Heterogeneous rating categories and the credit spread puzzle

Michele Breton and Yaovi Gassesse Siliadin*

Wednesday 14$^{\text{th}}$ January, 2015

Abstract

Can the heterogeneity of rating categories explain why many studies have found that default risk accounts for only a small fraction of corporate bond spreads? To answer the question, we test structural models of default using firm level data, thereby accounting for the large dispersions, skewnesses and kurtoses of firm fundamentals inside rating categories. We find that once the model risk premia is specified consistently with historical default experience, the Merton [1974] model and the Leland and Toft [1996] model cannot explain the level of spreads for investment grade bonds. Furthermore, when the representative firm is calibrated to historical default experience and one ignores the heterogeneity of rating categories, the models performance is actually biased upward. The findings are robust to the uncertainty about ex-ante default rates.

*Both authors are from HEC Montreal. We are grateful to Pascal François, Georges Dionne, and Jan Ericsson, for their comments on the earlier versions of this paper. We also thank participants of the 2014 CEQURA risk conference in Munich, Germany, for their insightful discussion of this work. We thank the Montreal Exchange for their generous financial support.

1 Introduction

Rating categories are heterogeneous. Firms with different business risks and leverages might have the same rating and a firm might keep a constant rating while its fundamentals are fluctuating. This is not surprising because ratings categories were designed in first place to be in few numbers and to be a stable indication of a firm credit worthiness.

The heterogeneity of rating categories is closely related to the credit spread puzzle. This puzzle illustrates the poor performance of structural models in explaining investment grade bonds spreads. Many studies have found that heterogeneity helps address the puzzle [Feldhutter and Schaefer, 2014, David, 2008, Chen et al., 2009, Bhamra et al., 2010]. Other authors have found that the credit spread puzzle is robust to heterogeneity, in the sense that even after accounting for heterogeneity, the size of default risk in corporate spreads is tiny [Jones et al., 1984, Leland, 2004, Huang and Huang, 2012, Elton et al., 2001].

The goal of this paper is to bring some consensus on the role of heterogeneity in the credit spread puzzle. To achieve this goal we must adopt a framework where the impact of heterogeneity on the default risk size can be isolated unambiguously. There are two ways that heterogeneity can be introduced in the performance test of structural models. The first approach is to extend pure structural models by explicitly incorporating heterogeneity in the model structure. It is the approach pursued by David [2008], Chen et al. [2009], Bhamra et al. [2010]. The other approach is to apply the model on individual bonds, then compare average model spreads to average historical spreads for each category. It is the approach adopted by Jones et al. [1984], Elton et al. [2001], Huang and Huang [2012], Feldhutter and Schaefer [2014]. While the focus of studies adopting the first approach is to develop new structural models where there is no puzzle, that of studies adopting the second approach is to test existing models without modifying them.

The problem with the first approach is that the structure of the original structural model is modified when heterogeneity is added. The resulting models are complex and understanding how heterogeneity operates in those models becomes a puzzle in a puzzle. The perfect illustration of the puzzle in a puzzle is the fact that David [2008] found that heterogeneity leads to higher spreads because of the convexity of structural models while Chen et al. [2009] argued that the higher spreads obtained by David [2008] are not due to convexity but to time varying Sharpe ratios. The puzzling fact is that structural models are convex, hence it seems natural to observe a convexity effect. To avoid these unnecessary complications we adopt the second approach. The advantage of this approach is that the structure of the model is preserved.

There are several issues with previous studies adopting the second approach. The goal of Jones et al. [1984] was to test the Merton [1974] model but they ended up testing an extension of the original model for callable bonds. Their work failed to convince because of an unnecessary anti-dilution assumption in their extension which causes debt value not to increase monotonically with asset value. Elton et al. [2001] did not test structural models of default. They tested the Nelson and Siegel [1987] yield curve model with which they captured the expected loss component of spreads by assuming that investors are risk-neutral. But this expected loss is different from the spread predicted by a structural model because in the structural modeling framework investors are not risk neutral. Huang and Huang [2012] tested structural models but their main discussion was based on the assumption of homogeneity of rating categories. In an attempt to address the heterogeneity issue, they studied the robustness of their results when some form of heterogeneity is introduced in the data. The problem is that they did not consider the full distribution of fundamentals inside the rating categories. For the Aa category for instance, they assumed that half of the sample have a lower average of 17.2% ratio and the other half have 25.2%. Not only the full distribution is not reflected by this extreme simplification of heterogeneity, but the implied range of leverage is only 8% which is much smaller than what we observe in the data. Feldhutter and Schaefer [2014] also tested structural models using the second approach but they did not calibrate the average model default rates to historical default rates. This calibration is essential for two reasons. First, individual default rates do not need to match the historical average but the average of these individual default rates should match the historical average. Second, whether rating categories are heterogeneous or not, without calibration to historical default experience, there is no puzzle. It is known that structural models can generate high enough spreads. The puzzle is that they cannot generate those high spreads when they are constrained to fit historical default rates.

In the present study we embrace all these issues. We test pure structural models of default, we estimate and consider the full distribution of fundamentals inside rating categories and we calibrate the average of individual default rates to historical default rates. Our evidences are based on a sample of daily data on 1286 firms for a total of 2735873 firm-days from December 26, 1984 to December 31, 2013. To our knowledge, this is the longest dataset (in terms of period covered) ever used for this type of study. Our conclusions were subjected to several robustness checks regarding the key constants of our study such as historical default rates, recovery rates, the calibration procedure, the period of study and the considered structural model. Our findings are the following.

Rating categories are highly heterogeneous. The range of dispersion is much larger than that considered in previous studies such as those of Huang and Huang [2012] and Leland [2004]. The BBB category appears to be the most heterogeneous one. Overall, high grade rating categories are less heterogeneous than low quality rating categories. The categories are highly overlapping. For instance, many BBB rated firms have lower business risks than AA rated firms and many B rated firms are less levered than BBB firms.

When heterogeneity is accounted for, without the calibration, the size of default risk in spreads is large, just as Feldhutter and Schaefer [2014] have found. However, once default rates are calibrated, we find default risk sizes pretty much in the same range as Huang and Huang [2012], even after accounting for the large dispersions, skewnesses and kurtoses in rating categories.

We discuss the convexity puzzle: Is there any convexity effect or not due to heterogeneity of rating categories? We show that the existence of the convexity bias depends on the choice of typical firm. If the typical firm is the statistical average firm, then there is a convexity bias, though its size is small for high quality bonds. If the typical firm is not the statistical average firm then there is not necessarily a convexity bias. The convexity puzzle arises because in some studies one assumes that the typical firm is the statistical average firm while in others the typical firm is carefully chosen by economic and calibration arguments. In the latter case, the bias due to heterogeneity can be even negative. A negative sign of the bias is not a concavity effect, it simply reflects the fact that the average spread is smaller than the spread of the typical firm.

Our findings have important implications for credit risk modeling. First, the credit spread puzzle is not a statistical artifact due to the heterogeneity of rating categories. The problem seems to come from the type of risk adjustment that is implied by pure structural models of default. Second, the development of new structural models where one accounts for the comovements between the value of the default option and macroeconomic fluctuations on one hand and the comovements between the value of the default option and market liquidity conditions must be encouraged. Third, assuming homogeneity of rating categories can be judicious when testing structural models, provided that the typical firm is carefully calibrated.

The remainder of the paper is organized as follow. In section 2, we discuss the methodology of our analysis. First, we clarify what we mean by heterogeneity of rating categories and how we measure it. Second we discuss the issue of calibration to historical default rates and the measure of the bias due to heterogeneity. In section 3, we present details about the construction of our dataset, we provide some summary statistics and we study the heterogeneity of the rating categories. Section 4 presents the measures of the size of default risk in presence of heterogeneity, and discusses the issues related to the convexity bias. In section 5 we detail the robustness checks that we have performed. Section 6 presents further discussion of our results and section 7 concludes.

2 Methodology

2.1 Measure of heterogeneity

We measure heterogeneity by estimating the distribution of main firm fundamentals inside rating categories. We deliberately focus on the following fundamentals: asset volatility, leverage or equivalently solvency ratio, asset drift and recovery rates. We chose these fundamentals first because they are the common denominator of all structural models of credit risk. Second, our goal is to study the impact of this heterogeneity on the empirical test of a Merton Model with bankruptcy costs and these fundamentals are those required by the Merton Model (See appendix A for details on the Merton Model). After estimating a distribution, we are interested in its dispersion, its skewness and its kurtosis. These statistics give as an overall picture of how heterogeneous the rating categories are. In addition to the heterogeneity of each rating category, we are also interested in their overlap. Overlapping distributions across rating categories mean that these rating categories share common values of fundamentals. For instance it is common to assume that an asset volatility of 21% is typical to AAA rated firm. This assumption would come under question if we find that a significant proportion of B rated firms also have an asset volatility of 21%.

It is also possible to measure heterogeneity in rating categories by measuring heterogeneity in historical default rates and spreads. For instance, Longstaff and

Schwartz [1995] showed that spreads are heterogeneous across economic sectors. Proceeding this way would be interesting if we were interested in measuring the impact of heterogeneity of rating categories on empirical estimates of spreads and default rates. But this is not the goal of the current paper. Our goal is to study the impact of heterogeneity of the inputs of structural models on the outputs of structural models. Besides, it would be very surprising if firm fundamentals were heterogeneous inside rating categories but spreads are not.

Estimation of the fundamentals is standard (See Appendix B). This paper is not the first to estimate distributions of fundamentals but it is the first to do so for each rating category, using daily data over the period 1983 to 2013, and using the standard estimation procedure of fundamentals. This work is more importantly the first to emphasize the overlap between distributions. Elkamhi et al. [2012] estimated the fundamentals of the Leland and Toft [1996] model including asset volatility, distance to default, default boundary and total debt and found a large dispersion in their sample of firm-quarters covering the period 1970-2007. However, they did not compute the distribution inside rating categories. Bhamra et al. [2010] estimated the distribution of leverage ratios for a sample of BBB firms over the period 1997 and 2004 (see their figure 1) and found a large dispersion and a positive skewness. Chen et al. [2009] found that the average leverage ratio for Baa rated firms was time varying over the period 1975-1998 (see their figure 1). David [2008] also presented some empirical evidence of time varying leverage ratios for the average Baa rated firm using data from 1975 to 2001 (see their figure 1). Schaefer and Strebulaev [2008] estimated the distribution of asset volatility and leverage inside rating categories using data covering the period 1996-2006 but found rather low dispersions. Interestingly, though they do not mention it, their table 7 suggests that the distributions of leverage and asset volatilities are highly overlapping. They found that average asset volatility was in the range 21-23% for all rating categories with standard deviations ranging from 5% to 8%. It is noteworthy that their estimation procedure is not standard. Feldhutter and Schaefer [2014] estimated the distribution of fundamentals inside rating categories. But their estimates are based on industrial bonds only and their samples covers the very recent period 2002-2012 only. They also made the problematic assumption that market value of assets is equal to book value of debt plus market value of equity.

2.2 Heterogeneity bias and calibration to historical default rates

2.2.1 Should we calibrate structural models to historical default rates?

An issue that deserves special attention when dealing with the credit spread puzzle is the calibration to historical default rates. Structural models yield three main outputs: the spread, the risk-adjusted default rate, and the objective default rate. Should we require that the model objective default rate matches its historical counterpart? We might be tempted to answer no by invoking the following popular argument. For pricing purposes, it is the risk-adjusted default rates that matters and for risk management purposes it is the objective default rate that matters (see for instance Hull [2012]). We would then conclude that we should not care about historical default rates because we would like to test the ability of structural models to price securities.

There is nothing fundamentally wrong with this argument except that the credit spread puzzle is neither a pure pricing problem nor a pure risk management problem. It is an Economics problem which involves both pricing and risk management. It is about explaining the observed spreads on corporate bonds. Since Economics rests on empirical observations, it is only natural to require that structural models match empirical default rates when attempting to explain the credit spread puzzle. That's when the dilemma arises. Structural models seem unable to generate spreads as high as those observed empirically once they are forced to generate objective default rates as low as those observed empirically. This is the credit spread puzzle as defined by Huang and Huang [2012]. The credit spread puzzle is therefore, by definition, closely related to the calibration to historical default rates. No calibration, no puzzle. As a result, any paper dealing with the puzzle should calibrate the model objective default rates to historical default rates.

In a recent paper about the credit spread puzzle, Feldhutter and Schaefer [2014] found that the puzzle might be a myth. This should surprise noone since they did not calibrate their structural models to historical default rates. But, they justified their choice not to calibrate by the uncertainty about the true average default rates. Obviously, if historical default rates are biased estimates of the true default rate then it does not make sense to rely on them for calibration purposes. The authors argued that ex-post default rates can be significantly different from ex-ante default rates due to systematic risk in the economy. Thus, the true default rate could be higher or lower than the realized default rate. The question is that why would a rational investor assume that the true default rate is higher than the realized default rate, ignoring it could be lower? We believe that, in the wake of uncertainty, the realized default rate which is at mid-way between conservative and exaggerated estimates is a reasonable and compromising choice. Another problem with their argument is that the large uncertainty about the true default rate arises only when they introduce systematic risk in their simulated economy, the size of which is controlled by a correlation parameter. They considered a range of correlation parameters without any empirical backing. The considered range might be much higher than actual correlations. Without empirical evidence on the actual size of this systematic risk, the uncertainty about the true average default rates might have been overstated. Finally their conclusions are based on a simulated economy and thus cannot claim generality. Later, in our analysis, we show that even when one accounts for the uncertainty over the true default rate estimates, corporate credit spreads are still a puzzle.

2.2.2 The standard calibration approach

Let's denote by Φ_T the objective default rate function of a particular structural model for a given horizon T. Let's denote by Ψ_T the spread function of the same structural model for a given horizon T. For the Merton Model, Φ_T is a function of asset volatility σ, asset drift μ, and leverage λ. Ψ_T is a function of asset volatility σ, leverage λ, recovery rate ρ and the risk free rate r. For the exact expressions of the functions Φ_T and Ψ_T in the Merton model, refer to appendix A.

The standard way of gauging the performance of a structural model is the typical firm based approach. For instance, the average historical spread for BBB rated firms is known. The idea is to consider values of σ, ρ and λ that are typical to BBB firms and show that $\Psi_T(\lambda, \sigma, \rho, r)$ is close to the observed BBB spread. There is no standard way of defining the typical firm, in some studies, one uses statistical averages, in others one uses statistical medians, in others one uses plausible values backed by some economic intuition.

The observation of Huang and Huang [2012] is that it is possible to find values of σ, ρ and λ that make $\Psi_T(\lambda, \sigma, \rho, r)$ sufficiently high but this might be implying a too high $\Phi_T(\lambda, \sigma, \mu)$ for plausible values of μ. To prove their point, they calibrated $\Phi_T(\lambda, \sigma, \mu)$ to its historical level before comparing $\Psi_T(\lambda, \sigma, \rho, r)$ to its historical counterpart. The calibration consists of the following. Consider a given rating category. Denote the historical average leverage, asset volatility, recovery rate, asset drift, default rate and spread for this rating category by $\overline{\lambda}, \overline{\sigma}, \overline{\rho}, \overline{\mu}, \overline{\phi}_T, \overline{\psi}_T$ respectively. Calibrating to historical default rates consists of setting the asset volatility to the value of σ which solves the following equation

$$\Phi_T\left(\overline{\lambda}, \sigma^*, \overline{\mu}\right) = \overline{\phi}_T.$$

In practice because $\Phi_T\left(\overline{\lambda}, \overline{\sigma}, \overline{\mu}\right)$ is very low, one finds that $\sigma^* > \overline{\sigma}$.

The size of default risk explained by the model is then computed as the ratio

$$\frac{\Psi_T\left(\overline{\lambda}, \sigma^*, \overline{\rho}, r\right)}{\overline{\psi}_T}$$

Huang and Huang [2012] found that this gauge ratio is typically low.

2.2.3 The heterogeneity issue

Since $\overline{\phi}$ is an average default rate, the correct calibration equation is to require that

$$\overline{\Phi}_T = \overline{\phi}_T,$$

where $\overline{\Phi}_T$ is the average of individual default rates predicted by the model for firms inside the rating category.

$$\overline{\Phi}_T = \frac{1}{n}\sum_{i=1}^{n}\Phi_T\left(\lambda_i, \sigma_i, \mu_i\right).$$

And the correct gauging ratio is then

$$\frac{\overline{\Psi}_T}{\overline{\psi}_T},$$

where

$$\overline{\Psi}_T = \frac{1}{n}\sum_{i=1}^{n}\Psi_T\left(\lambda_i, \sigma_i, \rho_i, r\right),$$

n is the sample size. Unless the rating category is homogeneous, due to the non-linearity of Φ_T,

$$\overline{\Phi}_T \neq \Phi_T\left(\overline{\lambda}, \overline{\sigma}, \overline{\mu}\right).$$

We also have that:

$$\overline{\Psi}_T \neq \Psi_T\left(\overline{\lambda}, \sigma^*, \overline{\rho}, r\right),$$

and this second inequality has nothing to do with the non-linearity of the spread function Ψ_T because it would hold even if Ψ_T were linear. The reason is that σ^* is different from $\overline{\sigma}$.

Accounting for heterogeneity might therefore change the conclusions of Huang and Huang [2012]. The hope is that we might be able to calibrate $\overline{\Phi}_T$ to $\overline{\phi}_T$ and still have

$$\frac{\overline{\Psi}_T}{\overline{\psi}_T} > \frac{\Psi_T\left(\overline{\lambda}, \sigma^*, \overline{\rho}, r\right)}{\overline{\psi}_T}.$$

2.2.4 Measuring the heterogeneity bias

Some of the authors that discussed the issue of heterogeneity in empirical tests of structural models argued that heterogeneity causes a bias because the spread function is convex in its arguments [David, 2008, Feldhutter and Schaefer, 2014]. They define the convexity bias as:

$$\overline{\Psi}_T - \Psi_T\left(\overline{\lambda}, \overline{\sigma}, \overline{\rho}, r\right).$$

Indeed, If Ψ_T is convex in (λ, σ, ρ) then:

$$\frac{\overline{\Psi}_T}{\overline{\psi}_T} > \frac{\Psi_T\left(\overline{\lambda}, \overline{\sigma}, \overline{\rho}, r\right)}{\overline{\psi}_T}.$$

This suggests that the bias due to heterogeneity is positive, i.e, if one takes into consideration heterogeneity, one should find larger size of default risk than when one ignores heterogeneity.

However, in many cases, the bias due to heterogeneity cannot be signed or even sized using convexity arguments. This is because, as we explained earlier, the typical firm is not always assumed to be the average firm. It is often carefully calibrated by economic arguments and it changes from study to study and from model to model (see table 1). In the case of Huang and Huang [2012] for instance, the typical firm is calibrated to match historical default experience. The bias is this case is equal to:

$$\overline{\Psi}_T - \Psi_T \left(\overline{\lambda}, \sigma^*, \overline{\rho}, r\right).$$

Even if Ψ_T is convex, we cannot sign $\overline{\Psi}_T - \Psi_T \left(\overline{\lambda}, \sigma^*, \overline{\rho}, r\right)$ because σ^* is different from $\overline{\sigma}$.

More generally we define the heterogeneity bias as:

$$\overline{\Psi}_T - \Psi_T \left(\lambda^*, \sigma^*, \rho^*, r\right),$$

where the parameters $(\lambda^*, \sigma^*, \rho^*)$ represent the fundamentals that are assumed by a given researcher to be representative of a particular rating category. These representative parameters are not necessarily the statistical average fundamentals inside the considered rating category.

The heterogeneity bias can be negative depending on the choice of representative parameters. For instance keeping other parameters at their average, one can easily solve for a leverage ratio $\widetilde{\lambda}$, such that:

$$\overline{\Psi}_T - \Psi_T \left(\widetilde{\lambda}, \overline{\sigma}, \overline{\rho}, r\right) = 0.$$

If one chooses $\lambda^* > \widetilde{\lambda}$, then the bias becomes negative because, Ψ_T is strictly increasing in λ. A negative bias, implies that accounting for heterogeneity would actually reduce the size of default risk compared to the homogeneity case.

3 Data and evidence of heterogeneity

3.1 Sources

Our sample covers all North American firms common to the databases COMPUSTAT and CRSP for which we have available data, excluding financial firms. The sample includes daily data on 1286 firms for a total of 2735873 firm-days from december 26, 1984 to december 31, 2013. To construct the database, we retrieved quaterly data on debt in current liabilities (COMPUSTAT variable name: dlcq) and long term debt (COMPUSTAT variable name: dlttq), economic sector (COMPUSTAT variable name: spcseccd) from COMPUSTAT. Regarding credit ratings, we used S&P credit rating on domestic long term debt issuer (COMPUSTAT label: splticrm). The rating data is available at monthly frequency. We also retrieved daily data on share closing price (CRSP variable name PRC), daily number of outstanding shares (CRSP variable name: SHROUT), share code (CRSP variable name: SHRCD), and ex-dividend daily returns (CRSP variable name: RETX) from CRSP. For the risk-free rate r, we used the monthly data on one year Treasury Constant Maturity Rate obtained from the Board of Governors of the Federal Reserve system.[1]

After first round cleaning (keeping common shares only, merging the dataset and keeping observations with no missing data and strictly positive debt principal), the dataset consisted of 1536 firms from January 2, 1981 to December 31, 2013 for a total of 4794915 firm-days. Using this dataset, we estimated daily firm asset values and asset volatilities using a rolling window of 4 years (See appendix B for details on estimation procedure). We then excluded financial firms and observations for which the estimation procedure did not converge. Finally, to ensure that our analysis is not driven by extreme values and outliers, we winsorized the data by eliminating the first and last percentiles. Table 2 shows the number of firms and observations in the final sample by rating and

[1] Available at http://research.stlouisfed.org/fred/data/irates/gs1 (H.15 Release).

economic sector.

For historical default experience, we used historical default rate data over the period 1981 to 2013 from the table 24 of Standard and Poors [2014]. The historical default rate data is reprinted in table 3. Average recovery rates by rating categories for senior unsecured bonds over the period 1982-2010 are retrieved from Exhibit 22 of Moody's [2011]. Average recovery rates by economic sector are retrieved from table 4 of Standard and Poors [2013]. Tables 4 and 5 show the recovery rates by rating and by economic sector respectively. Unfortunately, we did not have data on average recovery rates by economic sectors inside a given rating category, but this is exactly what we need. So we suggest a procedure described in appendix B.3 to generate these recovery rates. This procedure allows us to assign different recovery rates to firms inside a given rating category based on their economic sector, while ensuring that the average recovery rate for the rating category exactly match its historical counterpart. The obtained recovery rates are shown in table 6. Later in our robustness checks, we consider an alternative way of introducing recovery rates which consists of attributing random recovery rates to firms inside a given rating categories based on a normal distribution with mean equal to the historical average recovery rate for that rating category and a large coefficient of variation of 50%.

For historical spreads, we rely on the estimates of Huang and Huang [2012]. There are many reasons for this. First, we could have used data on actual transaction prices from Trace database but these prices are available only from 2002. Using TRACE would mean throwing away almost two third of the period that we cover. Second, historical average spreads in our study serve only as benchmark for comparison purposes. Since many studies rely on the Huang and Huang [2012] estimates, they seem to be a good choice of benchmark. In particular, using our own estimates of the historical spreads as benchmark would be problematic. It would be hard to tell whether our findings are due to a different benchmark than Huang and Huang [2012] or to heterogeneity. Finally, we have performed some robustness checks regarding the major problem that the use of the Huang and Huang [2012] estimates can pose. That problem arise because the Huang and Huang [2012] estimates are based on a different period than the one we consider in this study. In our robustness checks, we consider the case where we reduce our sample to the same period used by Huang and Huang [2012]. This did not change our conclusions.

3.2 Summary statistics and heterogeneity of rating categories

Tables 7 provides summary statistics of asset volatilities, asset drifts and leverage ratios by rating category. As one can see, there is substantial heterogeneity inside rating categories. Consider the BBB category for instance. Asset volatilities vary from 11.47% to 363% with an interquartile range of 22.64% to 36.16%. Asset drifts vary from -28.19% to 220.93% with an interquartile range of 3.41% to 19.50%. Leverage ranges from 4.88% to 1044.69% with an interquartile range of 19.85% to 54.69%.

By all measures of heterogeneity for each fundamental, one can conclude that there are important dissimilarities between firms inside rating categories. Coefficients of variations are large suggesting that there is a sizable variation around the means. The ranges are substantial even when we discard the first and last percentiles. The interquartile ranges are also important indicating that the observed large dispersions are not driven by extreme values. We also observe substantial positive skewness and kurtosis in leverage, asset drift and asset volatility. This shows that required risk premia can be substantially large for some firms at some times and leverage ratios and asset volatilities can be extremely high relative to the average. The largest skewnesses and kurtoses are observed for BBB rated firms. Globally, heterogeneity is less important in high quality rating categories than in low quality rating categories.

Table 7 also shows that distributions are highly overlapping. We often think of the median asset volatility of a given rating category to be exclusively typical to firms inside that category. However, the overlapping distributions contradict this common assertion. For instance, the median asset volatility for BBB rating category is 28.68% in our sample. It is true that this asset volatility is typical to BBB firms but it is not typical to BBB firms only. This value of 28.68% is close to the center of the distribution of AA and A rating categories also. In other words, a significant proportion of AA firms operate riskier businesses than typical BBB firms. And a significant proportion of BBB rated businesses are safer than AA rated firms.

Overlapping distribution may lead to a selection bias problem. To test whether a model correctly predicts the spread for a firm with 28.68%[2] asset volatility, many authors would compare the model prediction at the 28.68% asset volatility to the average spread observed for the BBB category (see for instance Longstaff and Schwartz [1995], Leland [1994], Leland and Toft [1996], He and Xiong [2012]). They would do so on the ground that 28.68% is typical to BBB firms. However, because the distributions are overlapping, a significant number of firms with different ratings than BBB also have asset volatilities of 28.68%. Thus these firms should also be included in the sample used in the test. Systematically restricting the sample to BBB firms may bias the test. One cannot easily predict the sign and size of this bias because the unbiased sample would consist of firms with lower ratings than BBB but also of firms with higher ratings than BBB. Firms with lower credit quality would tend to increase the average spread and default rate, while those with higher credit quality would tend to decrease them. The net effects on spreads and default rates are hardly predictable. Studying this selection bias would prove interesting in that respect. However, this would require empirical observations of spreads and defaults for firms with the same asset volatility but different ratings and this it is out of the

[2]This median asset volatility differ from study to study.

scope of this paper.

The large dispersions, skewnesses, kurtoses and the highly overlapping distributions suggest that the assumption of homogeneity can be hardly justified. However, would considering the whole distribution of fundamentals inside each rating category change conclusions about the credit spread puzzle? We attempt to answer this question in the next section.

4 The heterogeneity bias and the credit spread puzzle

One has to appreciate the following findings on the heterogeneity bias in light of the credit spread puzzle. This puzzle refers to the fact that many authors have found that structural models of default, in particular the Merton Model, are unable to fit the level of historical spreads for investment grade bonds. This under-performance is more pronounced at the short end of the yield curve and for high quality bonds. These studies have concluded that there must be other factors at play such as liquidity risk and macro-economic risks. However, many studies that have pioneered the credit spread puzzle have made the assumption of homogeneity of rating categories. The main question is whether relaxing this assumption would void the puzzle.

Some of the authors that have dealt with the heterogeneity issue have downplayed its magnitude. They include Leland [2004] who considered only a 10% variation around average leverage and Huang and Huang [2012] who found no sizable convexity effect after accounting for heterogeneity. Huang and Huang [2012] considered only a rough approximation of the actual heterogeneity. For the Aa category for instance, they assumed that half of the sample have a lower average of 17.2% ratio and the other half have 25.2%. Not only the full distribution is not reflected by this simplification, but the implied range is only 8%

which is roughly 10 times lower than the actual range as estimated in the previous section. It is therefore of interest to study the impact of the full distribution inside rating categories on the credit spread puzzle.

Others have found that heterogeneity helps in solving the puzzle but additional features have to be included in structural models for it to have a sizable positive effect on the spreads. Among those studies we can cite David [2008], Chen et al. [2009] and Bhamra et al. [2010]. David [2008] found that introducing time varying leverage ratio in the Merton Model increases the BBB-spread by 59%. But the obtained spread of 79 basis points is still far from the historical average BBB spread which is around 158 basis points. To get a spread of around 106 basis points a stochastic asset volatility combined with some signaling effect of inflation is needed. Chen et al. [2009] found that if the Merton Model is calibrated to match historical default experience and sharpe ratios, then ignoring heterogeneity in solvency ratios would actually increase the spreads by a few basis points. To get higher spreads Chen et al. [2009] introduced a time-varying risk premium in the model by means of a habit persistence pricing kernel. Bhamra et al. [2010] found that heterogeneity in the BBB sample helps calibrate their model to historical default rates while keeping the spreads high. However, they had to incorporate in their model an Epstein-Zin utility function with intertemporal macro-economic risk.

Our approach to capturing the effect of heterogeneity on the performance of the Merton Model is quite different from the previous studies. On a daily basis, we price different bonds using the Merton Model. We do not introduce any time-varying macroeconomic conditions or changing price of risk in the pricing model. It is really the basic Merton model. We compute the spread implied by the Merton model for each of these observations and compute the average of these spreads rating wise. It is this model average spread that should be compared to the historical average spread and it is what we do. When the typical firm is the statistical average firm, the difference between our model average spread and the spread for the typical firm would reflect a convexity effect just like the one highlighted by David [2008]. The difference with David [2008] though is that we do not incorporate time-varying solvency ratios in the Merton model, and we do not claim that investors, learning about the macro-economic condition through the signaling effect of inflation, demand any additional compensation for time-varying solvency ratios.

Feldhutter and Schaefer [2014] also adopts our approach but their measure of heterogeneity bias is upward biased. They made the assumption that all the parameters of the typical firm are the statistical averages of the considered rating category. As our table 1 shows, the typical firm is not always the statistical average firm, especially in the paper of Huang and Huang [2012] where the typical firm is calibrated to historical default rates. To measure the true bias due to heterogeneity in the work of Huang and Huang [2012], we need to calibrate the typical firm to the historical default rate before measuring the bias. Second, Feldhutter and Schaefer [2014] did not ensure that the average physical default rate generated by their model matches the historical average default rate. Not surprisingly, they found that the convexity bias is very large and that the credit spread is highly likely just a myth. To isolate the true impact of heterogeneity, we must make sure that in both the homogeneity and heterogeneity cases, the model average default rate exactly matches the historical default rates. We expect these calibrations to reduce the bias due to heterogeneity and eventually make it negative. Finally, contrarily to us, they paid no attention to the heterogeneity in recovery rates. We have introduced heterogeneity in recovery rates across economic sectors as detailed in appendix B. We are unable to predict however what the effect of the heterogeneity in recovery rates would be, so we let the data decide.

We now present our empirical measures of the size of default risk and the heterogeneity bias. As explained in section 2.2.4, we distinguish between the convexity bias and the heterogeneity bias. In the strict sense, convexity bias is a particular case of heterogeneity bias where one assumes that the typical firm is the average firm. An the bias in the case of the paper of Huang and Huang [2012] is also a particular of heterogeneity bias where the typical firm is calibrated to match historical default rates. We refer to this second bias as the heterogeneity bias, because it is the one that really matters for the credit spread puzzle.

4.1 Size of default risk

Table 8 compares the default sizes with and without calibration. When model default rates are not calibrated to historical default rates, default sizes are large except for the AA and AAA rated bonds (see panel A of table 8). In particular, there seems to be no credit spread puzzle for BBB rated bonds. The model generates average BBB spreads that are 36% larger than historical spreads at the 4 year horizon and that are only 8% less than historical spreads at the 10 year horizon. The puzzle appears to be also less severe for A rated bonds. The model explains 55.4% of the historical A spreads at the 4 year horizon and 45.4% of the historical A spreads at the 10 year horizon. In the homogeneity case, Huang and Huang [2012] estimated the default sizes for the A spreads to be only 10.3% and 19% respectively at the 4 year and the 10 year horizons.

For the default sizes without calibration to be satisfactory, the corresponding model physical default rates must match their empirical counterpart. Table 9 compares the average of individual default rates predicted by the uncalibrated model to the historical average. Except for the AAA category at the 4 year horizon, the model physical default rates are much larger than the historical averages. For the BBB category, the model predicts default rates that 6.5 times larger than historical default rates at the 4 year horizon, and 3.6 times larger at the 10 year horizon. For the A category, the model predicts default rates that 6.4 times larger than historical default rates at the 4 year horizon, and

4.1 times larger at the 10 year horizon. Thus, without calibration to historical default rates, spreads are large but the corresponding default rates are too high.

We now calibrate the model to historical default rates as explained in appendix B.2. The calibration essentially consists of adjusting individual leverages to the level imposed by historical default rates. In practice, the amount of debt due at a given maturity is not known but it is what matters in the Merton model. The philosophy of the calibration is to let historical default rates decide the amount of the debt required at a given maturity. Later in the robustness checks, we show that alternative calibration approaches do not change our conclusions. The calibrated leverages are shown in table 10. Without surprise, the calibrated leverage ratios are much smaller than the actual total leverage ratios. For the BBB category, the actual average leverage is 40.1%, but the leverage that justifies the historically observed 1.53% default rate at the 4 year horizon is 14.14% and the leverage that justifies the historically observed 4.33% default rate at the 10 year horizon is 10.43%. For the A category, the actual average leverage is 26.44%, but the leverage that justifies the historically observed 0.43% default rate at the 4 year horizon is 12.95% and the leverage that justifies the historically observed 1.59% default rate at the 10 year horizon is 9.27%.

Once the model is calibrated, it can explain only a small fraction of the observed spreads (see panel B of table 8). The model now explains only 39% of BBB spreads at the 4 year horizon compared to 136% without the calibration. At the 10 year horizon, it now explains only 34% of the observed BBB spreads compared to 91.8% without the calibration. For A bonds, the share of the spreads explained by default risk is brought down from 55.4% to 11% at the 4 year horizon and from 45.4% to 11% at the 10 year horizon. For very high quality bonds i.e AA and AAA bonds, the model explains only a tiny fraction of the spreads with or without calibration. This is less of a puzzle because these bonds are very safe and one would not expect credit risk to explain their prices. These findings show that even after accounting for heterogeneity, there is a credit spread puzzle once the model is calibrated to historical default rates. In the rest of this paper, the model is calibrated to historical default rates.

4.2 The convexity bias issue

We now turn to the convexity bias issue. To examine the size of this bias, we must assume, by definition, that the typical firm is the statistical average firm, i.e, the firm whose fundamentals are the average fundamentals observed the rating category of interest. Table 11 shows the term structure of the convexity bias in basis points by rating category. Panel A assumes homogeneity of recovery rates while panel B relaxes this assumption by introducing heterogeneity in recovery rates across economic sectors. The bias ranges from 0 basis points for one-year and 2-year AAA-rated bonds to 362 basis points for 1-year B-rated bonds. Overall, when one uses the average firm as typical firm, there is a convexity bias especially at medium and long horizons and for low grade bonds. The bias is larger for junk bonds than for investment grade bonds. The bias is typically around 5 basis points for high quality bonds and typically above 100 basis points for junk bonds. For investment grade bonds, the bias is larger at long maturities than at short maturities. For low quality bonds, the bias is larger at short maturities than at long maturities. For BBB-rated bonds the bias is no larger than 59 basis points.

It is striking that heterogeneity in recovery rates has no impact on the size of the convexity bias. Introducing heterogeneity in recovery rates either does not change the convexity bias or increases the convexity bias by only 1 basis points. An exception is the case of B-rated bonds at the 4 year maturity where heterogeneity in recovery rates actually reduced the convexity bias by 1 basis points.

These results show that when the typical firm is the average firm, there is a convexity bias, the size of which can be relatively important, especially for junk bonds.

However, the typical firm is not always the statistical firm. In the benchmark paper of Huang and Huang [2012], the asset volatility of the typical firm is carefully calibrated to match historical default rates. Huang and Huang [2012] did not actually consider the Merton model in their paper but their calibration procedure can be roughly replicated for the Merton Model. Essentially, it consists at calibrating the asset volatility while fixing the other fundamentals of the typical firm at the rating average. If there is a bias attributable to heterogeneity in this procedure, it should be reflected in the difference between the calibrated model average spread and the spread at the calibrated average. The convexity bias does not fairly represent the heterogeneity bias because it relies on an asset volatility which is not calibrated to historical default rates. Because of that, we expect the veritable heterogeneity bias to be smaller than the convexity bias, if not negative. Indeed, the model default rate for the average firm is generally smaller than the historical average. Therefore, the calibration to historical default rates would require a higher asset volatility. This in turn would generate higher spreads for the typical firm and a lower bias.

Panels A and B of table 12 show our measure of the heterogeneity bias without and with heterogeneous recovery rates. As in the convexity bias case, heterogeneity in recovery rates has virtually no impact on the size of the bias. More importantly, as predicted, the heterogeneity bias is much smaller than the convexity bias and it is negative in most cases. The only exception is the Below-B category where the bias is still large. At the one year horizon, except for the A rating category, the size of the bias is still positive although it is reduced. For the BBB rating category, the bias is reduced but still positive at short horizons. The reason of the reduced bias size is that the calibrated asset volatility is higher than the average asset volatility. We show in table 13 the calibrated asset volatilities compared to the average asset volatilities.

The negative heterogeneity biases do not mean that after accounting for heterogeneity the model spreads would be necessary smaller than those obtained by Huang and Huang [2012]. They mean that had Huang and Huang [2012] tested the Merton model using our calibrated typical firm, they would have ended up with slightly higher default size than what they found in some cases. This would mean an upward bias regarding the true default size.

In conclusion, if a study assumes homogeneity of rating categories and defines the representative firm as the statistical average firm, then their model spread will be downward biased, though the size of the bias will be almost zero for high quality bonds. This bias can be correctly named a "convexity bias" because its existence can be mathematically demonstrated by convexity arguments. However, if the study uses a representative firm which is not the average firm, then there might be an heterogeneity bias but there is no ground to call the bias "convexity bias" because its existence cannot be mathematically demonstrated by convexity argument. There is no basis either to call the bias 'non-linearity bias' because even if structural models were linear, which is of course not the case, the bias would exist. In the case of Huang and Huang [2012], our findings show that the bias due to heterogeneity is actually negative in most cases.

5 Robustness checks

We now study the robustness of our findings with respect to some key assumptions. We consider the following cases:

- Case 1: Average default probability is increased by 50%.
- Case 2: Heterogeneity in recovery rates is increased by attributing truncated normally distributed random recovery rates to firms inside rating categories while still matching average historical recovery rates.
- Case 3: The model default rate for each individual bond is calibrated to match the historical default rate for the corresponding rating category and

horizon.

- Case 4: Leverages are fixed to our sample average and asset volatilities are adjusted to match historical default rates.
- Case 5: Leverages are fixed to the same average as Huang and Huang [2012] and asset volatilities are adjusted to match historical default rates.
- Case 6: The sample period is restricted to 1985-1998, the same as Huang and Huang [2012].
- Case 7: An advanced structural model: Leland and Toft [1996] model.

For clarity and space parsimony, we present for each case only the size of default risk explained by the model after accounting for heterogeneity. We do not present the convexity biases and heterogeneity biases for each case. The results are shown in table 14. The base case scenario is the one used in the earlier part of our study and is described in details in appendix B.

5.1 Sensitivity to historical default rates

A key parameter of our analysis is the average historical default rate. This parameter is not measured with certainty. Keenan et al. [1999] provided some estimates of the volatility of one year default rates and ten year default rates by rating categories. At the one year horizon, the volatility of default rates range from 0% for AAA bonds to 0.28% for BBB bonds and 4.99% for B rated bonds. At the 10 year horizon, the volatility of default rates range from 1.93% for AAA bonds to 10.23% for BBB bonds and 15.94% for B rated bonds. Overall the volatility of default rates tends to decrease with credit quality and to increase with time to maturity. The authors examined cohorts of bonds from 1970 to 1998 which makes roughly 28 cohorts. Assuming independence of the cohorts, the volatility of each sample average default rate could be taken to the volatility of the sample divided by the square root of the sample size. This yields an estimate of coefficients of variation of average default rates no greater than 25% for each rating category. These estimates show that the uncertainty about average default rates can be substantial. The true coefficient of variation could be even larger than 25% due to correlations between the cohorts. It is therefore possible that the puzzle is due to larger true default rates than what the historical averages. To control for this possibility, we consider a worst scenario case, and increase the average historical default probabilities by 50% and measure the bias-free default sizes. The results are shown in column labeled "case 1" of table 14. The obtained default sizes are larger as one could have expected. Compared to the Huang and Huang [2012] estimates, the default size for BBB bonds have more than doubled at the 4 year horizon and is close to 50%. However, compared to our base case there is no substantial increase that could challenge our previous conclusions. Default sizes for investment grade bonds are still small.

5.2 Recovery rates

It could be surprising that heterogeneity in recovery rates have no sizable impact. This might be due to the way we generated the heterogeneity in recovery rates. We used a highly stylized approach to assign different recovery rates to firms inside the same rating category according to their economic sectors, while ensuring that at each maturity the average recovery rate matches empirical observations. One might argue that this approach does not generate enough heterogeneity of recovery rates inside rating categories. The ideal would have been to have a specific recovery rate for each individual bond. However, it is impossible to observe recovery rates for individual firms while they have not defaulted. One must rely on categorical level measures which are provided by rating agencies. There exists data on average recovery rates by year but we did not attempt to use them because what matters in our simple Merton model with bankruptcy costs is the expected recovery rate at maturity. It is not the recovery rate the year of pricing. At the pricing moment, we know the year at which the bond matures but we do not yet know the recovery rate for that year. With these data constraints, our best guest for the recovery rate of a particular firm is the historical average default rate observed for the corresponding rating category, maturity and economic sector.

However, what if the actual recovery rates were more variable inside rating categories than our best guess implies? As an attempt to deal with this issue, we consider larger heterogeneity in recovery rates than in our base case. We assume that for each rating category and maturity, recovery rates follow a normal distribution with mean equal to the historical average for that rating category and maturity. We arbitrarily consider a coefficient of variation of 50%. Recovery rates for individual firms are drawn from this distribution but are truncated at the maximum of 99% and a minimum of 1%. This procedure leads to much more heterogeneity in recovery than in our base case. Even inside the same economic sector, for the same rating and horizon, firms have different recovery rates. The assumed range of dispersion of recovery rates inside rating categories is now 98% while it was roughly 30% in our base case.

The new default sizes are displayed in the column labeled “case 2” of table 14. One can see that there is no marked difference from our base case findings.

5.3 Alternative calibration approaches

We have calibrated the leverage ratios in our analysis so that the model average default rates for each rating category match their corresponding historical counterpart. This calibration procedure have two particularities. First we do not force individual default model rates to match historical default rates, we only require that average model default rates match historical default rates. The basis of this approach is that in presence of heterogeneity, historical default rates are not individual default rates, they are average default rates. The drawback however with this approach is that we impose a uniform adjustment to leverage ratios. What if the adjustment needed to be firm specific? Second, we calibrate leverage ratios instead of asset volatilities. We obtained lower leverage ratios than previous studies. These lower leverage ratios reflects the reality that at a given maturity, the firm will default only if it misses the payments due at that time. These payments are match lower than the overall principal. One might argue that our earlier findings are due to this atypical calibration procedure and to our low leverages.

To address these concerns, we consider three alternative calibration procedures. The first alternative is to force the model default rate of each individual firm to match historical levels, thereby allowing firm specific adjustments. The second alternative is to calibrate the asset volatility instead of leverage, in such a way that leverage is kept at our sample average. The third alternative is to consider the same average leverage ratios as Huang and Huang [2012] and calibrate asset volatilities to match historical default rates. The results are shown in columns labeled "case 3", "case 4" and "case 5" of table 14. The results speak for themselves. None of these alternative procedures have substantially increased the explained default size for investment grade bonds. In most cases, the model default size is even lower than in the base case. We conclude that our findings are neither due to our calibration approach, nor to the difference between our sample average leverage and the Huang and Huang [2012] average leverage.

5.4 Period of study

Our findings are based on the period 1984 to 2013. However, we compare our model spreads with historical spreads estimated by Huang and Huang [2012]. These historical estimates are based on the period 1973-1993 for 10 year investment grade bonds and on the period 1985-1998 for 4 year investment grade bonds. Despite the fact that there is little disagreement between different measures of historical spreads (see table 4 of Bhamra et al. [2010]), it could be the case that this difference in time frames bias our measured default sizes downward. To account for this possibility, we reconducted our study over the period 1985-1998, to be consistent for the time frame for 4 year investment grade bonds. The results are shown in the column labeled “case 6” of table 14. The new default sizes are smaller for high quality bonds and are larger for lower quality bonds than in the base case. Our conclusion remains the same as in the base case though. Heterogeneity alone cannot explain the credit spread puzzle.

5.5 Is it specific to the Merton Model?

Would our findings change if we consider more elaborated structural models? We consider here the Leland and Toft [1996] model, which is one of the most successful extensions of the Merton Model. Some particularities of this model relatively to the Merton Model are the following. Default occurs endogenously once the firm value falls below an optimally determined level. There are corporate taxes. The firm continuously pay a coupon and replaces maturing debt with new debt at par. We leave the details about the model and its calibration to appendix C. Results based on this advanced structural model are given in the column labeled “case 7” of table 14. The calibrated spreads are even lower than in our base case. This shows that our findings are not specific to the Merton Model.

6 Further discussions

Many studies have attempted to determine the portion of the yield spread on corporate bonds attributable to default risk. These studies could be traced back to Jones et al. [1984] who used data on 27 firms on a monthly basis from January 1975 through January 1981. They found that for investment grade bonds, their extension of the Merton model to callable bonds is not an improvement over a naive model which assumes that debt is risk free. Their findings failed to convince mainly because their extension implied that debt value does not increase monotonically with firm value. Our study does not suffer from the same problem since what we test is really the basic Merton model. It is interesting that we arrive roughly at the same conclusion as Jones et al. [1984] that the Merton model has difficulty in pricing investment grade bonds.

6.1 Expected losses, taxes, market factors and liquidity

Another important work on the credit spread puzzle is that of Elton et al. [2001]. They found that a risk-neutral investor would require only a fraction of actual spreads on investment grade bonds to be fully compensated for the risk of default in a tax-free world. Part of the remainder of the observed spreads is due to the tax differential between government bonds and corporate bonds. The coupons on the former are tax free while the coupons on the later are subjected to both state and federal taxes. They argued that the part unexplained by risk-neutral default risk and tax premium is a risk premium which is related to market risk factors of Fama and French [1992]. However, their work cannot be considered as a full test of the structural modeling approach to credit risk measure. In the structural modeling framework, investors are not risk-neutral. In the basic Merton model, they require a constant risk adjustment to expected returns. Since Elton et al. [2001] use historical default probabilities, what they measure is only the expected loss component of the default risk. Based on their results, we cannot conclude on the size of default risk explained by structural models. In our approach, which is an extension of the approach of Huang and Huang [2012], the constant default risk premia required in the Merton model as well as the expected default loss are fully taken into consideration. We found that when the constant default risk premia is specified consistently with historical default rates, the size of default risk predicted by structural models is considerably low. We do not attempt to explain the remainder of the yield spreads. Previous studies have shown known that the remainder is related to taxes, market risk, liquidity risk and macroeconomic risks and liquidity risk [Chen, 2010, Longstaff et al., 2005, Elton et al., 2001, Collin-Dufresne et al., 2001]. What we do find is that this remainder is large, even when we account for heterogeneity of rating categories.

6.2 Macro-structural models, convexity and time-varying risk premia

Though our work is similar in spirit to that of Huang and Huang [2012], there is a major difference in our two works. They ignored heterogeneity of rating categories in the main part of their work. Later in their analysis, they studied the impact of the heterogeneity on their results. But as we argued earlier, the range of dispersion that they consider is small compared to the dispersion that we observe in the data. The heart of our work is to study in details the impact of full distribution of firms fundamentals inside rating categories on the measured default sizes. Thus our work tests the robustness of the conclusions of Huang and Huang [2012] to heterogeneity. David [2008] was amongst the firsts to show that heterogeneity of rating categories must be considered in the discussion of the credit spread puzzle. He argued that heterogeneity leads to a convexity bias in the test of structural models. As we have shown, there is indeed a convexity bias only if the typical firm is assumed to be the average firm, though its size is tiny for high quality bonds. But in many studies the typical firm is not the average firm. The parameters are carefully calibrated using various economic arguments. We showed that when the typical firm is not the average firm, one cannot sign the bias due to heterogeneity by convexity arguments. The bias can be even negative if the spread for the typical firm is significantly higher than the average spread. Our empirical investigations have demonstrated that once the asset volatility of the typical firm is calibrated to match historical default rates, the heterogeneity bias becomes negative in many cases. Thus, we conclude that the higher spreads obtained by David [2008] in their paper are not due to a convexity effect. Using a different argument line, Chen et al. [2009] also arrived at the same conclusion.

There are two ways of accounting for heterogeneity of rating categories. The first approach is to extend pure structural models by incorporating into it time-varying or cross-sectionally varying fundamentals (see David [2008], Chen et al. [2009], Bhamra et al. [2010]). The second is to consider that pure structural models must be tested on individual bonds. In this paper, it is this second approach that we pursue. Technically, both approaches rely on the assumption that it is the average model default rate of individual observations in a given rating category that must be compared with historical spreads. This assumption is less obvious in the first approach because the average default rate of future individual observations in a particular rating category represents ex-ante the expected default rate for a representative individual bond of that rating category. The problem with the first approach is that the model has to be modified and it can become puzzling to isolate the effect of heterogeneity from the effects of the new assumptions introduced in the models. The debate over whether the findings of David [2008] were indeed due to convexity or not is an illustration of "the puzzle in the puzzle" created by the first approach. In the second approach that we adopt, the original structure is preserved so that one can easily measure the impact of heterogeneity. The implications of our findings is that in the macro-structural models of David [2008], Chen et al. [2009] and Bhamra et al. [2010], it is not heterogeneity per se that leads to higher spreads. It is the additional features of these models mainly the fact that there exist states of the world where default risk can be very high and risk premia is also very high.

Our findings therefore supports the development of macro-structural models exploring the impact of time varying risk premia and stochastic macro-economic environments.

6.3 Assuming homogeneity might be warranted

A side implication of our work is that assuming homogeneity of rating categories does not necessarily bias downward the performance of pure structural models. If anything, this would most likely bias upward the tests. In the first part of the paper we have shown that rating categories are highly heterogeneous and that the distributions are highly overlapping from one category to the other. This suggests that it is not correct to consider that a given value of asset volatility or leverage or asset drift is typical or specific to a particular rating category. However, our findings also imply that for test purposes only, given the computational and modeling costs associated with introducing heterogeneity in the tests of structural models, adopting the homogeneity assumption might be judicious, provided that the typical firm is carefully calibrated. If the typical firm is the statistical average firm, there will be a systematic downward bias in the tests due to convexity, especially for low quality bonds. In some studies where one is concerned with the size of spread attributable to other factors such as liquidity, an upward bias might be equally undesirable as a downward bias. In such cases, it is advisable to consider the full consideration of fundamentals in rating categories to avoid any upward or downward bias.

6.4 It is all about risk adjustment

What our research and previous works on the credit spread point out is that there is something wrong with the type of risk-adjustment made in pure structural models. These models are based on the work of Black and Scholes [1973]. In the Black and Scholes [1973] derivation, a position in a call option can be perfectly hedged by a position in the underlying stock. The hedging argument is that if the stock price falls, the call price also falls and the value of a portfolio in both the stock and option with optimal weights will not change. What if due a change of macro-economic conditions, the stock price remains constant but the call price increases due to a greater risk of it being exercised? The portfolio is no more a perfect hedge. In the Black and Scholes [1973] world such a scenario cannot occur because changes in option value happen only because maturity is closer or because the stock price has changed. In other words, in the Black and Scholes [1973] and Merton [1974] framework, the systematic risk priced in bonds comes only from the comovement between the firm asset value and the option to default. The framework thus correctly adjusts for the fact that default risk is substantial when asset value is low. However, the framework ignores other sources of systematic risk such as macro-economic risk and liquidity. It ignores that default risk can be substantial when market returns are low, even if asset value is constant. By the same token, default risk can be substantial when liquidity conditions deteriorates, even if asset value remains constant. That is also ignored in the Black and Scholes [1973] and Merton [1974]. Our work therefore encourages the development of new structural models where these additional sources of systematic risk are taken into consideration.

7 conclusion

The rating through the cycle policy of rating agencies and the few number of rating categories makes rating categories heterogeneous. Many studies have found that accounting for that heterogeneity helps solve the credit spread puzzle. Others have argued of the opposite. What exactly is the role of heterogeneity in explaining corporate spreads? This is the question we sought to answer in this paper.

Three ingredients are necessary to answer the question: the true extent of the heterogeneity inside rating category, the calibration of the model to historical default experience in presence of heterogeneity and a methodology to clearly single out the impact heterogeneity on the measured spreads. Previous authors dealing with the issue have failed to deal with one or more of these aspects of the question. The present study brings together these three ingredients in an attempt to definitively answer the question.

We have found that asset volatilities, leverages, asset drifts and recovery rates are highly dispersed inside rating categories with large skewnesses and kurtoses. Furthermore, the distributions of these fundamentals are highly overlapping across rating categories. Notwithstanding this substantial heterogeneity, the credit spread puzzle is not a statistical artifact due to the heterogeneity of rating categories. After several robustness checks, we conclude that pure structural models can explain only a small fraction of observed spreads for investment grade bonds.

The problem lies in the type of risk adjustment made in structural models. When we do not require this risk adjustment to be consistent with historical default rates, structural models perform well. However, once we constraint the risk adjustment to reflect historical default rates, the model performs poorly for investment grade bonds.

We have also shown that assuming homogeneity does not systematically bias downward measured spreads through the so-called convexity effect. This only happens when the typical firm is the statistical average firm. When the typical firm is carefully calibrated with economic arguments, there is still a bias due to heterogeneity but we found that this bias is negative. The negative sign is not a concavity effect and it has nothing to do with the non-linearity of structural model. It simply reflects the fact that the average spread is smaller than the spread of the calibrated typical firm.

Overall, our findings support the development of structural models where investors require compensation for other sources risk such as macroeconomic risks, market risks and liquidity risk. They also suggest that despite the heterogeneity of rating categories, assuming their homogeneity might be judicious given the computational burden that might be associated with heterogeneity for a large sample size.

References

Harjoat S Bhamra, Lars-Alexander Kuehn, and Ilya A Strebulaev. The levered equity risk premium and credit spreads: A unified framework. *The Review of Financial Studies*, 23(2):645–, February 2010. ISSN 08939454. URL http://proxy2.hec.ca/login?url=http://search.proquest.com/docview/230030950?accountid=11357.

Sreedhar T Bharath and Tyler Shumway. Forecasting default with the merton distance to default model. *The Review of Financial Studies*, 21(3):1339–1369, May 2008. ISSN 08939454. URL http://proxy2.hec.ca/login?url=http://search.proquest.com/docview/230017066?accountid=11357.

Fischer Black and Myron Scholes. The pricing of options and corporate liabilities. *The Journal of Political Economy*, 81(3):637–, 1973. ISSN 00223808. URL http://proxy2.hec.ca/login?url=http://search.proquest.com/docview/195408552?accountid=11357.

Hui Chen. Macroeconomic conditions and the puzzles of credit spreads and capital structure. *The Journal of Finance*, 65(6):2171–2212, December 2010. ISSN 1540-6261. URL http://dx.doi.org/10.1111/j.1540-6261.2010.01613.x.

Long Chen, Pierre Collin-Dufresne, and Robert S. Goldstein. On the relation between the credit spread puzzle and the equity premium puzzle. *Review of Financial Studies*, 22(9):3367–3409, 2009. doi: 10.1093/rfs/hhn078. URL http://rfs.oxfordjournals.org/content/22/9/3367.abstract.

Pierre Collin-Dufresne, Robert S Goldstein, and J Spencer Martin. The determinants of credit spread changes. *The Journal of Finance*, 56(6):2177–2207, December 2001. ISSN 00221082. URL http://proxy2.hec.ca/login?url=http://search.proquest.com/docview/194716693?accountid=11357.

Peter J. Crosbie and Jeffrey R. Bohn. Modeling default risk. *KMV, LLC*, 2002.

Alexander David. Inflation uncertainty, asset valuations, and the credit spreads puzzle. *The Review of Financial Studies*, 21(6):2487–2534, November 2008. ISSN 08939454. URL http://proxy2.hec.ca/login?url=http://search.proquest.com/docview/230016968?accountid=11357.

Redouane Elkamhi, Jan Ericsson, and Christopher A Parsons. The cost and timing of financial distress. *Journal of Financial Economics*, 105(1):62–, July 2012. ISSN 0304405X. URL http://proxy2.hec.ca/login?url=http://search.proquest.com/docview/1020879189?accountid=11357.

Edwin J Elton, Martin J Gruber, Deepak Agrawal, and Christopher Mann. Explaining the rate spread on corporate bonds. *The Journal of Finance*, 56(1):247–277, February 2001. ISSN 00221082. URL http://proxy2.hec.ca/login?url=http://search.proquest.com/docview/194720318?accountid=11357.

Eugene F Fama and Kenneth R French. The cross-section of expected stock returns. *The Journal of Finance*, 47(2):427–, June 1992. ISSN 00221082. URL http://proxy2.hec.ca/login?url=http://search.proquest.com/docview/194708140?accountid=11357.

Peter Feldhutter and The Credit Spread Puzzle Myth or Reality? Schaefer, Stephen M. The credit spread puzzle - myth or reality? *Working Paper*, 2014. doi: http://dx.doi.org/10.2139/ssrn.2363081. URL http://ssrn.com/abstract=2363081.

Zhiguo He and Wei Xiong. Rollover risk and credit risk. *The Journal of Finance*, 67(2):391–, April 2012. ISSN 00221082. URL http://proxy2.hec.ca/login?url=http://search.proquest.com/docview/963343706?accountid=11357.

Jing-Zhi Huang and Ming Huang. How much of the corporate-treasury yield spread is due to credit risk? *Review of Asset Pricing Studies*, 2(2):153–202, 2012. doi: 10.1093/rapstu/ras011. URL `http://raps.oxfordjournals.org/content/2/2/153.abstract`.

Hull. *Risk Management and Financial Institutions*. Wiley Finance. Wiley, 3 edition, 2012. ISBN 9781118286388. URL `http://books.google.ca/books?id=MqdrF2R5eTIC`.

E Philip Jones, Scott P Mason, Eric Rosenfeld, and Lawrence Fisher. Contingent claims analysis of corporate capital structures: An empirical investigation/discussion. *The Journal of Finance*, 39(3):611–, July 1984. ISSN 00221082. URL `http://proxy2.hec.ca/login?url=http://search.proquest.com/docview/194703866?accountid=11357`.

Sean C. Keenan, Igor Shtogrin, and Jorge Sobehart. Historical default rate of corporate bond issuers, 1920-1998. *Moody's Investor Service, Special comments*, January 1999. URL `http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.195.2907&rep=rep1&type=pdf`.

Hayne E. Leland. Corporate debt value, bond covenants, and optimal capital structure. *The Journal of Finance*, 49(4):1213–1252, 1994. ISSN 1540-6261. doi: 10.1111/j.1540-6261.1994.tb02452.x. URL `http://dx.doi.org/10.1111/j.1540-6261.1994.tb02452.x`.

Hayne E. Leland. Predictions of default probabilities in structural models of debt. *Journal of Investment Management*, 2:5–20, 2004.

Hayne E Leland and Klaus Bjerre Toft. Optimal capital structure, endogenous bankruptcy, and the term structure of credit spreads. *The Journal of Finance*, 51(3):987–1019, July 1996. ISSN 00221082. URL `http://proxy2.hec.ca/login?url=http://search.proquest.com/docview/194710294?accountid=11357`.

Francis A Longstaff and Eduardo S Schwartz. A simple approach to valuing risky fixed and floating rate debt. *The Journal of Finance*, 50(3):789–, July 1995. ISSN 00221082. URL http://proxy2.hec.ca/login?url=http://search.proquest.com/docview/194713489?accountid=11357.

Francis A Longstaff, Sanjay Mithal, and Eric Neis. Corporate yield spreads: Default risk or liquidity? new evidence from the credit default swap market. *The Journal of Finance*, 60(5):2213–2253, October 2005. ISSN 00221082. URL http://proxy2.hec.ca/login?url=http://search.proquest.com/docview/194722304?accountid=11357.

Pierre Mella-Barral and William Perraudin. Strategic debt service. *The Journal of Finance*, 52(2):531–556, June 1997. ISSN 00221082. URL http://proxy2.hec.ca/login?url=http://search.proquest.com/docview/194715020?accountid=11357.

Robert C. Merton. On the pricing of corporate debt: The risk structure of interest rates*. *The Journal of Finance*, 29(2):449–470, 1974. ISSN 1540-6261. doi: 10.1111/j.1540-6261.1974.tb03058.x. URL http://dx.doi.org/10.1111/j.1540-6261.1974.tb03058.x.

Moody's. Corporate default and recovery rates, 1920-2010. 2011.

Charles R. Nelson and Andrew F. Siegel. Parsimonious modeling of yield curves. *Journal of Business*, 60(4):473–489, October 1987. ISSN 00219398. URL http://proxy2.hec.ca/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=ent&AN=4584556&lang=fr&site=ehost-live.

Stephen M. Schaefer and Ilya A. Strebulaev. Structural models of credit risk are useful: Evidence from hedge ratios on corporate bonds. *Journal of Financial Economics*, 90(1):1 – 19, 2008. ISSN 0304-405X. doi: http://dx.doi.

org/10.1016/j.jfineco.2007.10.006. URL http://www.sciencedirect.com/science/article/pii/S0304405X08001001.

Standard and Poors. Recovery study (u.s.): Are second liens and senior unsecured bonds losing ground as recoveries climb? 2013.

Standard and Global Fixed Income Research Poors. Default, transition and recovery: 2013 annual global corporate default study and rating transition. 2014. URL http://www.maalot.co.il/publications/FTS20140324161422.pdf.

A The Merton model with Bankruptcy costs

The model of Merton [1974] relies on the following assumptions. The market is frictionless: no taxes, no indivisibility problem, no transaction costs, no restrictions on short-selling. Securities are perfectly liquid. There exists a market where one can lend and borrow at the same riskless rate. The Modigliani-Miller theorem holds in such a way that the firm value is independent of the firm capital structure. Trading takes place in continuous time. The riskless rate is constant across all maturities. The firm value V follows a Gauss-Wiener Process:

$$dV = (\mu V - C)\, dt + \sigma V dz, \tag{1}$$

where μ is the instantaneous rate of return of the firm value, C is the firm instantaneous net payout to its stakeholders (interest payments + dividends), σ is the instantaneous standard deviation of the firm value and dz is the standard Gauss-Wiener process. In this tax free word, the firm value is equal to the value of its assets.

Following Black and Scholes [1973], Merton [1974] used no arbitrage arguments to show that the price of any security Y whose value could be written as $F(V,t)$ satisfies the following partial differential equation

$$\frac{1}{2}\sigma^2 V^2 F_{VV} + (\mu V - C)\, F_V + F_t = rF - C_y, \tag{2}$$

where the subscripts denote partial derivatives and C_y is the instantaneous net dollar payout to the security holder.

Consider now a firm with only two types of securities: equity and a zero-coupon debt of principal B due at date T. The firm cannot issue any new senior or equivalent debt, nor can it pay dividends or do share repurchases prior to the maturity of the existing debt. In the event of default, the bondholders immediately take over the company at no cost and the shareholders receive nothing. Thus $C = C_y = 0$, there is no default cost and the absolute priority rule is respected. In the original version of the Merton Model there are no bankruptcy costs. Here we introduce some bankruptcy costs in the form of a recovery rate ρ in case of default. The equity price is given by

$$\begin{aligned}
E &= e^{-rT}E^Q\left[(V_T - B)\times 1_{\{V_T>B\}}\right],\\
&= V_0Q^S\left(V_T > B\right) - Be^{-rT}Q\left(V_T > B\right),\\
&= V_0N\left(d_1\left(\lambda\right)\right) - Be^{-rT}N\left(d_2\left(\lambda\right)\right).
\end{aligned}$$

$$\begin{aligned}
d_1\left(x\right) &= \frac{-\ln\left(x\right) + \left(r + \frac{\sigma^2}{2}\right)T}{\sigma\sqrt{T}},\\
d_2\left(x\right) &= \frac{-\ln\left(x\right) + \left(r - \frac{\sigma^2}{2}\right)T}{\sigma\sqrt{T}},\\
\text{and } \lambda &= \frac{B}{V_0} \text{ is the leverage ratio.}
\end{aligned}$$

The bond price is given by :

$$\begin{aligned}
F &= e^{-rT}E^Q\left[B\times 1_{\{V_T>B\}} + \min\left(\rho B, V_T\right)\times 1_{\{V_T<B\}}\right],\\
&= Be^{-rT}Q\left(V_T > B\right) + e^{-rT}E^Q\left[V_T\times 1_{\{V_T<\rho B\}} + \rho B\times 1_{\{\rho B<V_T<B\}}\right],\\
&= Be^{-rT}N\left(d_2\left(\lambda\right)\right) + V_0N\left(-d_1\left(\rho\lambda\right)\right) + \rho Be^{-rT}Q\left(\rho B < V_T < B\right),\\
F &= Be^{-rT}N\left(d_2\left(\lambda\right)\right) + V_0N\left(-d_1\left(\rho\lambda\right)\right) + \rho Be^{-rT}\left(N\left(d_2\left(\lambda\right)\right) - N\left(d_2\left(\rho\lambda\right)\right)\right).
\end{aligned}$$

$$\frac{F}{B} = e^{-rT}N\left(d_2\left(\lambda\right)\right) + \frac{V_0}{B}N\left(-d_1\left(\rho\lambda\right)\right) + \rho e^{-rT}\left(N\left(d_2\left(\lambda\right)\right) - N\left(d_2\left(\rho\lambda\right)\right)\right).$$

The objective default rate is given by:

$$\Phi_T(\lambda, \sigma, \mu) = N\left(-\frac{-\ln(\lambda) + \left(\mu - \frac{\sigma^2}{2}\right)T}{\sigma\sqrt{T}}\right).$$

The spread is given by

$$\begin{aligned}\Psi_T(\lambda, \sigma, \rho, r) &= -\frac{1}{T}\ln\left(\frac{F}{B}\right) - r \\ &= -\frac{1}{T}\ln\left(e^{-rT}N(d_2(\lambda)) + \frac{1}{\lambda}N(-d_1(\rho\lambda)) + \rho e^{-rT}(N(d_2(\lambda)) - N(d_2(\rho\lambda)))\right) - r.\end{aligned}$$

B Computation of firm fundamentals

In this section, we discuss the calculation of firm fundamentals including leverage, asset value, asset volatility, asset drift and recovery rate.

B.1 Asset value, asset volatility and asset drift

A key input to structural models is the asset volatility which is not directly observable. One could estimate it from the time series of market value of assets, if such a times series were available. In the absence of taxes, the market value of assets can be taken to the sum of the market value of equity and the market value of debt. However, while market data on equity is available, only a small portion of total debt is traded on the market. Thus, computing the market value this way is not directly feasible. Several solutions to this problem exist in the literature. The easiest solution would be to proxy the market value of debt by its book value. The drawback to this solution is that market value of equity is observed daily while book value of debt is at best observed quarterly. Another solution suggested by Jones et al. [1984] is to assume that the market value of non-traded debt is equal to the book value of non-traded debt scaled by the ratio of book to market of traded debt. Implicit to this solution, is the assumption that the book to market ratio of traded debt and non-traded debt are similar. However, there is no reason to believe that this assumption is true. Moreover, to implement this solution one must be able to disentangle book value of traded debt from book value of non traded debt. Such disentangled data is not readily available. A third solution is to estimate separately equity return volatility using time series of equity returns and debt return volatility using available data on publicly traded debt. The estimate of the asset volatility in this third approach would be a weighted average of the equity return volatility and the debt return volatility[3]. The drawback of this approach is that the weights are often chosen arbitrarily with no economic argument.

A fourth approach which relies on the Merton [1974] model exists. It consists of solving a system of two equations as described in Crosbie and Bohn [2002] for the implied volatility of asset returns. Indeed, the value of equity E is

$$E = VN(d_1) - Be^{-r\tau}N(d_2), \tag{3}$$

and the volatility of equity returns σ_E is

$$\sigma_E = \sigma E_V \frac{V}{E} = \sigma N(d_1) \frac{V}{E}. \tag{4}$$

If the equity value E, the equity volatility σ_E, the horizon T, the debt principal B and the riskless rate r are known, one can solve simultaneously and numerically equations (3) and (4) for the asset volatility σ and the asset value V. In practice, one must first choose an horizon T. Here we use $T = 4$, in such a way that the assumption that debt principal is equal to short term debt by half long term debt is plausible. Second, the equity value E and the riskless rate are observed on the market and historical return on equity is used to estimate

[3]Schaefer and Strebulaev [2008] proposed an extended version of the weighted average approach where it is possible to take into consideration the correlation between bond returns and equity returns.

equity volatility σ_E.

According to Crosbie and Bohn [2002], the asset volatility obtained from the simultaneous solution of equations (3) and (4) is not robust to changes in the firm leverage. So they suggested an alternative complicated iterative procedure that is better described by Bharath and Shumway [2008]. The procedure is the following. One first guesses a value of σ. Equation (3) is then used to find the corresponding V for every day of the previous four years. The resulting asset returns are used to estimate a new asset volatility σ which becomes the input of the next iteration. Bharath and Shumway [2008] suggested to use $\sigma_E \frac{E}{E+B}$ as a starting value of σ. We adopt this iterative procedure to calculate daily asset values and yearly asset volatilities. We assume that convergence is achieved either when the norm of the difference between two consecutive vectors of daily asset values is less than 0.001. Crosbie and Bohn [2002] argued that this procedure converges in few iterations. In practice, in few cases, convergence may be slow. To avoid infinite loops, we stop the iteration procedure if the number of iterations exceeds 100, and we discard the corresponding observation. We measure the firm drift by the annualized average daily asset returns calculated from the above iterative procedure.

B.2 Leverage and calibration to historical default rates

Leverage is calculated as the ratio of debt principal to the initial asset value. Many studies make the assumption that the debt principal is the same for every maturity. This lead to the same leverage ratio for all considered horizons. This hardly makes economic sense because only a small proportion of the debt is due in short-term. The asset value needs only to exceed this short-term debt in order to avoid default in the short term. If one assumes that the whole debt is due in the short term, then one would overestimate the likelihood of a default in the short term. In our study we consider four horizons 1, 2, 4 and 10 year. We do not know the exact proportion of the total debt that is due at each of these maturities. We decided to recover this information from the historical default rates. For each rating and maturity, we required that the debt principal due is the amount of debt that would make the model average historical default rate equal to the historical default rate. Our results are however independent of this choice.

In practice, for each rating category and maturity, we solve the following equation for x:

$$\frac{1}{n}\sum_{i=1}^{n}\Phi_T\left(\lambda_i e^{-x},\sigma_i,\mu_i\right)=\overline{\phi}_T,$$

which is equivalent to solving:

$$\frac{1}{n}\sum_{i=1}^{n}N\left(-\frac{-\ln\left(\lambda_i\right)+\left(\mu_i-\frac{\sigma_i^2}{2}\right)T}{\sigma_i\sqrt{T}}-\frac{x}{\sigma_i\sqrt{T}}\right)=\overline{\phi}_T.$$

n is the number of observation in that rating category. The index i is for individual observations in that rating category.

Thus our calibration procedure is similar in spirit to that of Moody's KMV. We make an adjustment to the distance to default implied by the model to match historical default rates. The adjustment that we make is proportional to asset volatility and to maturity. A lower leverage is necessary to compensate for a too high asset volatility and a longer maturity that both increase the risk of default. The adjustment is uniform for all observations in the rating category. This is done to guarantee a unique solution since we only have one calibrating equation. The adjusted leverage is equal to $\lambda_i e^{-x}$. We could have chosen to adjust asset volatilities instead of leverage. However, our results are not dependent on our choice to adjust the leverage instead of the asset volatility. We repeated the study by applying the adjustment to asset volatility instead of leverage and we found similar quantitative results.

B.3 Recovery rates

Moody's annual reports provide average recovery rates by economic sector $\overline{\rho}^k, k \in K$ and average recovery rates by rating $\overline{\rho}^l, l \in L$. K and L are the sets of economic sectors and rating categories respectively. However, these reports do not provide average recovery rates stratified by rating and economic sector simultaneously ρ_{kl}. Using $\overline{\rho}^k$ and $\overline{\rho}^l$, we suggest a way to generate heterogeneity inside rating categories.

We assume that the recovery rate of a firm in a given economic sector with a given rating is equal to average recovery rate inside that economic sector scaled by an adjustment factor which depends of the firm rating. That is ρ_{kl} has the following form:

$$\rho_{kl} = \alpha_l \overline{\rho}^k.$$

With this form, two firms with the same rating but different economic sector will have different recovery rates.

To identify α_l, we require that the average recovery rate inside a given rating category in our sample matches the historical average reported by the rating agencies.

Let Ω^k be the set of all observations falling in sector $k, k \in K$, in our sample and let Ω^l be the set of all observations falling in rating category $l, l \in L$ in our sample. The identification condition is:

$$\sum_{k \in K} \frac{card\left(\Omega^k \cap \Omega^l\right)}{card\left(\Omega^l\right)} \alpha_l \overline{\rho}^k = \overline{\rho}^l, l \in L.$$

This implies that:

$$\alpha_l = \frac{\overline{\rho}^l}{\sum_{k \in K} \frac{card(\Omega^k \cap \Omega^l)}{card(\Omega^l)} \overline{\rho}^k}, l \in L.$$

C The model of Leland and Toft [1996]

The optimal default boundary is given by:

$$V_B = \frac{\frac{C}{r}\left(\frac{A}{rT} - B\right) - \frac{AP}{rT} - \frac{\tau C x}{r}}{1 + \alpha x - (1-\alpha) B},$$

Where:

C is the total coupon payment,

r is the risk free rate,

T is the maturity of new issues,

P is the total outstanding debt principal,

τ is the tax rate,

α is the bankruptcy cost parameter as a proportion of asset value at default,

$$\begin{aligned} A &= 2ae^{-rT} N\left(a\sigma\sqrt{T}\right) - 2zN\left(z\sigma\sqrt{T}\right) \\ &\quad - \frac{2}{\sigma\sqrt{T}} n\left(z\sigma\sqrt{T}\right) + \frac{2e^{-rT}}{\sigma\sqrt{T}} n\left(a\sigma\sqrt{T}\right) + (z-a), \end{aligned}$$

$$B = -\left(2z + \frac{2}{z\sigma^2 T}\right) N\left(z\sigma\sqrt{T}\right) - \frac{2}{\sigma\sqrt{T}} n\left(z\sigma\sqrt{T}\right) + (z-a) + \frac{1}{z\sigma^2 T},$$

$$\begin{aligned} z &= \frac{\left(\left(a\sigma^2\right)^2 + 2r\sigma^2\right)^{1/2}}{\sigma^2}, \\ a &= \frac{r - \delta - \frac{\sigma^2}{2}}{\sigma^2}, \end{aligned}$$

$$x = z + a,$$

σ is the asset volatility,

δ is the asset payout rate,

$N(.)$ is the cumulative standard normal distribution,

$n(.)$ is the standard normal density function.

The price of a new issue of maturity T is given by:

$$d(V, V_B, T) = \frac{c}{r} + e^{-rT}\left(p - \frac{c}{r}\right)(1 - F(T)) + \left(\rho V_B - \frac{c}{r}\right)G(T),$$

$$F(T) = N(h_1(T)) + \left(\frac{V}{V_B}\right)^{-2a} N(h_2(T)),$$

$$G(T) = \left(\frac{V}{V_B}\right)^{-a+z} N(q_1(T)) + \left(\frac{V}{V_B}\right)^{-a-z} N(q_2(T)),$$

$$q_1(T) = \frac{-b - z\sigma^2 T}{\sigma\sqrt{T}},$$

$$q_2(T) = \frac{-b + z\sigma^2 T}{\sigma\sqrt{T}},$$

$$h_1(T) = \frac{-b - a\sigma^2 T}{\sigma\sqrt{T}},$$

$$h_2(T) = \frac{-b + a\sigma^2 T}{\sigma\sqrt{T}},$$

$$b = \ln\left(\frac{V}{V_B}\right),$$

c is the coupon rate on the new issue : $c = \frac{C}{T}$,

p is the principal of the new issue: $p = \frac{P}{T}$,

$\rho = 1 - \alpha$.

The cumulative physical probability of default at time t is given by

$$\Phi_T\left(P, \alpha, \tau, V, \sigma, \delta, r, \mu\right) = N\left(\frac{-b - \lambda t}{\sigma\sqrt{t}}\right) + e^{\frac{-2\lambda b}{\sigma^2}} N\left(\frac{-b + \lambda t}{\sigma\sqrt{t}}\right),$$

$$\lambda = \mu - \delta - \frac{\sigma^2}{2}.$$

In the original model, the coupon rates are determined in such a way the new issues sell at par. This involves solving numerically the following equation for the coupon rate.

$$d\left(V, V_B, T\right) = p.$$

The spread on new issues is then computed as

$$\Psi_T\left(P, \alpha, \tau, V, \sigma, \delta, r\right) = \frac{c}{p} - r.$$

Calibration

When we calibrate to historical default rates, we solve the following equation for ϵ:

$$\frac{1}{n}\sum_{i=1}^{n} \Phi_T\left(\exp\left(\ln\left(P_i\right) + \epsilon\right), \alpha_i, \tau_i, V_i, \sigma_i, \delta_i, r, \mu_i\right) = \phi_T.$$

Following Elkamhi et al. [2012] we choose:

$$C = rP.$$

We determine the yield y on new issues by solving the classical discounting equation.

$$c\frac{1-e^{yT}}{y}+pe^{-yT}=d\left(V,V_B,T\right).$$

The yield spread is given by

$$\Psi_T\left(P,\alpha,\tau,\sigma,\delta,r\right)=y-r.$$

Following Leland [2004] we choose $\tau=15\%$, $\delta=6\%$, $\alpha=30\%$. The other parameters are the same as in the Merton model.

Table 1: Typical leverage, asset volatility and bankruptcy cost for the Baa rating category

	HH -LT	HH-MP	HH-LS	L-LT	LT	SS
Leverage	43.28%	43.28%	43.28%	43.30%	49%	36%
Asset volatility	25.05%	16.11%	29.10%	23%	20%	22%
Bankruptcy cost	87.29%	N.A.	N.A.	30%	50%	N.A.

HH-LT is the Leland and Toft [1996] model calibrated by Huang and Huang [2012]. HH-MP is the Mella-Barral and Perraudin [1997] model calibrated by Huang and Huang [2012]. HH-LS is the Longstaff and Schwartz [1995] model calibrated by Huang and Huang [2012]. L-LT is the Leland and Toft [1996] model calibrated by Leland [2004]. LT is the Leland and Toft [1996] as calibrated in their paper. SS is the estimation of Schaefer and Strebulaev [2008].

Table 2: Number of firms and observations by rating category and economic sector

Sector	AAA	AA	A	BBB	BB	B	Below B	Total
Panel A: Number of firms by rating category and economic sector								
Transportation	1	2	6	20	22	25	13	48
Utilities	0	14	62	81	27	13	4	120
Health care	5	10	23	29	43	31	2	94
Capital Goods	1	10	39	75	100	46	14	180
Energy	1	3	15	41	45	30	4	96
Technology	2	9	23	41	53	45	11	101
Basic Materials	0	6	39	69	63	54	20	148
Communication services	1	5	12	7	16	21	9	43
Consumer cyclicals	1	11	53	113	169	159	55	290
Consumer Staples	2	17	50	63	66	60	27	166
Total	14	87	322	539	604	484	159	1286
Panel B: Number of firms and observations by economic sector								
Transportation	1037	1771	10401	41343	25708	23531	4797	108588
Utilities	0	19219	87189	153839	22403	8959	475	292084
Health care	15180	17432	44691	39793	35817	20733	95	173741
Capital Goods	720	21014	90153	133967	121065	42107	2449	411475
Energy	2236	5361	35049	67866	60227	22651	384	193774
Technology	3561	15598	47182	48530	48090	30866	1961	195788
Basic Materials	0	14663	88234	134157	75762	51307	3772	367895
Communication services	2193	1650	15957	9722	6798	18722	2336	57378
Consumer cyclicals	436	16891	104274	154378	180342	105702	11927	573950
Consumer Staples	1945	35792	95620	100187	64604	56560	6492	361200
Total	27308	149391	618750	883782	640816	381138	34688	2735873

Note about the last column of Panel A: The total exclude duplicates due to changing firm ratings.

Table 3: Historical default rates (%)

Rating	Maturity in years			
	1	2	4	10
AAA	0	0.03	0.24	0.74
AA	0.02	0.07	0.24	0.84
A	0.07	0.17	0.43	1.59
BBB	0.21	0.6	1.53	4.33
BB	0.8	2.46	6.29	14.39
B	4.11	9.27	16.99	26.97
Below B	26.87	36.05	44.27	51.35

Source: Standard and Poors [2014]

Table 4: Average recovery rates by rating categories and maturity

Rating	number of years prior to default			
	1	2	4	10
AAA	0.3724	0.4015	0.5043	0.3880
AA	0.3724	0.4015	0.5043	0.3880
A	0.3177	0.4756	0.3990	0.4182
BBB	0.4147	0.4302	0.4457	0.4269
BB	0.4711	0.4461	0.4081	0.4080
B	0.3790	0.3606	0.3806	0.4135
Below B	0.3550	0.3481	0.3533	0.3496

Note: Due to extremely small sample sizes in the AAA category, we assume the same recovery rates as AA firms. Due to data unavailability at the 10 year horizon, we use the 5 year horizon recovery rate as a proxy. Source: Moody's [2011]

Table 5: Average recovery rates by economic sectors

Sector	recovery rate
Basic Materials	0.353
Capital goods	0.353
Communication services	0.391
Consumer Staples	0.336
Consumer Cyclicals	0.336
Energy	0.461
Health Care	0.347
Technology	0.328
Transportation	0.379
Utilities	0.641
Source: Standard and Poors [2013]	

Table 6: Heterogeneous recovery rates across economic sectors

sector	AAA	AA	A	BBB	BB	B	Below B	AAA	AA	A	BBB	BB	B	Below B
	T=1							T=2						
Transportation	39.45%	36.63%	30.62%	38.66%	48.79%	39.89%	38.02%	42.53%	39.49%	45.83%	40.11%	46.20%	37.96%	37.29%
Utilities		61.95%	51.78%	65.39%	82.52%	67.47%	64.31%		66.79%	77.52%	67.83%	78.14%	64.19%	63.06%
Health care	36.12%	33.53%	28.03%	35.40%	44.67%	36.52%	34.81%	38.94%	36.15%	41.96%	36.72%	42.30%	34.75%	34.14%
Capital Goods	36.74%	34.11%	28.52%	36.01%	45.44%	37.16%	35.42%	39.61%	36.78%	42.69%	37.36%	43.03%	35.35%	34.73%
Energy	47.98%	44.55%	37.24%	47.03%	59.35%	48.52%	46.25%	51.73%	48.03%	55.75%	48.79%	56.20%	46.17%	45.35%
Technology	34.14%	31.70%	26.50%	33.46%	42.22%	34.52%	32.91%	36.81%	34.17%	39.67%	34.71%	39.98%	32.85%	32.27%
Basic Materials		34.11%	28.52%	36.01%	45.44%	37.16%	35.42%		36.78%	42.69%	37.36%	43.03%	35.35%	34.73%
Communication services	40.69%	37.79%	31.59%	39.89%	50.33%	41.16%	39.23%	43.87%	40.74%	47.28%	41.38%	47.66%	39.16%	38.47%
Consumer cyclicals	34.97%	32.47%	27.14%	34.28%	43.25%	35.37%	33.71%	37.70%	35.01%	40.63%	35.56%	40.96%	33.65%	33.06%
Consumer Staples	34.97%	32.47%	27.14%	34.28%	43.25%	35.37%	33.71%	37.70%	35.01%	40.63%	35.56%	40.96%	33.65%	33.06%
	T=4							T=10						
Transportation	53.42%	49.60%	38.45%	41.55%	42.26%	40.06%	37.84%	41.10%	38.16%	40.30%	39.80%	42.25%	43.52%	37.45%
Utilities		83.89%	65.03%	70.28%	71.48%	67.75%	64.00%		64.54%	68.16%	67.31%	71.46%	73.61%	63.33%
Health care	48.91%	45.41%	35.20%	38.04%	38.70%	36.68%	34.65%	37.63%	34.94%	36.90%	36.44%	38.69%	39.85%	34.28%
Capital Goods	49.75%	46.20%	35.81%	38.70%	39.37%	37.31%	35.25%	38.28%	35.54%	37.54%	37.07%	39.36%	40.54%	34.88%
Energy	64.97%	60.33%	46.77%	50.54%	51.41%	48.73%	46.03%	49.99%	46.42%	49.02%	48.41%	51.40%	52.94%	45.55%
Technology	46.23%	42.92%	33.28%	35.96%	36.58%	34.67%	32.75%	35.57%	33.03%	34.88%	34.44%	36.57%	37.67%	32.41%
Basic Materials		46.20%	35.81%	38.70%	39.37%	37.31%	35.25%		35.54%	37.54%	37.07%	39.36%	40.54%	34.88%
Communication services	55.11%	51.17%	39.67%	42.87%	43.60%	41.33%	39.04%	42.40%	39.37%	41.58%	41.06%	43.59%	44.90%	38.63%
Consumer cyclicals	47.36%	43.97%	34.09%	36.84%	37.47%	35.52%	33.55%	36.44%	33.83%	35.73%	35.28%	37.46%	38.59%	33.20%
Consumer Staples	47.36%	43.97%	34.09%	36.84%	37.47%	35.52%	33.55%	36.44%	33.83%	35.73%	35.28%	37.46%	38.59%	33.20%

Table 7: Distribution of asset volatility and asset drifts by rating category

Rating	mean	sd	CV	range	p99-p1	skewness	kurtosis	min	p1	p25	p50	p75	p99	max
Panel A: Asset Volatility σ_V (in percent per annum)														
AAA	25.08	6.67	0.27	45.31	35.42	1.56	7.93	13.22	13.84	20.61	24.19	28.32	49.26	58.52
AA	25.39	6.98	0.27	57.41	34.87	0.89	4.73	11.47	12.13	20.39	24.42	29.53	47.00	68.88
A	27.73	10.92	0.39	111.10	54.16	3.15	22.17	11.47	12.51	21.08	26.00	31.69	66.67	122.58
BBB	31.32	19.43	0.62	351.73	73.38	9.31	136.01	11.47	12.47	22.64	28.68	36.16	85.84	363.21
BB	40.31	19.21	0.48	313.75	101.05	4.26	39.05	11.48	15.76	29.36	36.83	46.37	116.81	325.23
B	49.71	28.87	0.58	398.54	196.34	4.69	38.33	11.82	16.67	34.63	44.18	56.23	213.01	410.37
Below B	65.34	53.99	0.83	391.93	286.66	3.60	17.63	18.49	19.43	40.00	52.33	67.70	306.09	410.43
Panel B: Asset drift μ_V (in percent per annum)														
AAA	14.24	12.62	0.89	77.35	60.79	0.46	3.21	-17.59	-12.00	5.16	13.11	21.92	48.78	59.77
AA	12.51	11.30	0.90	101.28	54.41	0.79	4.09	-20.63	-9.17	5.04	10.67	18.56	45.24	80.65
A	12.61	12.87	1.02	137.66	66.96	1.60	9.29	-28.16	-11.23	4.75	10.90	18.42	55.73	109.50
BBB	13.03	17.10	1.31	249.11	83.83	3.03	25.65	-28.19	-15.74	3.41	10.48	19.50	68.09	220.93
BB	17.54	20.97	1.20	249.07	105.98	1.43	9.40	-28.19	-20.95	3.65	15.05	28.07	85.03	220.88
B	13.95	24.44	1.75	246.23	115.82	1.82	10.52	-28.19	-24.56	-2.25	9.61	25.40	91.26	218.04
Below B	8.10	34.56	4.27	221.78	184.43	2.48	10.13	-28.19	-27.18	-12.09	-1.29	14.03	157.25	193.58
Panel C: Total leverage (in percent per annum)														
AAA	8.76	8.49	0.97	64.36	38.44	1.80	6.81	0.10	0.14	2.90	6.41	11.39	38.58	64.46
AA	16.42	13.59	0.83	99.69	65.90	1.62	5.94	0.23	1.01	7.20	11.97	20.84	66.91	99.93
A	26.44	19.05	0.72	269.62	81.48	1.16	4.59	0.05	1.07	12.21	21.14	37.16	82.56	269.67
BBB	40.10	29.81	0.74	1044.56	125.14	4.88	95.13	0.13	2.82	19.85	33.82	54.69	127.96	1044.69
BB	58.95	47.91	0.81	946.99	221.05	2.64	19.40	0.13	4.02	26.97	46.11	78.02	225.06	947.12
B	107.62	85.82	0.80	1519.63	400.07	3.04	23.88	1.62	8.20	50.16	87.24	142.07	408.27	1521.26
Below B	219.11	261.44	1.19	2695.82	1590.32	5.00	36.73	3.67	14.01	89.23	151.48	263.80	1604.33	2699.48

This table shows summary statistics and quantiles of firm fundamentals across S&P credit ratings. CV stands for Coefficient of Variation and is defined as the ratio of the standard deviation to the mean. Range is the maximum minus the minimum. Panels A, B and B display statistics for asset volatility, asset drift and total leverage respectively. The first column presents the rating categories considered. μ_V, the drift of asset returns (in percent per annum), is calculated for each firm and year as the mean of daily asset returns multiplied by 252. σ_V, the instantaneous volatility of asset returns (in percent per annum), is calculated for each firm as the standard deviation of the daily asset return multiplied by the square root of 252. See appendix B for details about the estimation procedure. Total leverage is short term debt plus long term debt divided by asset value.

Table 8: Comparison of calibrated Vs uncalibrated default sizes

Rating	Spread (basis points)		Historical spread (basis points)		Default size		Default sizes of Huang and Huang [2012]	
	T=4	T= 10	T=4	T= 10	T=4	T= 10	T=4	T= 10
	Panel A: Uncalibrated spreads and default sizes							
AAA	0	3	55	63	0.8%	4.7%	2.1%	15.8%
AA	9	15	65	91	13.2%	16.1%	9.2%	15.6%
A	53	56	96	123	55.4%	45.4%	10.3%	19.0%
BBB	215	178	158	194	136.0%	91.8%	20.3%	29.0%
BB	604	396	320	320	188.6%	123.6%	53.9%	60.0%
B	1474	841	470	470	313.7%	179.0%	94.8%	82.5%
	Panel B: Calibrated spreads and default sizes							
AAA	1	1	55	63	1%	2%	2.1%	15.8%
AA	5	5	65	91	8%	5%	9.2%	15.6%
A	11	14	96	123	11%	11%	10.3%	19.0%
BBB	62	66	158	194	39%	34%	20.3%	29.0%
BB	159	185	320	320	50%	58%	53.9%	60.0%
B	447	376	470	470	95%	80%	94.8%	82.5%

Table 9: Comparison of uncalibrated model physical default rates and historical default rates

Rating	Uncalibrated model DP		Historical DP		Ratio	
	T=4	T= 10	T=4	T= 10	T=4	T= 10
AAA	0.20%	1.41%	0.24%	0.74%	0.9	1.9
AA	0.54%	2.51%	0.24%	0.84%	2.3	3.0
A	2.76%	6.55%	0.43%	1.59%	6.4	4.1
BBB	9.92%	15.64%	1.53%	4.33%	6.5	3.6
BB	21.99%	26.20%	6.29%	14.39%	3.5	1.8
B	47.03%	49.68%	16.99%	26.97%	2.8	1.8

Table 10: Distribution of calibrated leverage ratios by rating category

Rating	mean	sd	CV	range	p99-p1	skewness	kurtosis	min	p1	p25	p50	p75	p99	max
Panel A: Leverage at 1 year maturity														
AAA	7.49	6.99	0.93	62.00	35.95	2.05	9.75	0.09	0.11	2.53	5.73	10.30	36.06	62.09
AA	11.77	9.91	0.84	92.18	45.34	1.82	7.57	0.23	0.76	5.01	8.87	15.21	46.09	92.41
A	15.68	11.90	0.76	155.00	53.15	1.43	5.77	0.05	0.62	7.05	12.22	21.54	53.77	155.06
BBB	8.53	6.54	0.77	219.74	27.58	5.19	98.42	0.03	0.68	4.17	7.14	11.58	28.26	219.76
BB	15.00	12.50	0.83	275.96	59.46	2.88	24.48	0.05	1.03	6.73	11.52	19.83	60.49	276.00
B	23.81	19.17	0.81	536.31	89.68	3.09	25.34	0.37	1.81	11.01	19.21	31.42	91.49	536.68
Below B	62.56	74.91	1.20	1078.50	400.00	4.56	35.52	1.21	3.26	23.31	40.80	74.79	403.26	1079.70
Panel B: Leverage at 2 year maturity														
AAA	8.45	7.88	0.93	69.91	40.53	2.05	9.75	0.10	0.12	2.85	6.46	11.61	40.65	70.01
AA	12.62	10.62	0.84	98.87	48.63	1.82	7.57	0.24	0.81	5.37	9.52	16.32	49.44	99.11
A	14.82	11.25	0.76	146.50	50.23	1.43	5.77	0.05	0.59	6.67	11.55	20.36	50.82	146.55
BBB	13.44	10.31	0.77	346.09	43.43	5.19	98.42	0.04	1.07	6.56	11.25	18.25	44.51	346.13
BB	18.29	15.24	0.83	336.39	72.48	2.88	24.48	0.05	1.26	8.20	14.04	24.17	73.74	336.45
B	27.35	22.03	0.81	616.10	103.02	3.09	25.34	0.42	2.08	12.64	22.07	36.10	105.10	616.52
Below B	65.89	78.89	1.20	1135.80	421.25	4.56	35.52	1.28	3.43	24.55	42.96	78.76	424.68	1137.10
Panel C: Leverage at 4 year maturity														
AAA	8.99	8.39	0.93	74.40	43.13	2.05	9.75	0.11	0.13	3.04	6.87	12.36	43.27	74.51
AA	12.36	10.40	0.84	96.80	47.61	1.82	7.57	0.24	0.79	5.26	9.32	15.98	48.40	97.04
A	12.95	9.83	0.76	128.03	43.90	1.43	5.77	0.04	0.51	5.83	10.09	17.79	44.41	128.08
BBB	14.14	10.84	0.77	364.04	45.69	5.19	98.42	0.04	1.13	6.90	11.83	19.19	46.82	364.08
BB	20.75	17.29	0.83	381.69	82.24	2.88	24.48	0.06	1.43	9.30	15.93	27.42	83.67	381.75
B	28.69	23.10	0.81	646.25	108.06	3.09	25.34	0.44	2.18	13.26	23.15	37.86	110.24	646.70
Below B	59.80	71.60	1.20	1030.90	382.33	4.56	35.52	1.16	3.11	22.28	38.99	71.49	385.45	1032.00
Panel C: Leverage at 10 year maturity														
AAA	5.69	5.30	0.93	47.06	27.28	2.05	9.75	0.07	0.08	1.92	4.35	7.81	27.37	47.13
AA	8.45	7.11	0.84	66.17	32.55	1.82	7.57	0.16	0.54	3.59	6.37	10.92	33.09	66.34
A	9.27	7.04	0.76	91.62	31.41	1.43	5.77	0.03	0.37	4.17	7.22	12.73	31.78	91.65
BBB	10.43	7.99	0.77	268.40	33.68	5.19	98.42	0.03	0.83	5.09	8.72	14.15	34.52	268.44
BB	21.65	18.04	0.83	398.23	85.80	2.88	24.48	0.07	1.49	9.71	16.62	28.61	87.29	398.29
B	22.63	18.22	0.81	509.65	85.22	3.09	25.34	0.35	1.72	10.46	18.26	29.86	86.94	510.00
Below B	31.43	37.63	1.20	541.75	200.92	4.56	35.52	0.61	1.64	11.71	20.49	37.57	202.56	542.36

This table shows summary statistics and quantiles of leverage ratios across S&P credit ratings at different maturities. CV stands for Coefficient of Variation and is defined as the ratio of the standard deviation to the mean. Range is the maximum minus the minimum. For each maturity and rating, the leverage ratio is calibrated to match the historical default rate report by S&P for this rating and maturity. See appendix B for details on the calculation and calibration of leverage.

Table 11: Convexity bias

	Maturity (Years)			
Rating	1	2	4	10
Panel A: Homogeneous recovery				
AAA	0	0	1	1
AA	4	6	5	5
A	6	7	10	12
BBB	34	52	59	56
BB	75	109	93	46
B	362	266	132	98
below B	353	96	172	302
Panel B: Heterogeneous recovery				
AAA	0	0	1	1
AA	4	6	5	5
A	7	7	11	12
BBB	34	52	60	57
BB	77	110	94	47
B	359	265	132	99
below B	358	98	173	303

Table 12: Heterogeneity bias

	Maturity (Years)			
Rating	1	2	4	10
Panel A: Homogeneous recovery				
AAA	0	-2	-9	-24
AA	2	1	-4	-19
A	-2	-3	-9	-28
BBB	16	19	5	-29
BB	5	-30	-94	-154
B	33	-22	-55	-61
below B	635	468	378	335
Panel B: Heterogeneous recovery				
AAA	0	-2	-9	-24
AA	2	1	-4	-19
A	-1	-2	-9	-28
BBB	16	19	6	-29
BB	6	-28	-93	-154
B	30	-23	-55	-61
below B	640	470	379	336

Table 13: Typical firm: Calibrated and uncalibrated asset volatilities

	Calibrated asset volatility to historical default rates(%)				
	Maturity (years)				
Rating	1	2	4	10	Average asset volatility(%)
AAA	25.08	51.35	45.47	43.43	25.08
AA	59.05	46.55	40.21	39.11	25.39
A	56.88	46.87	41.83	41.14	27.73
BBB	79.48	55.23	47.05	46.13	31.32
BB	74.51	60.52	54.72	54.12	40.31
B	74.59	62.96	58.63	58.95	49.71
Below B	59.99	54.93	57.93	63.81	65.34

Table 14: Robustness checks

rating	H-H	base case	case 1	case 2	case 3	case 4	case 5	case 6	case 7
					Maturity = 4 years				
AAA	2%	1%	2%	1%	0%	4%	4%	0%	1%
AA	9%	8%	11%	9%	5%	8%	6%	3%	7%
A	10%	11%	15%	11%	11%	9%	3%	11%	8%
BBB	20%	39%	47%	39%	38%	22%	3%	56%	22%
BB	54%	50%	72%	51%	109%	34%	16%	120%	39%
B	95%	95%	143%	98%	140%	55%	55%	164%	73%
					Maturity = 10 years				
AAA	16%	2%	3%	2%	7%	2%	0%	0%	1%
AA	16%	5%	8%	6%	19%	5%	1%	3%	5%
A	19%	12%	16%	12%	37%	6%	0%	11%	9%
BBB	29%	34%	43%	34%	74%	12%	0%	42%	14%
BB	60%	58%	95%	59%	151%	48%	30%	94%	44%
B	83%	80%	131%	82%	136%	50%	43%	114%	53%

This table shows results of our robustness checks. Column 2 labeled H-H reports the base case default size of Huang and Huang [2012]. Column 3 reports our base case default size with heterogeneity. Case 1: Average default probability is increased by 50%. Case 2: Heterogeneity in recovery rates is increased by attributing truncated normally distributed random recovery rates to firms inside rating category while matching average historical recovery rates. Case 3: The model default rate for each bond is calibrated to match the historical default rate for the corresponding rating category and horizon. Case 4: Leverages are fixed to the sample average and asset volatilities are adjusted to match historical default rates. Case 5: Leverages are fixed to the same average as Huang and Huang [2012] and asset volatilities are adjusted to match historical default rates. Case 6: The sample period is restricted to 1985-1998, the same as Huang and Huang [2012]. Case 7: Biased free default sizes based on the Leland and Toft [1996] model.