	Sample	Average loss	99.9% quantile of losses	99% quantile of losses
Universe: Investment grade				
1year	1983-2001	0.052%	2.00%	1.00%
3years	1983-2001	0.297%	4.00%	2.00%
5years	1983-2001	0.609%	5.00%	4.00%
1year	1991-2001	0.042%	2.00%	1.00%
3years	1991-2001	0.185%	3.00%	2.00%
5years	1991-2001	0.312%	4.00%	3.00%
Universe: EDF<=0.50%				
1year	1983-2001	0.025%	1.00%	1.00%
3years	1983-2001	0.296%	4.00%	2.00%
5years	1983-2001	0.680%	6.00%	4.00%
1year	1991-2001	0.021%	1.00%	1.00%
3years	1991-2001	0.356%	4.00%	3.00%
5years	1991-2001	0.706%	6.00%	4.00%
Universe: EDF<=0.75%				
1year	1983-2001	0.038%	2.00%	1.00%
3years	1983-2001	0.333%	4.00%	3.00%
5years	1983-2001	0.781%	7.00%	4.00%
1year	1991-2001	0.032%	2.00%	1.00%
3years	1991-2001	0.384%	4.00%	3.00%
5years	1991-2001	0.803%	7.00%	5.00%
Universe: EDF<=2%				
1year	1983-2001	0.128%	2.00%	2.00%
3years	1983-2001	0.840%	6.00%	4.00%
5years	1983-2001	1.732%	9.00%	7.00%
1year	1991-2001	0.083%	2.00%	1.00%
3years	1991-2001	0.693%	6.00%	4.00%
5years	1991-2001	1.285%	9.00%	6.00%
Universe: Rating better than Ba1				
1year	1983-2001	0.234%	3.00%	2.00%
3years	1983-2001	1.157%	7.00%	5.00%
5years	1983-2001	2.007%	10.00%	7.00%
1year	1991-2001	0.151%	3.00%	2.00%
3years	1991-2001	0.682%	6.00%	4.00%
5years	1991-2001	1.045%	7.00%	5.00%

Table 3: Statistical significance of differences in average loss rates of buy-and-hold portfolios

Based on the analysis of Table 2, I examine differences between average loss rates of the investment-grade strategy and three EDF-based strategies. Positive differences mean that the rating-based strategy lead to larger losses. All analyses are for the entire sample 1981-2001.

Holding period	Average difference of losses	t-statistic (Newey-West)	Fraction positive in bootstrap
Panel A: Investment-grade versus 0.5% EDF cutoff			
1year	0.026%	1.47	86.5%
3years	0.001%	0.01	30.1%
5years	-0.071%	-0.38	9.8%
Panel B: Investment-grade versus 0.75% EDF cutoff			
1year	0.014%	1.05	74.9%
3years	-0.036%	-0.38	18.7%
5years	-0.172%	-0.89	6.2%
Panel C: Investment-grade versus 2% EDF cutoff			
1year	-0.076%	-3.10	0.0%
3years	-0.543%	-5.24	0.0%
5years	-1.123%	-4.75	0.0%

Table 4: Average loss rates of simulated portfolios with portfolio adjustments

The set-up and presentation is similar to Table 2. The main difference is that issuers are randomly replaced if they no longer meet the criterion that defines the universe, and that trading costs are charged for such revisions. Strategies differ in the speed with which replacement occurs in this instance (immediately/after 6 months/after 12 months provided issuers do not re-enter the universe within the specified tolerance period).

Holding Period	Sample	No trading cost			trading cost 1%		
		Sell after 0m	Sell after 6m	Sell after 12m	Sell after 0m	Sell after 6m	Sell after 12m
Universe: Investment grade							
1year	1983-2001	0.005%	0.041%	0.052%	0.029%	0.050%	0.052%
3years	1983-2001	0.020%	0.155%	0.172%	0.094%	0.206%	0.208%
5years	1983-2001	0.044%	0.229%	0.268%	0.159%	0.315%	0.339%
1year	1991-2001	0.000%	0.038%	0.042%	0.022%	0.046%	0.042%
3years	1991-2001	0.000%	0.156%	0.161%	0.064%	0.200%	0.193%
5years	1991-2001	0.000%	0.240%	0.240%	0.099%	0.312%	0.300%
Universe: EDF<=0.5%							
1year	1983-2001	0.018%	0.023%	0.025%	0.280%	0.069%	0.025%
3years	1983-2001	0.067%	0.085%	0.107%	0.667%	0.312%	0.245%
5years	1983-2001	0.112%	0.130%	0.148%	0.950%	0.493%	0.385%
1year	1991-2001	0.007%	0.016%	0.021%	0.240%	0.059%	0.021%
3years	1991-2001	0.022%	0.041%	0.091%	0.579%	0.265%	0.228%
5years	1991-2001	0.036%	0.073%	0.126%	0.829%	0.431%	0.357%
Universe: EDF<=0.75%							
1year	1983-2001	0.016%	0.030%	0.038%	0.219%	0.067%	0.038%
3years	1983-2001	0.063%	0.100%	0.142%	0.515%	0.278%	0.248%
5years	1983-2001	0.101%	0.155%	0.213%	0.725%	0.435%	0.394%
1year	1991-2001	0.008%	0.018%	0.032%	0.193%	0.053%	0.032%
3years	1991-2001	0.023%	0.046%	0.132%	0.456%	0.225%	0.238%
5years	1991-2001	0.036%	0.066%	0.202%	0.649%	0.349%	0.380%
Universe: Buy EDF<=0.50%, Sell EDF>0.75%							
1year	1983-2001	0.018%	0.022%	0.025%	0.135%	0.037%	0.025%
3years	1983-2001	0.065%	0.084%	0.131%	0.388%	0.211%	0.207%
5years	1983-2001	0.106%	0.146%	0.201%	0.575%	0.361%	0.340%
1year	1991-2001	0.009%	0.015%	0.021%	0.117%	0.030%	0.021%
3years	1991-2001	0.024%	0.041%	0.132%	0.344%	0.176%	0.211%
5years	1991-2001	0.038%	0.069%	0.195%	0.518%	0.294%	0.335%
Universe: EDF<=2%							
1year	1983-2001	0.027%	0.102%	0.128%	0.136%	0.120%	0.128%
3years	1983-2001	0.075%	0.278%	0.523%	0.317%	0.380%	0.582%
5years	1983-2001	0.112%	0.381%	0.836%	0.439%	0.536%	0.935%
1year	1991-2001	0.022%	0.050%	0.083%	0.131%	0.070%	0.083%
3years	1991-2001	0.050%	0.141%	0.340%	0.296%	0.245%	0.401%
5years	1991-2001	0.064%	0.199%	0.512%	0.404%	0.356%	0.613%
Universe: Rating better than B1							
1year	1983-2001	0.080%	0.210%	0.234%	0.110%	0.220%	0.234%
3years	1983-2001	0.279%	0.707%	0.938%	0.368%	0.765%	0.978%
5years	1983-2001	0.456%	1.086%	1.466%	0.590%	1.183%	1.540%
1year	1991-2001	0.025%	0.142%	0.151%	0.050%	0.152%	0.151%
3years	1991-2001	0.074%	0.445%	0.520%	0.144%	0.492%	0.554%
5years	1991-2001	0.119%	0.600%	0.687%	0.218%	0.669%	0.744%

Table 5: Newey-West t-values for differences in average loss rates (1983-2001)

Based on the analysis of Table 4, I test differences between average loss rates of the investment-grade strategy and four EDF-based strategies. Positive differences mean that the rating-based strategy leads to larger losses.

	No trading cost			trading cost 1%		
Holding Period	Sell after 0m	Sell after 6m	Sell after 12m	Sell after 0m	Sell after 6m	Sell after 12m
Universe: EDF<=0.5%						
1year	-1.88	1.03	1.47	-7.01	-1.17	1.47
3years	-2.10	1.30	2.48	-5.31	-3.20	-1.31
5years	-1.68	1.65	2.37	-4.90	-3.38	-1.16
Universe: EDF<=0.75%						
1year	-1.55	0.62	1.05	-6.54	-1.03	1.05
3years	-2.03	0.99	2.36	-5.08	-1.86	-1.37
5years	-1.67	1.06	2.16	-4.85	-2.15	-1.53
Universe: Buy EDF<=0.50%, Sell EDF>0.75%						
1year	-1.76	1.07	1.50	-4.89	0.78	1.50
3years	-1.85	1.28	2.25	-4.22	-0.16	0.04
5years	-1.53	1.26	2.15	-3.99	-1.00	-0.03
Universe: EDF<=2%						
1year	-1.99	-2.27	-3.10	-5.16	-2.66	-3.10
3years	-2.08	-1.40	-4.96	-4.08	-2.03	-3.89
5years	-2.06	-1.23	-5.13	-4.10	-1.78	-3.10

Table 6: Transaction costs that equate average loss rates under EDF rules with the investment-grade rule

Based on the analysis used for Table 4, transaction costs are chosen to make loss rates of EDF-based strategies equal to loss rates of corresponding rating-based strategies with investment-grade cutoff. Cells with no entry mark cases in which transaction cost would have to be negative in order to equate loss rates, or in which there is no trading.

Holding Period	Sample	Sell after 0m	Sell after 6m	Sell after 12m
Universe: EDF<=0.50%				
1year	1983-2001		0.48%	
3years	1983-2001		0.40%	0.64%
5years	1983-2001		0.36%	0.72%
1year	1991-2001		0.62%	
3years	1991-2001		0.64%	0.67%
5years	1991-2001		0.59%	0.67%
Universe: EDF<=0.75%				
1year	1983-2001		0.38%	
3years	1983-2001		0.43%	0.43%
5years	1983-2001		0.38%	0.50%
1year	1991-2001		0.72%	
3years	1991-2001		0.82%	0.40%
5years	1991-2001		0.82%	0.32%
Universe: Buy EDF<=0.50%, Sell EDF>0.75%				
1year	1983-2001		3.34%	
3years	1983-2001		0.93%	1.03%
5years	1983-2001		0.64%	0.98%
1year	1991-2001		3.57%	
3years	1991-2001		1.26%	0.62%
5years	1991-2001		1.12%	0.56%
Universe: EDF<=2%				
1year	1983-2001			
3years	1983-2001			
5years	1983-2001			
1year	1991-2001			
3years	1991-2001		0.25%	
5years	1991-2001		0.48%	

Table 7: Loss quantiles for simulated portfolios (1983-2001)

The Table shows 99.9% and 99% loss quantiles for the simulated strategies already analyzed in Tables 4 to 6. A 99.9% quantile of 2.00% means that in 99.9% of all trials losses are smaller than 2.00%. Gray shaded cells mark entries where EDF-based strategies lead to lower quantiles (=lower risk) than investment-grade strategies.

		no trading cost			1% trading cost		
Holding Period	Quantile	Sell after 0m	Sell after 6m	Sell after 12m	Sell after 0m	Sell after 6m	Sell after 12m
Universe: Investment grade							
1year	99.9%	1.00%	2.00%	2.00%	1.04%	2.00%	2.00%
3years	99.9%	1.00%	3.00%	3.00%	1.16%	3.06%	3.06%
5years	99.9%	2.00%	3.00%	3.00%	2.12%	3.16%	3.14%
1year	99%	0.00%	1.00%	1.00%	0.12%	1.02%	1.00%
3years	99%	1.00%	2.00%	2.00%	1.06%	2.06%	2.04%
5years	99%	1.00%	2.00%	2.00%	1.16%	2.14%	2.14%
Universe: EDF<=0.50%							
1year	99.9%	1.00%	1.00%	1.00%	1.72%	1.12%	1.00%
3years	99.9%	2.00%	2.00%	2.00%	2.88%	2.46%	2.38%
5years	99.9%	3.00%	3.00%	3.00%	3.94%	3.48%	3.36%
1year	99%	1.00%	1.00%	1.00%	1.22%	1.02%	1.00%
3years	99%	1.00%	1.00%	1.00%	2.00%	1.48%	1.38%
5years	99%	1.00%	2.00%	2.00%	2.60%	2.40%	2.34%
Universe: EDF<=0.75%							
1year	99.9%	1.00%	2.00%	2.00%	1.52%	2.02%	2.00%
3years	99.9%	2.00%	2.00%	2.00%	2.56%	2.38%	2.34%
5years	99.9%	2.00%	2.00%	3.00%	2.98%	2.70%	3.40%
1year	99%	1.00%	1.00%	1.00%	1.20%	1.04%	1.00%
3years	99%	1.00%	1.00%	2.00%	1.72%	1.42%	2.12%
5years	99%	1.00%	2.00%	2.00%	2.12%	2.28%	2.34%
Universe: Buy EDF<=0.50%, Sell EDF>0.75%							
1year	99.9%	1.00%	2.00%	2.00%	1.42%	2.00%	2.00%
3years	99.9%	2.00%	2.00%	2.00%	2.46%	2.26%	2.28%
5years	99.9%	3.00%	3.00%	3.00%	3.36%	3.26%	3.28%
1year	99%	1.00%	1.00%	1.00%	1.14%	1.00%	1.00%
3years	99%	1.00%	1.00%	2.00%	1.58%	1.32%	2.06%
5years	99%	1.00%	2.00%	2.00%	1.98%	2.22%	2.30%
Universe: EDF<=2%							
1year	99.9%	2.00%	2.00%	2.00%	2.18%	2.08%	2.00%
3years	99.9%	2.00%	4.00%	5.00%	2.40%	4.12%	5.04%
5years	99.9%	2.00%	4.00%	6.00%	2.76%	4.28%	6.16%
1year	99%	1.00%	1.00%	2.00%	1.20%	1.10%	2.00%
3years	99%	1.00%	2.00%	3.00%	1.52%	2.28%	3.12%
5years	99%	1.00%	3.00%	4.00%	1.94%	3.20%	4.22%
Universe: Rating better than B1							
1year	99.9%	2.00%	3.00%	3.00%	2.12%	3.02%	3.00%
3years	99.9%	4.00%	5.00%	6.00%	4.12%	5.14%	6.06%
5years	99.9%	4.00%	6.00%	8.00%	4.30%	6.20%	8.12%
1year	99%	1.00%	2.00%	2.00%	1.10%	2.02%	2.00%
3years	99%	2.00%	4.00%	5.00%	2.26%	4.08%	5.04%
5years	99%	3.00%	5.00%	6.00%	3.24%	5.12%	6.12%

Table 8: Performance of simulated portfolios (Prices derived from EDFs and a structural model)

Portfolio construction is the same as for Table 4. Prices are model prices derived from EDFs. The performance figures stated in the table average performance figures across trials; one trial comprises 19 years. Gray shaded cells mark entries where EDF-based strategies lead to better performance than investment-grade strategies.

Performance measure	Sample	no trading cost Sell after 0m	no trading cost Sell after 6m	no trading cost Sell after 12m	0.5% trading cost Sell after 0m	0.5% trading cost Sell after 6m	1% trading cost Sell after 0m	1% trading cost Sell after 6m
Universe: Investment grade								
Mean return	1983-2001	5.65%	5.63%	5.63%	5.64%	5.63%	5.62%	5.62%
Sharpe	1983-2001	1.86	1.58	1.57	1.82	1.56	1.78	1.55
Sharpe (Semi-variance)	1983-2001	3.06	2.45	2.44	3.01	2.44	2.95	2.42
Mean return	1991-2001	5.58%	5.56%	5.56%	5.56%	5.56%	5.55%	5.55%
Sharpe	1991-2001	2.03	1.68	1.66	2.01	1.67	1.99	1.66
Sharpe (Semi-variance)	1991-2001	3.16	2.57	2.55	3.13	2.55	3.09	2.54
Universe: EDF<=0.50%								
Mean return	1983-2001	5.30%	5.35%	5.35%	5.17%	5.32%	5.04%	5.30%
Sharpe	1983-2001	1.97	2.23	2.20	0.41	2.09	-0.40	1.78
Sharpe (Semi-variance)	1983-2001	2.96	3.31	3.26	0.58	3.11	-0.52	2.65
Mean return	1991-2001	5.27%	5.31%	5.32%	5.16%	5.29%	5.04%	5.27%
Sharpe	1991-2001	2.55	2.91	2.88	0.36	2.70	-0.53	2.18
Sharpe (Semi-variance)	1991-2001	3.93	4.31	4.26	0.50	4.09	-0.72	3.30
Universe: EDF<=0.75%								
Mean return	1983-2001	5.39%	5.43%	5.43%	5.29%	5.42%	5.18%	5.40%
Sharpe	1983-2001	2.25	2.34	2.21	1.14	2.25	0.33	2.09
Sharpe (Semi-variance)	1983-2001	3.38	3.40	3.17	1.68	3.28	0.46	3.06
Mean return	1991-2001	5.35%	5.40%	5.39%	5.26%	5.38%	5.16%	5.36%
Sharpe	1991-2001	2.82	3.14	2.83	1.27	3.01	0.26	2.76
Sharpe (Semi-variance)	1991-2001	4.40	4.67	4.15	1.80	4.55	0.36	4.21
Universe: Buy EDF<=0.50%, Sell EDF>0.75%								
Mean return	1983-2001	5.33%	5.35%	5.35%	5.27%	5.34%	5.21%	5.33%
Sharpe	1983-2001	2.10	2.15	2.08	1.25	2.12	0.57	2.02
Sharpe (Semi-variance)	1983-2001	3.12	3.19	3.07	1.80	3.15	0.78	3.00
Mean return	1991-2001	5.30%	5.31%	5.31%	5.24%	5.31%	5.19%	5.30%
Sharpe	1991-2001	2.75	2.80	2.65	1.43	2.78	0.55	2.61
Sharpe (Semi-variance)	1991-2001	4.22	4.16	3.95	2.01	4.17	0.74	3.92
Universe: EDF<=2%								
Mean return	1983-2001	5.65%	5.67%	5.65%	5.59%	5.66%	5.54%	5.65%
Sharpe	1983-2001	2.61	1.87	1.63	2.26	1.84	1.86	1.81
Sharpe (Semi-variance)	1983-2001	3.77	2.43	2.07	3.30	2.39	2.72	2.36
Mean return	1991-2001	5.58%	5.64%	5.63%	5.53%	5.63%	5.47%	5.62%
Sharpe	1991-2001	2.99	2.82	2.30	2.65	2.80	2.16	2.77
Sharpe (Semi-variance)	1991-2001	4.49	4.09	3.26	3.90	4.07	3.11	4.04
Universe: Rating better than B1								
Mean return	1983-2001	5.97%	5.91%	5.90%	5.96%	5.91%	5.94%	5.90%
Sharpe	1983-2001	1.74	1.36	1.30	1.72	1.35	1.69	1.35
Sharpe (Semi-variance)	1983-2001	2.70	2.02	1.93	2.66	2.01	2.62	2.00
Mean return	1991-2001	5.91%	5.86%	5.86%	5.90%	5.85%	5.88%	5.85%
Sharpe	1991-2001	2.07	1.68	1.66	2.05	1.67	2.04	1.66
Sharpe (Semi-variance)	1991-2001	3.37	2.68	2.66	3.35	2.67	3.32	2.66

Table 9: Performance of simulated portfolios (Prices derived from the empirical relationship between EDFs and spreads)

Portfolio construction is the same as for Table 4. Prices are derived from the study of Bohn (2000). Performance figures stated in the table average performance figures across trials; one trial comprises 19 years. Gray shaded cells mark entries where EDF-based strategies lead to better performance than investment-grade strategies.

		no trading cost			0.5% trading cost		1% trading cost	
Performance Measure	Sample	Sell after 0m	Sell after 6m	Sell after 12m	Sell after 0m	Sell after 6m	Sell after 0m	Sell after 6m
Universe: Investment grade								
Mean return	1983-2001	5.80%	5.77%	5.77%	5.78%	5.77%	5.77%	5.76%
Sharpe	1983-2001	4.06	3.04	2.95	4.01	3.02	3.93	3.00
Sharpe (Semi-variance)	1983-2001	6.92	4.72	4.52	6.86	4.70	6.74	4.68
Mean return	1991-2001	5.76%	5.73%	5.73%	5.75%	5.73%	5.74%	5.73%
Sharpe	1991-2001	5.19	3.77	3.65	5.26	3.79	5.27	3.80
Sharpe (Semi-variance)	1991-2001	8.21	5.80	5.60	8.30	5.81	8.30	5.81
Universe: EDF<=0.50%								
Mean return	1983-2001	5.58%	5.61%	5.62%	5.45%	5.59%	5.32%	5.56%
Sharpe	1983-2001	9.39	9.56	9.48	2.92	8.38	1.06	6.07
Sharpe (Semi-variance)	1983-2001	13.97	14.44	14.30	3.88	12.35	1.39	8.31
Mean return	1991-2001	5.58%	5.60%	5.60%	5.46%	5.58%	5.34%	5.55%
Sharpe	1991-2001	14.50	15.28	14.42	3.86	12.20	1.41	8.22
Sharpe (Semi-variance)	1991-2001	20.99	22.95	21.59	5.24	18.12	1.94	11.45
Universe: EDF<=0.75%								
Mean return	1983-2001	5.63%	5.65%	5.65%	5.53%	5.63%	5.42%	5.61%
Sharpe	1983-2001	7.84	7.53	6.76	3.62	7.10	1.75	6.05
Sharpe (Semi-variance)	1983-2001	11.85	11.21	9.90	4.96	10.54	2.31	8.76
Mean return	1991-2001	5.61%	5.64%	5.64%	5.52%	5.62%	5.42%	5.60%
Sharpe	1991-2001	11.65	12.03	10.38	4.65	11.06	2.14	8.95
Sharpe (Semi-variance)	1991-2001	17.39	18.02	15.45	6.29	16.94	2.88	13.40
Universe: Buy EDF<=0.50%, Sell EDF>0.75%								
Mean return	1983-2001	5.60%	5.61%	5.62%	5.54%	5.61%	5.48%	5.60%
Sharpe	1983-2001	9.30	9.49	9.17	4.91	9.19	2.76	8.42
Sharpe (Semi-variance)	1983-2001	13.84	14.35	13.80	6.66	13.70	3.61	12.29
Mean return	1991-2001	5.59%	5.60%	5.60%	5.53%	5.59%	5.48%	5.59%
Sharpe	1991-2001	14.26	15.29	14.22	6.34	14.06	3.43	12.33
Sharpe (Semi-variance)	1991-2001	20.92	22.92	21.39	8.49	21.23	4.57	18.30
Universe: EDF<=2%								
Mean return	1983-2001	5.76%	5.75%	5.73%	5.70%	5.74%	5.65%	5.73%
Sharpe	1983-2001	5.37	2.96	2.40	4.33	2.91	3.28	2.84
Sharpe (Semi-variance)	1983-2001	7.80	3.84	3.00	6.24	3.79	4.57	3.70
Mean return	1991-2001	5.73%	5.76%	5.74%	5.67%	5.75%	5.62%	5.74%
Sharpe	1991-2001	7.35	6.30	4.52	5.97	6.33	4.29	6.25
Sharpe (Semi-variance)	1991-2001	10.84	9.33	6.60	8.49	9.36	5.95	9.24
Universe: Rating better than B1								
Mean return	1983-2001	5.96%	5.87%	5.86%	5.95%	5.87%	5.93%	5.86%
Sharpe	1983-2001	2.36	1.60	1.48	2.33	1.60	2.29	1.59
Sharpe (Semi-variance)	1983-2001	3.46	2.19	2.00	3.42	2.17	3.36	2.16
Mean return	1991-2001	5.95%	5.87%	5.87%	5.94%	5.86%	5.93%	5.86%
Sharpe	1991-2001	3.41	2.26	2.19	3.41	2.25	3.41	2.24
Sharpe (Semi-variance)	1991-2001	5.57	3.37	3.30	5.57	3.36	5.57	3.34

Table 10: Loss quantiles for simulated portfolios (1983-2001)

The Table shows 0.1% and 1% return quantiles for the simulated strategies already analyzed in tables 8 and 9. A 0.1% quantile of 4.61% means that in 0.1% of all simulated returns were below 4.61%. Gray shaded cells mark entries where EDF-based strategies lead to higher return quantiles (=lower risk) than investment-grade strategies.

Table 10.1 Prices derived from EDFs and a structural model

	no trading cost			1% trading cost	
	Sell after 0m	Sell after 6m	Sell after 12m	Sell after 0m	Sell after 6m
Universe: Investment grade					
0.1% quantile	4.61%	4.35%	4.33%	4.59%	4.35%
1% quantile	5.21%	4.61%	4.61%	5.19%	4.60%
Universe: EDF<=0.50%					
0.1% quantile	4.17%	4.19%	4.20%	3.73%	4.15%
1% quantile	4.31%	4.31%	4.31%	3.94%	4.27%
Universe: EDF<=0.75%					
0.1% quantile	4.22%	4.24%	4.23%	3.90%	4.20%
1% quantile	4.41%	4.40%	4.40%	4.12%	4.37%
Universe: Buy EDF<=0.50%, Sell EDF>0.75%					
0.1% quantile	4.20%	4.19%	4.19%	3.97%	4.18%
1% quantile	4.33%	4.30%	4.30%	4.15%	4.29%
Universe: EDF<=2%					
0.1% quantile	3.98%	3.57%	3.56%	3.84%	3.55%
1% quantile	4.63%	4.50%	4.44%	4.43%	4.49%
Universe: Rating better than B1					
0.1% quantile	3.80%	3.07%	2.94%	3.76%	3.07%
1% quantile	4.63%	4.03%	3.94%	4.59%	4.02%

Table 10.2 Prices derived from the empirical relationship between EDFs and spreads (Bohn, 2000)

	no trading cost			1% trading cost	
	Sell after 0m	Sell after 6m	Sell after 12m	Sell after 0m	Sell after 6m
Universe: Investment grade					
0.1% quantile	4.76%	4.63%	4.62%	4.73%	4.61%
1% quantile	5.56%	4.75%	4.76%	5.52%	4.74%
Universe: EDF<=0.50%					
0.1% quantile	4.51%	4.52%	4.53%	3.98%	4.46%
1% quantile	4.58%	4.59%	4.59%	4.30%	4.55%
Universe: EDF<=0.75%					
0.1% quantile	4.53%	4.55%	4.55%	4.13%	4.51%
1% quantile	4.63%	4.63%	4.63%	4.42%	4.60%
Universe: Buy EDF<=0.50%, Sell EDF>0.75%					
0.1% quantile	4.52%	4.53%	4.53%	4.25%	4.51%
1% quantile	4.60%	4.60%	4.60%	4.44%	4.58%
Universe: EDF<=2%					
0.1% quantile	4.16%	3.70%	3.70%	4.03%	3.68%
1% quantile	4.74%	4.68%	4.66%	4.54%	4.66%
Universe: Rating better than B1					
0.1% quantile	3.85%	2.99%	2.92%	3.81%	2.98%
1% quantile	4.76%	3.99%	3.94%	4.72%	3.98%

Figure 1: Number of issuers across time

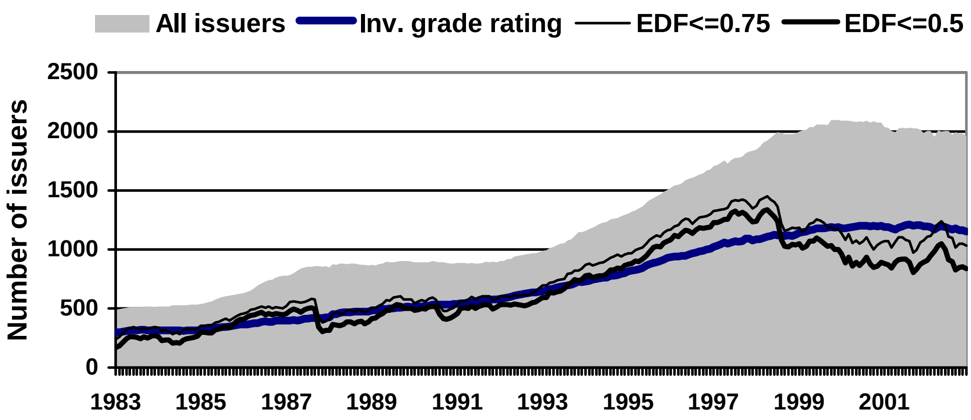


Figure 2: Median EDFs within universes

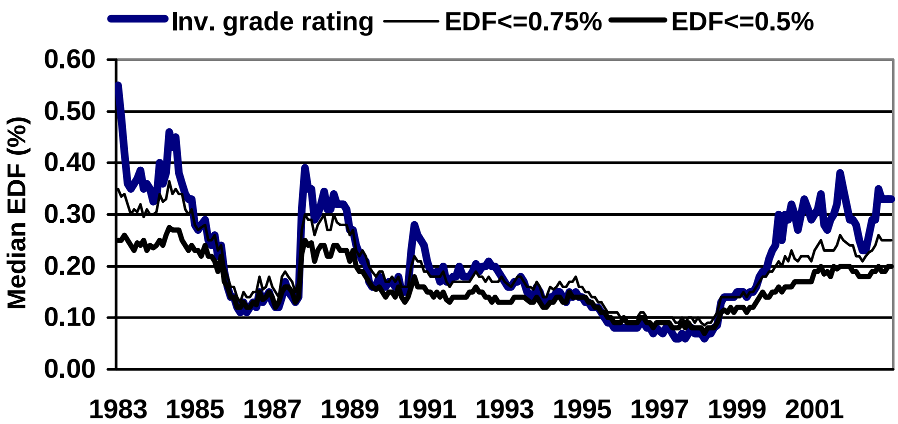


Figure 3: Cumulative accuracy profiles

Cumulative accuracy profiles plot the percentage of defaulters included given a percentage of observations included, where observations are ranked from riskiest (lowest rating or highest EDF) to safest. The fraction of observations with non-investment grade rating is 41%; the fraction of observations with an EDF below 0.75% is 37%.

One-year horizon | *One-year horizon (detail)* | *Five-year horizon*

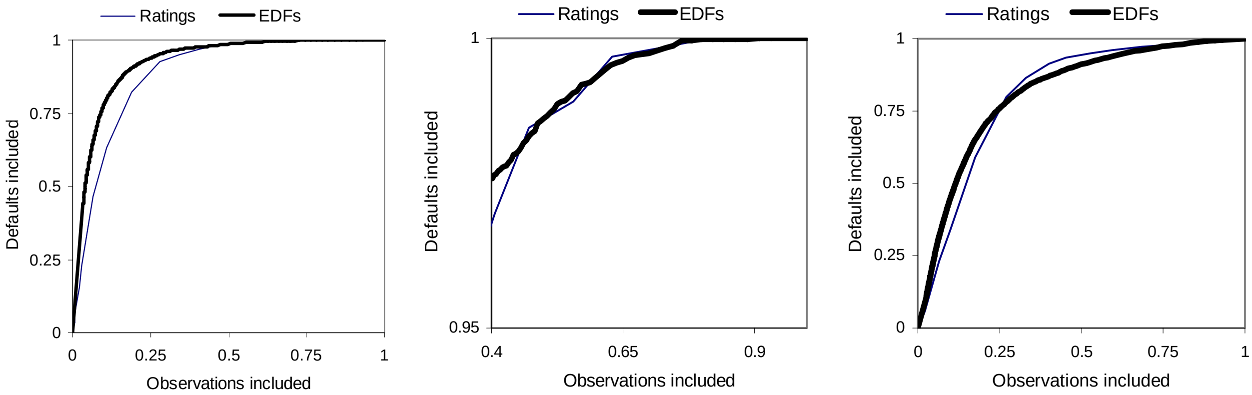
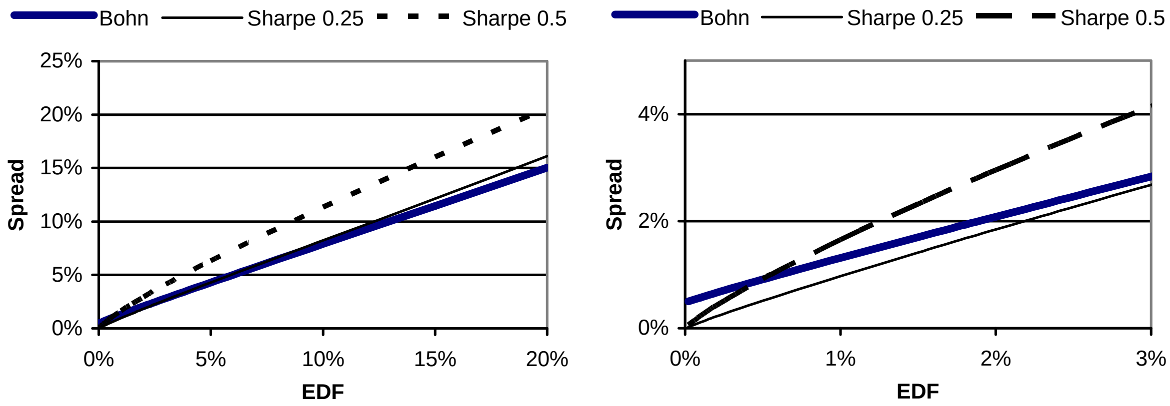


Figure 4: Credit spreads used in the analysis

Spreads are derived from EDFs, either based on the empirical study of Bohn (2000), or from a contingent-claim model with the asset Sharpe ratio set to 0.25 or 0.5. The figure on the right-hand side shows a detail of the left-hand side figure.



Appendix A: Modifications of the data

i) In the data set, the minimum EDF is 0.01% until April 1997 and 0.02% thereafter. Since the minimum EDF is a value set by KMV, I use a consistent minimum of 0.02%.

ii) There are gaps in the data in the sense that data for a specific issuer is available for, say, May and July 2002, but not for June 2002. Most of these gaps appear to be due to data problems. Since gaps can prevent the purchase of bonds and trigger their sale, I linearly interpolate missing EDFs if

- the rating did not change during the gap (to check this, I use the detailed rating history)
- the issuer did not default at any time after the gap
- the gap extends at most over three months
- the percentage change in EDF across the gap is smaller than 25%.

Overall this adjustment increases the number of observations by 608.

iii) There are companies that have EDFs of 20% for an extended period, and then jump to very low EDF levels. In principle, this can happen, but the following observations suggest that these jumps represent data errors:

- The majority of those companies are rated Aaa (224 observations with Rating=Aaa and EDF=20%; 76 observations with rating=Aa1 and EDF=20%; 37 observations with rating=Aa2 and EDF=20%)
- The majority of the jumps occur in a single month, February 1988.
- In most cases, the time series for companies with jumps starts with EDF being at 20%.

Since I cannot determine the true EDF for these instances, I discard questionable observations. Specifically, I discard observations if each of the following conditions is met:

- the EDF is 20%.
- Observations are followed by a change in EDF larger than 18 percentage points.
- the rating is better than A3

The following example shows which observations would be discarded

	Month												
	8702	8703	8704	8705	8706	8707	8708	8709	8710	8711	8712	8801	8802
EDF (%)	-	20	20	20	20	20	20	20	20	20	20	20	0.3
Rating	-	Aaa	Aaa	Aaa	Aaa	Aaa	Aaa	Aaa	Aaa	Aaa	Aaa	Aaa	Aaa
Discard	-	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	no

Overall, 411 observations are discarded.

Appendix B: Portfolio standard deviation and Sharpe ratios

The analysis is done within the structural bond pricing model used in the main part of the paper. Having specified the asset correlation between issuers *i* and *j,* the joint default probability JDP_{ij} follows from the bivariate normal distribution. Define the default correlation:

$$\rho_{ij} = \frac{JDP_{ij} - PD_i PD_j}{\sqrt{PD_i(1-PD_i)PD_j(1-PD_j)}}$$

where PD_i is the default probability of issuer *i*. Let *LGD* denote the loss in case of default. The variance of a portfolio with *N* obligors is:

$$\sigma_N^2 = \sum_{i,j=1}^{N} LGD_i LGD_j \sqrt{PD_i(1-PD_i)PD_j(1-PD_j)}\rho_{ij} .$$

The computation of a portfolio's expected return, which is also needed for the Sharpe ratio, is described in the text.

The following table gives expected Sharpe ratios for equally-weighted portfolios with 50 issuers and a one-year investment horizon with no portfolio adjustments; for a given specification, issuers have identical default probabilities and asset correlations. The assumed asset Sharpe Ratio is 0.5:

	Asset correlation	
Default probability	10%	1%
0.25%	0.63	0.68
0.50%	0.89	1.00
0.75%	1.01	1.18