Submission for the:

Q-Group Jack Treynor Prize

“The Credit Spread Puzzle in the Merton Model – Myth or Reality?

Peter Feldhutter and Stephen Schaefer
London Business School

Overview and Relevance to Q-Group Sponsors and Mission

To insurance companies and pension funds as well as many other major investment managers, corporate debt is an asset class is of huge importance. The value of corporate (including financial sector) debt outstanding in the US is the same order of magnitude as the value of the US stocks. (The value of non-financial corporate debt is about a third of this value). Yet, despite its importance and a large amount of recent research, the pricing of corporate debt remains a controversial topic.

A common view among both academics and practitioners is that corporate bond spreads are too high relative to the value and risk of the corporate assets that collateralize the debt. This is the “credit spread puzzle” and, if true, would mean that investors obtain a higher Sharpe ratio from corporate debt than from equities.

Despite employing the same (Merton) model that many other researchers have used, our paper reaches a different conclusion: for the most part, corporate debt is fairly priced relative to equity. There are two main reasons why our results are different. First, several earlier papers have based their analysis on a “representative firm” – one with average leverage and average asset risk – and this leads to a bias that we correct. Second, and more importantly, we benchmark our model to estimates of average default rates obtained from a much longer (90 year) period than used in previous studies. Long-run average default rates are higher than those over the past 30-40 years that much of the academic research has used in benchmarking. Also, even over periods as long as 30 years, the distribution of realized default rates is highly skewed with the result that the observed average default rate is likely to be significantly below the long-run mean. Using long-run default rates we find that actual market spreads are in line with the long-run average. This means that Sharpe ratios in the corporate bond market are in line with those in the equity market and investors in corporate credit do not, on average, benefit from a significant risk premium that is specific to credit markets.

The Credit Spread Puzzle in the Merton Model - Myth or Reality? *

Peter Feldhütter†
London Business School

Stephen Schaefer‡
London Business School

July 1, 2015

Abstract

We test the Merton model of credit risk using data on individual firms for the period 1987-2012. We find that the model matches the average level of investment grade spreads and furthermore captures the time series variation of the BBB-AAA spread well with a correlation between 84% and 92%. A crucial ingredient to the success of the model is that we use default rates for a long period of 92 years to calibrate the model. In simulations we show that such a long history of *ex post* default rates is essential to obtain estimates of *ex ante* default probabilities that have a reasonable level of precision; using default rates from shorter periods as done in most existing studies will often lead to the conclusion that spreads in the Merton model are too low even if the Merton model is the true model. Finally, we show that using data on individual firms - rather than a "representative firm" - is important for matching the slope of the term structure of credit spreads.

Keywords: Credit spread puzzle, Merton model, Structural models, Corporate bond spreads, Realized default frequencies;

JEL: C23; G01; G12

* We are grateful for valuable comments and suggestions received from Antje Berndt (discussant), Harjoat Bhamra (discussant), Hui Chen (discussant), Mike Chernov, Darrell Duffie, Stefano Giglio (discussant), Jean Helwege (discussant), Ralph Koijen, David Lando, Erik Loualiche (discussant), Gustavo Manso (discussant), Jens-Dick Nielsen, Lasse Heje Pedersen, Scott Richardson, and seminar participants at the NBER Asset Pricing 2015 conference, Utah Winter Finance Conference 2015, HKUST Finance Symposium 2014, WFA 2014, SFS Cavalcade 2014, ESSFM 2014 in Gerzensee, Fifth Risk Management Conference 2014, UNC's Roundtable 2013, UCLA Anderson School of Management, UC San Diego, Financial Conduct Authority, AQR, London Business School, Vienna Graduate School of Finance, Stockholm School of Economics, Tilburg University, Duisenburg School of Finance Amsterdam, NHH Bergen, Rotterdam School of Management, University of Southern Denmark, Copenhagen Business School, Brunel University, Henley Business School, Bank of England, and University of Cambridge. We are particularly grateful for the research assistance provided by Ishita Sen.

†London Business School, Regent's Park, London, NW1 4SA, `pfeldhutter@london.edu`

‡London Business School, Regent's Park, London, NW1 4SA, `sschaefer@london.edu`

1. Introduction

Merton (1974)'s model of credit risk is the classic textbook framework for studying the default risk of corporations and pricing of their debt. While the model is intuitive and simple, many studies find that credit spreads calculated using the Merton model are lower than actual spreads, particularly for investment grade bonds, giving rise to the so-called credit spread puzzle in the Merton model.[1] In this paper, we test the Merton model using a longer history of data than in previous studies and find that the model is able to capture not only the level but also the time series variation of investment grade credit spreads. This result diverges from much of the literature and we explain carefully why we reach this different conclusion.

We implement the Merton model using 286,234 individual bond yield observations for industrial firms in the period 1987-2012. On average, we find that the Merton has little difficulty in capturing credit spreads to AAA yields for investment grade bonds. The average actual long-term credit spread on BBB-rated bonds is 121 basis points while the Merton model implies an average credit spread of 113 basis points. For A-rated bonds the long-term credit spread is 40 basis points in the data and 51 basis points in the Merton model. Only for short maturities of less than two years does the Merton model predict significantly lower spreads than in the data. For speculative grade bonds, spreads in the Merton model are lower than those in the data which is consistent with an illiquidity premium for speculative grade bonds documented in Dick-Nielsen, Feldhütter, and Lando (2012).

Having shown that the Merton can capture average investment grade credit spreads, we put the model to a harder test and ask if the model can also match variations in spreads over time. We calculate the average BBB-AAA spread on a monthly basis and find that there is high degree of co-movement between actual and model-implied BBB-AAA spreads; the

[1] Papers finding that the Merton model (or close variants thereof) underpredict credit spreads include Eom, Helwege, and Huang (2004), Leland (2006), Cremers, Driessen, and Maenhout (2008), Chen, Collin-Dufresne, and Goldstein (2009), Chen (2010), Huang and Huang (2012), and McQuade (2013).

correlation is between 84% and 92% depending on maturity. Thus, the BBB-AAA spread – which has received most attention in the literature – is tracked well by the Merton model.

Our implementation allows for both heterogeneity in firms and variation in leverage ratios over time while at the same time historical default rates are matched. We do so by implying out the default boundary of firms such that model-implied default probabilities match historical default rates. This implementation combines two different approaches in the literature. The first is to use a "representative" firm and match historical default rates. The second is to use individual firms, thus allowing for heterogeneity, but disregard the level of model-implied default probabilities[2]. Departing from the representative firm approach enables us to study time series variation in spreads, while matching default rates ensures that the amount of default risk is consistent with historical experience.

Our results are different from most of the literature because we use a *long* (92-year) history of default rates to calibrate the model. We show in simulations that a long history of *ex post* default rates is crucial in estimating *ex ante* default probabilities with a reasonable degree of precision. In contrast, using a short history of around 30 years, as in much of the literature, is likely to lead to the conclusion that the Merton model underpredicts spreads even if the model is correct. The reason is that the distribution of default rates is very skewed: most of the time we see few defaults but occasionally we see many defaults. Key to this result is that defaults are correlated. To see this, we can think of a large number of firms with a default probability (over some period) of 5% and where their defaults are perfectly correlated. We will see no defaults 95% of the time (and 100% defaults 5% of the time) so the realized default rate will underestimate the default probability 95% of the time. If an average default rate is calculated over two independent periods, the realized default rate will only underestimate the default probability $0.95^2 = 90.25\%$ of the time; thus a longer history

[2]See Cremers, Driessen, and Maenhout (2008), Chen, Collin-Dufresne, and Goldstein (2009), Chen (2010), and Huang and Huang (2012) for examples of the first approach and Eom, Helwege, and Huang (2004), Ericsson, Reneby, and Wang (2007), and Bao (2009) for the second approach.

reduces the skewness.

To illustrate the importance of using a long history of default rates, we revisit the results in Chen, Collin-Dufresne, and Goldstein (2009). They find a credit spread puzzle in the Merton model by using default rates from the period 1970-2001. When we use default rates from the period 1920-2001 instead, the puzzle disappears.

Using individual firm data instead of a representative firm is important for predicting the term structure of credit spreads. We show this by simulating an economy with a cross section of firms that are identical except that they have different leverage ratios. If one calculates spreads using a representative firm with average leverage, spreads at all maturities are typically too low due to a "convexity effect" (Leland (2006) and McQuade (2013) are examples of this approach). If a representative firm is calibrated to match the average T-year default probability, the representative firm's T-year credit spread provides a reasonable estimate of the average T-year credit spread in the economy. However, the credit spread of the same representative firm at shorter maturities is much lower than the average economy-wide credit spread at these maturities (Cremers, Driessen, and Maenhout (2008) and Zhang, Zhou, and Zhu (2009) are examples of this approach). Thus, allowing for firm heterogeneity is crucial for calculating the term structure of credit spreads.

There is a large literature on testing the Merton model. Leland (2006), Cremers, Driessen, and Maenhout (2008), Chen, Collin-Dufresne, and Goldstein (2009), Chen (2010), and Huang and Huang (2012) use a representative firm to test variants of the Merton model and find that the model underpredicts spreads. In contrast to these papers, we allow for heterogeneity in leverage ratios and asset volatility across firms and over time, calibrate to more robust "long-term" rather than "short-term" default rates and report the time series variation of spreads. Eom, Helwege, and Huang (2004), Ericsson, Reneby, and Wang (2007), and Bao (2009) allow for heterogeneity in firms and variations in leverage ratios, but none of these calibrates to historical default rates. Eom, Helwege, and Huang (2004) and Bao (2009) also do not look at the time series variation in credit spreads. It appears to us that much of the literature has not recognised the necessity of using a long history of default rates to estimate default probabilities with a reasonable degree of precision, although Bhamra, Kuehn, and Strebulaev (2010) make a related point. In a structural-equilibrium model with macro-economic risk, they simulate default rates over 5 and 10 years and find a wide distribution. In their model, uncertainty in default rates arises because of systematic cash flow risk, refinancing, and an endogenous default boundary that jumps when macro-economic conditions change according to a regime-switching model. We show that even with no refinancing, a constant default boundary, and average default rates measured over *several decades* there is a large amount of variation in realized default rates. The convexity effect when using a representative firm is well-known and David (2008) and Bhamra, Kuehn, and Strebulaev (2010) find the effect to be crucial while Bhamra, Kuehn, and Strebulaev (2010) and Chen, Collin-Dufresne, and Goldstein (2009) argue that the effect is negligible. We clarify why there is disagreement in the literature about the importance of convexity, and focus on the term structure implications where we show the convexity bias is severe.[3]

The organization of the article is as follows: Section 2 explains the data and how the Merton model is implemented. Section 3 tests the Merton model using individual bond data for the period 1987-2012. Section 4 examines why our results are different from those commonly found in the literature and Section 5 concludes.

[3]Bhamra, Kuehn, and Strebulaev (2010) also look at the term structure effect of the convexity bias, but since they do not hold the default probability constant when comparing a representative firm with a cross section of firms, it is not clear if their results are due to the convexity bias or a change in default probabilities.

2. The Merton model: basics and implementation

2.1 Data

For the period January 1, 1997 to July 1, 2012, we use daily quotes provided by Merrill Lynch (ML) on all corporate bonds included in the ML investment grade and high-yield indices. These data are used by, among others, Schaefer and Strebulaev (2008) and Acharya, Amihud, and Bharath (2013). We obtain bond information from the Mergent Fixed Income Securities Database (FISD) and limit the sample to senior unsecured fixed rate or zero coupon bonds. We exclude bonds that are callable, convertible, putable, perpetual, foreign denominated, Yankee, have sinking fund provisions, or have covenants.[4] For the period April 1987 to December 1996 we use monthly data from the Lehman Brothers Fixed Income Database. This data is used by among others Duffee (1998), Huang and Huang (2012), and Acharya, Amihud, and Bharath (2013). We include only data that are actual quotes (in contrast to data based on matrix-pricing) and exclude bonds that are callable or contain sinking fund provisions. The database starts in 1973, but there are two reasons why we start from April 1987. First, there are only few noncallable bonds before the mid-80s (see Duffee (1998)) and second, we do not have swap rates prior to April 1987. We use only bonds issued by industrial firms. The sample from ML contains 235,419 observations and the sample from Lehman 50,815 observations; in total we have 286,234 observations.

To price a bond in the Merton model we need the issuing firm's asset volatility, leverage ratio, and payout ratio. *Leverage ratio* is calculated as the book value of debt divided by firm value (where firm value is calculated as book value of debt plus market value of equity). *Payout ratio* is calculated as the sum of interest payments to debt, dividend payments to equity, and net stock repurchases divided by firm value. *Asset volatility* is not directly

[4] For bond rating, we use the lower of Moody's rating and S&P's rating. If only one of the two rating agencies have rated the bond, we use that rating. We track rating changes on a bond, so the same bond can appear in several rating categories over time.

observable and is estimated from equity volatility and leverage as we will explain in Section 2.2. All firm variables are obtained from CRSP and Compustat and details are given in Appendix A.

2.2 The Merton model

In this section we describe our implementation of the Merton model and provide the associated formulae for credit spreads and default probabilities. We omit derivations and refer to Chen, Collin-Dufresne, and Goldstein (2009) Appendix A for details.

Asset value in the Merton model follows a Geometric Brownian Motion under the natural measure,

$$\frac{dV_t}{V_t} = (\mu - \delta)dt + \sigma dW_t^P \tag{1}$$

where δ is the payout rate, μ is the expected return on the firm's assets and σ is the volatility of asset value.

The firm is financed by equity and a single zero-coupon bond with face value F and maturity T.[5] If the asset value falls below some default boundary at the bond's maturity, $V_T < D$, the firm cannot repay its bondholders and the firm defaults. We let the default boundary be a fraction d of the face value of debt such that $D = dF$. As we explain later we estimate d such that model-implied default probabilities are consistent with historical

[5]To keep the model as simple as possible, we assume that the bond is a zero-coupon bond and implicitly account for coupons by estimating the payout rate as the total payout to debt and equity holders. An alternative would be to assume that payout is only dividends to equity holders and the firm refinances coupons and repays them at bond maturity. In this case the drift of the firm would be higher, but the amount of debt would also be higher and these two effects offset each other and the model is the same as the one we present. Finally, we could - at the expense of simplicity - allow for coupon payments as in Eom, Helwege, and Huang (2004). This generally leads to higher spreads because the firm can default not only at bond maturity but also before bond maturity.

default rates. The default probability is

$$\pi^P = N\Big[-\Big(\frac{1}{\sqrt{\sigma^2 T}}\Big)\Big(\log\big(\frac{V_0}{dF}\big) + \big(\mu - \delta - \frac{\sigma^2}{2}\big)T\Big)\Big]. \tag{2}$$

To be consistent with empirical recovery rates we follow Chen, Collin-Dufresne, and Goldstein (2009) and assume that the bondholders receive R if the firm defaults. We set $R = 37.8\%$ which is Moody's (2013)'s average recovery rate, as measured by post-default trading prices, for senior unsecured bonds for the period 1982-2012.

Under the risk-neutral measure asset value follows

$$\frac{dV_t}{V_t} = (r - \delta)dt + \sigma dW_t^Q \tag{3}$$

and the credit spread is given as

$$(y - r) - \frac{1}{T}\log\big[1 - (1 - R)\pi^Q\big]$$

where the risk-neutral probability of default π^Q is

$$\pi^Q = N\Big[-\Big(\frac{1}{\sqrt{\sigma^2 T}}\Big)\Big(\log\big(\frac{V_0}{dF}\big) + \big(r - \delta - \frac{\sigma^2}{2}\big)T\Big)\Big].$$

We set the asset Sharpe ratio, $\theta = \frac{\mu - r}{\sigma}$, to Chen, Collin-Dufresne, and Goldstein (2009)'s estimate of 0.22. Chen, Collin-Dufresne, and Goldstein (2009) show that the credit spread is given as

$$(y - r) = -\Big(\frac{1}{T}\Big)\log\Big(1 - (1 - R)N\big[N^{-1}(\pi^P) + \theta\sqrt{T}\big]\Big).$$

The last formula nicely shows that the credit spread can be expressed in terms of only the default probability, recovery rate, and Sharpe ratio.

An important parameter is the volatility of assets and here we follow the approach of Schaefer and Strebulaev (2008) in calculating asset volatility. Since firm value is the sum of the debt and equity values, asset volatility is given by:

$$\sigma^2 = (1 - L)^2\sigma_E^2 + L^2\sigma_D^2 + 2L(1 - L)\sigma_{ED}, \tag{4}$$

where σ is the volatility of assets, σ_D volatility of debt, σ_{ED} the covariance between the returns on debt and equity, and L is leverage ratio. If we assume that debt volatility is zero, asset volatility reduces to $\sigma = (1-L)\sigma_E$. This is a lower bound on asset volatility. Schaefer and Strebulaev (2008) (SS) compute this lower bound along with an estimate of asset volatility that implements equation (4) in full. They find that for investment grade companies the two estimates of asset volatility are similar while for junk bonds there is a significant difference. We compute the lower bound of asset volatility, $(1-L)\sigma_E$, and multiply this lower bound with SS's estimate of the ratio of asset volatility computed from equation (4) to the lower bound. Specifically, we estimate $(1-L)\sigma_E$ and multiply this by 1 if $L < 0.25$, 1.05 if $0.25 < L \leq 0.35$, 1.10 if $0.35 < L \leq 0.45$, 1.20 if $0.45 < L \leq 0.55$, 1.40 if $0.55 < L \leq 0.75$, and 1.80 if $L > 0.75$[6]. This method has the advantage of being transparent and easy to replicate.

Finally, the riskfree rate r is chosen to be the swap rate for the same maturity as the bond. For maturities shorter than one year we use LIBOR rates.

2.3 Summary statistics

Summary statistics for the firms in our sample are shown in Table 1. Since the sample contains few AAA-rated firms, particular in the later years, we combine AAA and AA into one rating group and refer to the combined group as AAA-rated bonds later in the paper. The average leverage ratios of 0.14 for AAA/AA, 0.28 for A, and 0.38 for BBB are similar to those found in other papers.[7] The payout ratio is similar across rating, with an average

[6]These fractions are based on Table 7 in SS apart from 1.80 which we deem to be reasonable. Results are insensitive to other choices of values for $L > 0.75$. See also Correia, Kang, and Richardson (2014) for an assesment of different approaches to calculating asset volatility.

[7]Huang and Huang (2012) use a leverage ratio of 0.13 for AAA, 0.21 for AA, 0.32 for A, and 0.43 for BBB while Schaefer and Strebulaev (2008) find an average leverage of 0.10 for AAA, 0.21 for AA, 0.32 for A, and 0.37 for BBB.

across rating of 4.7%, implying that higher coupon payments for riskier firms are offset by smaller dividend payments. Average equity volatilities are monotonically increasing with rating consistent with a leverage effect. The estimates are similar to those in SS for A-AAA ratings, while the average equity volatility for BBB firms of 0.38 is higher than the value of 0.33 estimated in SS. Asset volatilities are slightly increasing in rating consistent with the estimates in SS.

In Table 2 we see that the number of bonds with a low rating of B or C is small relative to the number of investment grade bonds. One reason for this is that speculative grade bonds frequently contain call options which leads to their exclusion from our sample. Since the number of speculative grade firm and bond observations in the sample is small relative to the number of observations in the investment grade segment, we place all speculative grade bonds in one group.

2.4 Estimating the default boundary

The level of the default boundary plays an important role in spread predictions: holding other parameters constant a higher default boundary leads to higher default probabilities and thus higher spreads. Direct estimates of the default boundary are difficult to obtain and a range of estimates has been used in the literature.[8] Huang and Huang (2012) show that holding expected default probabilities fixed at historical averages eliminates the dependence of the spread on the default boundary. In their implementation they fix the default boundary and imply out asset volatility by matching historical default rates. We follow their approach, but since the default boundary is harder to estimate than asset volatility, we imply out the default boundary instead of asset volatility.

Specifically, if we observe a spread observation of bond i with a time-to-maturity T issued

[8]Davydenko (2013) estimate the default boundary using market values of debt and equity, but his estimates depend on difficult-to-estimate costs of default.

by firm j on date t, we calculate the firm's T-year default probability $\pi^P(d, \Theta_{jt}, T)$ using equation (2). Here, Θ_{jt} contains the time-t estimate of asset volatility, payout rate, and leverage ratio.[9] For a given rating a and maturity T - rounded up to the nearest integer year - we find all N_{aT} bond observations in the sample with the corresponding rating and maturity and calculate the average default probability $\overline{\pi}^P_{aT}(d)$. We denote the corresponding historical default frequency for $\hat{\pi}^P_{aT}$ and use Moody's historical default frequencies for the period 1920-2012. For all investment grade ratings and horizons of 1-20 years (Moody's only report default rates for up to 20 years horizon) we find the value of d that minimizes the weighted sum of absolute errors between the historical and model-implied default rates by solving

$$\min_{\{d\}} \sum_{a=AAA}^{BBB} \sum_{T=1}^{20} N_{at} |\overline{\pi}^P_{aT}(d) - \hat{\pi}^P_{aT}|.$$

We compute the default boundary for each calendar year from estimated values of $\overline{\pi}^P$ in that year and our estimate of d is the average of the 26 yearly estimates from 1987 to 2012. This average value is 1.036 which implies that the default boundary is close to the face value of debt, as in the Merton model, and we therefore set $d = 1$[10]. By implying out the default boundary from historical default rates while allowing for heterogeneity in firms and leverage

[9] For simplicity we omit that the default probability depends on the riskfree rate and Sharpe ratio because they are not firm dependent.

[10] It may appear that, together, a default boundary of 1 and a recovery rate of 0.378 imply a large deadweight cost of bankruptcy, but this misses an important observation made by Bao (2009). Much corporate debt is bank debt that has a much higher recovery rate; Acharya, Bharath, and Srinivasan (2007) find that the average (median) recovery rate for bank debt is 81% (92%). Furthermore, much of total debt is bank debt; Houston and James (1996) find that 64% of total debt is bank debt. This implies that modest deadweight costs of bankruptcy measured as percentage of firm value can lead to low recovery rates on bonds. For example, suppose that the face value of debt is 100 of which bank debt is 64 and bond debt is 36, and 22% of firm value is lost at default. If the firm value is 100 at default and 22 is lost, this leaves 64 for the bank debt holders (because they are senior) and 14 for bond holders. Thus a 22% deadweight cost of bankruptcy leads, in this case, to a loss of 20 (36 - 16) or 61% of face value for bond holders.

ratios, we combine two strands of literatures examining structural models of credit risk. The first strand matches historical default rates but uses an average "representative firm" (Cremers, Driessen, and Maenhout (2008), Chen, Collin-Dufresne, and Goldstein (2009), Chen (2010), and Huang and Huang (2012)). Departing from the representative firm assumption allows us to study the time series variation of spreads. The second strand allows for cross-sectional heterogeneity in firms, but does not target historical default rates (Eom, Helwege, and Huang (2004), Ericsson, Reneby, and Wang (2007), and Bao (2009)). In Section 4.3 we discuss the problems in using a "representative firm" that result from a convexity bias.

3. Empirical results

In this section we compare spreads in the Merton model with actual bond spreads. Corporate bond spreads are sensitive to the choice of riskfree rate (typically either Treasury yields or swap rates) and so here we measure bond spreads to AAA-yields such that we cleanly separate our results from the specific choice of riskfree rate. We later investigate the spread between AAA-rated bonds and the riskfree rate.

3.1 Investment grade corporate bond credit spreads

Table 3 shows actual and model-implied bond spreads to AAA-rated bonds in our sample. To calculate an average spread, we first calculate a monthly average spread and then calculate the average spread across months. Credit spreads for investment grade ratings A and BBB for medium and long maturity bonds are on average well matched by the Merton model. For example, the average long maturity BBB-AAA spread is 121 basis points in the data and 113 in the Merton model. The empirical estimates of the BBB-AAA spread are slightly higher than the 100-110 basis points typically found in the literature and reflects that the 2008-2009 crisis saw the highest corporate bond spreads since the Great Depression; before 2008 actual spreads are close to those found in earlier studies. Overall, the correspondence between actual and model-implied investment grade spreads in the three sub-samples is good.

The finding that average spreads in the Merton model are similar to average actual spreads for investment grade bonds is surprising given that a number of papers have found that the Merton model underpredicts spreads on investment grade bonds. In Section 4 we discuss why our conclusions are different from most of the previous literature. For investment grade bonds it is only for maturities below two years that actual spreads are substantially higher than model-implied spreads in most of the periods in Table 3.

The literature has predominantly focused on the Merton model's ability to capture average spreads over long periods of time. Given our results that the model can capture average spreads for investment grade bonds, we next examine whether the model also captures the time series variation of credit spreads. The spread between BBB and AAA-rated bonds has received most attention in the literature and we therefore focus on this spread. In each month, we calculate the average BBB yield spread and subtract the average AAA yield spread and Figure 1 plots the time series variation in this BBB-AAA spread. The model-implied spread tracks the actual spread well for all maturities: the correlation is 86% for long-maturity bonds, 92% for medium-maturity bonds, and 84% for short-maturity bonds. The Merton model underpredicts spreads slightly during 2008-2009 consistent with findings in Bao, Pan, and Wang (2011) and Dick-Nielsen, Feldhütter, and Lando (2012) of an illiquidity premium during this period. If we restrict the sample to the period from 1987 to 2007, the correlations are 82% (long-maturity), 75% (medium-maturity), and 31% (short-maturity) and so, apart from the results for short maturities, the correlations are not driven by the large spikes during 2008-2009. Overall, we find that the Merton model captures both the average level and time series variation of the BBB-AAA spread.

3.2 Speculative grade corporate bond credit spreads

For some maturities there are few speculative grade bonds in the sample compared to investment grade bonds (Table 2). Among short-maturity bonds for example there are no C-rated bonds and only 15 B-rated bonds. We therefore group all speculative grade bonds into one rating category in Table 3. The table shows that actual speculative grade bond spreads are higher than those implied by the Merton model; the average long-term speculative yield spread for example is 435 basis points in the data and 248 basis points in the Merton model. This is surprising because Huang and Huang (2012) finds that the credit spread puzzle is primarily a phenomenon in investment grade bonds and not in speculative grade bonds.

To investigate in more detail the ability of the Merton model to capture speculative grade yield spreads, we compute the average difference between actual and model-implied yield spreads for rating notches. Each broad rating category has three rating notches and because we aggregate all yield spreads for bonds with a maturity between zero and 30 years we have sufficient observations to be able to compute the average spread difference by rating notch. Figure 2 shows the results and the vertical line separates investment-grade ratings on the left-hand side from speculative-grade ratings on the right-hand side. For investment grade ratings the average difference between actual and model-implied yield spreads is small for all rating notches and ranges from -16 to +14 basis points. For speculative grade ratings the actual spread is substantially larger than the model-implied spread and the difference ranges from 67 to 473 basis points[11]. Thus, as we move from a low investment grade rating to a high speculative grade rating the Merton model starts to underpredict spreads.

[11]The results are not sensitive to the choice of sample period. If we split the sample period into three subperiods of equal length, the average difference between actual and model-implied spreads is positive in 8 out of 18 instances for investment grade bonds (six rating notches times three time periods) while the difference is positive for 17 of 18 instances for speculative grade bonds.

3.3 Spread between the AAA yield and the riskless rate

Traditionally, Treasury yields have been used as riskfree rates, but recent evidence shows that swap rates are a better proxy than Treasury yields. A major reason for this is that Treasury bonds enjoy a convenience yield that pushes their yields below riskfree rates (Feldhütter and Lando (2008), Krishnamurthy and Vissing-Jorgensen (2012), and Nagel (2014)). Hull, Predescu, and White (2004) find that the riskfree rate is 63bps higher than Treasury yields and 7bps lower than swap rates, Feldhütter and Lando (2008) find riskfree rates to be 53bps higher than Treasury yields and 8bps lower than swap rates, while Krishnamurthy and Vissing-Jorgensen (2012) find that riskfree rates are on average 73bps higher than Treasury yields. However, on occasion, swap rates also deviate significantly from riskfree rates. Feldhütter and Lando (2008) find that during 2002-2003 swap rates were influenced by supply and demand factors and were pushed 40bps below the riskfree rate. More recently, long-term swap rates became extraordinary low following the Lehman default in September 2008. For example, the 30-year swap spread to Treasury was around 50bps in the two years before the Lehman default, but went negative and stayed negative until the end of the sample period. This implies that although recent papers find swap rates to be more appropriate riskfree rates than Treasury yields in general, swap rates are poor proxies for riskfree rates in certain periods, and it is difficult to identify a single rate that is a good proxy for the riskfree rate at all times.

Figure 3 plots the time series variation in the spread between AAA-yields and both the swap rate and the Treasury yield along with the model-implied AAA bond spread. When we use Treasury yields as the riskfree rate, the AAA-riskfree spread is substantially underpredicted by the Merton model for all maturities. For example, the average long-term AAA-Treasury spread is 77bps and only 10bps in the Merton model.

There is also an underprediction of long-term AAA spreads when using the swap rate as the riskfree rate, but the difference is relatively small. For long-maturity bonds, the difference in average spreads is 14bps (the actual spread to the swap rate is 25 bps versus 11bps in the Merton model), while for medium and short-maturity bonds the difference between the average empirical and model-implied spread is 12bps respectively 14bps. Figure 3 shows that for most of the sample period, AAA-spreads to the swap rate are not far from zero. The median spread to swap is 10bps for short maturity, 6bps for medium maturity, and 16bps for long maturity. This implies that when the swap rate is used as the riskfree rate, high-quality corporate bond spreads–particularly at shorter maturities–are close to zero most of the time and only deviate substantially around the recessions of 2001 and 2008-2009. It has previously been regarded as a failure of the Merton model that high-quality short-term spreads are close to zero, but these results suggest that most of the time this corresponds to the spreads to the swap rate that we actually observe.

4 Why our results are different from most of the literature

A number of papers find that the Merton model cannot match the level of credit spreads and in this section we explain why we reach a different conclusion.[12] We show that using a long history (around 90 years) of *ex post* default rates is crucial for pinning down *ex ante* default probabilities with a reasonable degree of precision. The literature has typically used a short history (around 30 years) of default rates and we show that the finding that the Merton model underpredicts spreads in Chen, Collin-Dufresne, and Goldstein(2009) disappears once we use a long history of default rates. We also clarify when and how the common use of a "representative firm" biases spread predictions.

[12] Papers finding that the Merton model (or variants thereof) underpredicts spreads include Leland (2006), Cremers, Driessen, and Maenhout (2008),Chen, Collin-Dufresne, and Goldstein (2009), Chen (2010), Huang and Huang (2012), and McQuade (2013). See also Sundaresan (2013) for a recent review.

4.1 A long history of realized default rates is important

In many empirical studies of credit spreads ex-post realized default rates play a crucial role since they are used in lieu of *ex ante* default probabilities.[13] Even abstracting from the problem of secular variation in the default rate, there are two significant problems in obtaining reliable *ex-ante* default probabilities from realized default frequencies. The first is that the low level of default frequency, particularly for investment grade firms, leads to a sample size problem with default histories as short as those typically used in the literature (around 30 years). The second is that, even though the problem of sample size is potentially mitigated by existence of a large cross-section of firms, defaults are correlated across firms, so the benefits of a large cross-section in improving precision are greatly reduced.

While both these points are "obvious", their importance in the context of studies of credit pricing may have been underestimated and, more to the point, the benefits of using a long history of defaults as a means of addressing both problems seems to have been overlooked.

There is a tradition in the literature to use realized default rates published by Moody's.[14] To explain how Moody's calculate default frequencies, let us consider the 10-year BBB cumulative default frequency of 4.39% used in Cremers, Driessen, and Maenhout (2008) and Huang and Huang (2012). This number is published in Keenan, Shtogrin, and Sobehart (1999) and is based on default data in the period 1970-1998. For the year 1970, Moody's define a cohort of BBB-rated firms and then report how many of these default over the next 10 years. The 10-year BBB default frequency for 1970 is the number of defaulted firms divided by the number in the 1970 cohort.[15] The average default rate of 4.39% is calculated

[13] See for example Leland (2006), Cremers, Driessen, and Maenhout (2008), Chen, Collin-Dufresne, and Goldstein (2009), Chen (2010), and Huang and Huang (2012).

[14] All articles mentioned in Footnote 13 use Moody's estimates of default frequencies.

[15] Some firms have their rating withdrawn and Moody's have incomplete knowledge of subsequent defaults once firms are no longer rated. Moody's adjust for this by assuming that firms with withdrawn ratings would have faced the same risk of default as other similarly rated issuers if they had stayed in the data sample. Evidence in Hamilton and Cantor (2006) suggests that this is a reasonable assumption.

as an average of the 10-year default rates for the cohorts formed at yearly intervals over the period 1970-1988[16].

To assess the statistical accuracy of the realized default rate of 4.39% as an estimate of the *ex ante* default rate we carry out a simulation. In an economy where the *ex ante* 10-year default probability is 4.39% for all firms, we simulate the ex post realized 10-year default frequency averaged over 28 years. We assume that in year 1 we have 1,000 identical firms[17], where firm i's value under the natural measure follows the GBM in equation (1)

$$\frac{dV_t^i}{V_t^i} = (\mu - \delta)dt + \sigma dW_{it}^P. \tag{5}$$

We assume every firm has one T-year bond outstanding, and a firm defaults if firm value is below bond face value at bond maturity, $V_T^i \leq F$. In the simulation $T = 10$. Using the properties of a Geometric Brownian Motion, the default probability is

$$p = P(W_{iT}^P - W_{i0}^P \leq c) \tag{6}$$

where $c = \frac{\log(F/V_0)-(\mu-\delta-\frac{1}{2}\sigma^2)T)}{\sigma}$. This implies that the unconditional default probability is $N(\frac{c}{\sqrt{T}})$ where N is the cumulative normal distribution. For a given default probability p we can always find c such that equation (6) holds, so in the following we use p instead of the underlying Merton parameters that give rise to p[18].

We introduce systematic risk by assuming that

$$W_{iT}^P = \sqrt{\rho} W_{sT} + \sqrt{1-\rho} W_{iT} \tag{7}$$

[16]In recent years Moody's calculate average default frequencies based on monthly cohorts instead of yearly cohorts; the difference between default frequencies using monthly and yearly cohorts is small.

[17]We choose 1,000 firms each month because the average number of firms in Moody's BBB cohorts during the last decade is close to 1,000. The average number of BBB cohort firms during 1970-2012 is 606. There has been an increasing trend from 372 in 1970 to 1,245 in 2012. The results are very similar if we use 600 or 1,250 firms instead of 1,000.

[18]Although only p is relevant for the simulation, we note that if we use median parameter values for BBB firms, $\sigma_A = 0.24$, $\delta = 0.045$, leverage=0.36, and $r = 0.05$ and a Sharpe Ratio of 0.22, the resulting default probability is 4.22%, very close to $p = 4.39\%$ used in the simulation.

where W_i is a Wiener process specific to firm i, W_s is a Wiener process common to all firms, and ρ is the pairwise correlation between percentage firm value changes. All the Wiener processes are independent. The realized 10-year default frequency in the year-1 cohort is found by simulating one systematic and 1,000 idiosyncratic processes in equation (7).

In year 2 we form a cohort of 1,000 new firms. The firms in year 2 have characteristics that are identical to those of the previous firms at the point they entered the index in year 1. We calculate the realized 10-year default frequency of the year-2 cohort as we did for the year-1 cohort. Crucially, the common shock for years 1-9 for the year-2 cohort is the same as the common shock for years 2-10 for firms in the year-1 cohort. We repeat the same process for 18 years and calculate the overall realized cumulative 10-year default frequency in the economy by taking an average of the default frequencies across the 18 cohorts. Finally, we repeat this entire simulation 100,000 times.

There are only two parameters in our simulation; the default probability p and the default correlation ρ. As mentioned we set $p = 4.39\%$. We assume $\rho = 0.25$; this is consistent with Cremers, Driessen, and Maenhout (2008) who find an average pairwise equity correlation of 25.4% for S&P 100 firms.

The top graph in Figure 4 shows the distribution of realized default rate in the simulation study.[19] A 95% confidence interval is [0.56%; 13.50%]. The black vertical line shows the ex ante default probability of 4.39%. Given that we simulate 18,000 firms over a period of 28 years, it might be surprising that the realized default rate can be far from the ex ante default probability. The reason is simply the presence of systematic risk in the economy which

[19]The distribution of the default frequency in Figure 4 is similar to the one derived analytically by Vasiček (1991). The reason that we cannot use Vasicek's result here is that the number reported by Moody's is the *average* of 10-year default rates while Vasicek's formula refers to the default rate over a single period. In a previous version of the paper we showed that if the correlation parameter in the Vasicek formula is set equal to the average correlation produced by Moody's overlapping cohorts this produces a good approximation to the distribution in Figure 4.

induces correlation in defaults among firms. If there is no systematic risk in the economy Table 4 shows that a 95% confidence interval for the realized default rate is $[4.11\%; 4.68\%]$.

We also see from Figure 4 that the default frequency is significantly skewed to the right, i.e., the modal value of around 2% is significantly below the mean of 4.39%. This means that the default frequency *most often* observed – e.g., long-run data from the rating agencies – is below the mean. This in turns implies that the number reported by Moody's (4.39%) is more likely to be below the true mean than above it and, in this case, if spreads reflect the true expected default rate, they will appear too high relative to the observed historical loss rate[20].

Moody's started to record default rates in 1920 so in practice we can increase statistical precision by extracting an estimate of the ex ante default probability from the average default frequency for the period 1920-2012 instead of for the period 1970-1998. In the bottom graph in Figure 4 we therefore repeat the simulation where we maintain the ex-ante default probability at 4.39% but, instead of 28 years, we use 90 years in each simulated economy. We see that the distribution is tighter and more symmetric; the modal value is close to 4% and the width of the 95% confidence band of $[1.70\%; 8.61\%]$, although still wide, is almost half that when using 28 years of data[21]. Thus, in those cases where historical default rates are employed in investigating the ability of structural models to price corporate bonds, using a long time series of defaults is of crucial importance.

In the simulations we assume that the cohorts are different each year and so a given firm only appears in a single cohort. This is a simplification because when Moody's form cohorts of BBB-rated firms from year to year, there is a substantial overlap of firms from one

[20]This is a point that Moody's KMV are aware of, see for example Kealhofer, Kwok, and Weng (1998) and Bohn, Arora, and Korablev (2005).

[21]Note that we hold p fixed at the historical default rate for the period 1970-1998, so $[1.70\%; 8.61\%]$ cannot be viewed as a confidence band for the 10-year BBB default probability for the long period because in such case we would need to set p equal to the default rate for the period 1920-2012–which is 7.11%–in the simulations. In this case the 95% band is $[3.10\%; 12.99\%]$.

cohort to the next. This overlap increases the correlation between the default frequencies in different cohorts in addition to that caused by systematic risk. If we allowed firms to stay in more than one cohort, the dispersion in the distribution of realised defaults would be even larger. Also, the amount of systematic risk is assumed to be $\rho = 0.25$, but as Table 4 shows, even with low levels of systematic risk, there is considerable uncertainty regarding default rates. Finally, it is important to note that as the horizon over which default rates are calculated (10 years in the simulation) is increased the statistical uncertainty increases. Intuitively, over a 30-year horizon, there are six independent observations of 5-year default rates while there are only three independent observations of 10-year default rates.

Overall, we show in this section that a long history of realized default rates is necessary to estimate 10-year default probabilities with a reasonable degree of precision. We conclude that it is important to use a "long" history (around 90 years), while existing studies typically use a "short" history (around 30 years)[22].

4.2 Is our use of more recent data important?

In this study we use a longer time series of bond yield data that includes both older and more recent data than most of the literature. In particular, our sample period includes recent recessions. Frequently-used estimates of yield spreads for different ratings and maturities are provided by Duffee (1998) and his average yield spreads are based on the period 1985-1995.[23]

[22]Although it is not the focus of their paper, Bhamra, Kuehn, and Strebulaev (2010) makes a related point. In their structural-equilibrium model with macro-economic risk, they simulate default rates over 5 and 10 years and find a wide distribution. In their model, uncertainty in default rates arises because of systematic cash flow risk, endogenous refinancing, and an endogenous default boundary that jumps when macro-economic conditions change according to a regime-switching model. We show that even with no refinancing, constant default boundary, and average default rates measured over *several decades* there is a large amount of variation in realized default rates.

[23]Papers using data from Duffee (1998) include Leland (2006), Cremers, Driessen, and Maenhout (2008), Chen (2010), Huang and Huang (2012), and McQuade (2013).

To what extent is our conclusion that the Merton model can price corporate bonds the result of using more recent data? We examine this question by revisiting the results in Chen, Collin-Dufresne, and Goldstein (2009). In particular, we calculate yield spreads exactly as in their paper and show the results in Figure 5. Importantly, they calculate credit spreads based on Moody's default rates for the period 1970-2001 and, as the figure shows, in this case the Merton model underpredicts spreads substantially. However, as we argued in the previous section, it is important to use a long history of default rates and Figure 5 also shows credit spreads when we use Moody's default rates for the period 1920-2001.[24] The figure shows that using a long history of default rates brings spreads in the Merton model in line with actual spreads for investment grade bonds for both the 4-year and 10-year maturities.[25] This shows that our conclusion that the Merton model can price corporate bonds is not due to the use of more recent data, but due to the use of a long history of default rates.[26]

4.3 On the importance of cross-sectional heterogeneity of firms

In our implementation of the Merton model we allow for both cross-sectional and time series variation in leverage ratios and firm characteristics. That is, we depart from the use the "representative firm" approach often used in the literature.[27] Using a representative firm with average characteristics within a given rating category gives rise to a bias due to

[24]We calculate default rates for 1920-2001 by taking a weighted average of the average default rates for 1920-1999 and the default rates for the 2000 and 2001 cohorts.

[25]Duffee (1998) does not provide yield spread estimates for speculative grade bonds.

[26]The average long-term BBB-AAA spread (using Moody's Seasoned corporate bond yields available from the Federal Reserve's webpage) for the period 1970-2001 is 110 basis points while it is 121 basis points for the period 1920-2001. Thus the higher default rates in the early half of the 20th century were not reflected in substantially higher spreads which suggests that expected default probabilities were quite similar pre- and post 1970.

[27]See for example Leland (2006), Cremers, Driessen, and Maenhout (2008),Chen, Collin-Dufresne, and Goldstein (2009), Chen (2010), Huang and Huang (2012), and McQuade (2013).

Jensen's inequality because spreads are typically convex in firm variables. This convexity bias is illustrated in Figure 6. David (2008) and Bhamra, Kuehn, and Strebulaev (2010) argue that taking into account this convexity bias is crucial for raising model-implied credit spreads to actual spread levels. On the other hand, Chen, Collin-Dufresne, and Goldstein (2009) and Huang and Huang (2012) argue that, once historical default rates are matched, the convexity bias will if anything result in higher spreads for a representative firm compared to average spreads for the cross-section of firms. In this section we clarify how and when firm heterogeneity impacts the size of model-implied credit spreads. Overall we show that using a representative firm does not matter much for credit spread predictions at a single bond maturity (although it leads to biased parameters), but it matters crucially when predicting the term structure of credit spreads.

To show how heterogeneity in the leverage ratio affects spreads and default probabilities we carry out a simulation of 100,000 firms. For each firm we us an asset volatility of 24%, a payout rate of 4.5%, a recovery rate of 37.8%, and a Sharpe Ratio of 0.22. The firms differ only in their leverage ratios and we draw 100,000 values from a normal distribution with mean 0.36 and a standard deviation of 0.16.[28] The standard deviation of 0.16 is equal to the standard deviation of all observations of leverage ratios of BBB firms in our sample. Finally, the riskfree rate is 5%. For each firm we calculate the default probability and credit spread for different maturities (assuming that a firm has only a single bond outstanding at the corresponding maturity. Panel A in Table 5 shows the results including the correct average credit spread (row 1) and default probability (row 2) from the model and the correct asset volatility of 24% for each maturity in row 3.

Leland (2006) and McQuade (2013) use values of the leverage ratio, payout rate, and asset volatility (obtained from Schaefer and Strebulaev (2008)) averaged over time and firms to calculate model-implied credit spreads for a representative firm and then compare these

[28]If a simulated leverage ratio is negative we set it to zero. This implies that the average leverage ratio is slightly higher than 0.36, namely 0.3614 in our simulation.

with historical averages of actual credit spreads. As Figure 6 illustrates this typically results in a downward bias of spreads. We calculate the term structure of spreads and default probabilities for a representative firm with a leverage ratio of 0.36 in Panel B of Table 5. There is a downward bias in default probabilities and spreads relative to the correct values given in Panel A and the bias becomes more pronounced at shorter maturities. For example, the one-year spread and default probability of the representative firm in Panel B are more than 100 times smaller than the average one-year spread and default probability in Panel A. This is why David (2008) and Bhamra, Kuehn, and Strebulaev (2010) argue that taking into account firm heterogeneity is important. This is also why we disagree with Leland (2006)'s conclusion - based on a representative firm - that a wider range of structural models "can imply reasonable credit spreads for *longer* maturities, but **all underestimate credit spreads and/or default risks for short horizons**" (emphasis in the original). As Panel A shows the Merton model can indeed generate significant credit spreads and default probabilities at short maturities once the variation in leverage ratios is taken into account.

The most common approach in the literature is to let a representative firm match historical default rates, typically by backing out asset volatility.[29] To examine the effect on spreads of doing so we proceed as follows. For a given maturity, we compute the asset volatility that allows the representative firm to match the average default probability in the economy at that given maturity. We then calculate the spread at the given maturity using the implied volatility. Panel C shows the term structure of credit spreads and the implied asset volatilities. The downward bias in credit spread has disappeared and there is now a slight upward bias: model spreads are now approximately 10% higher than average spreads. This is why Chen, Collin-Dufresne, and Goldstein (2009) and Huang and Huang (2012) argue that one can use a representative firm despite the convexity effect. Although spreads are not downward biased (at least in this example), there are two problems with this approach. The first

[29]Examples include Cremers, Driessen, and Maenhout (2008),Chen, Collin-Dufresne, and Goldstein (2009), Zhang, Zhou, and Zhu (2009), Chen (2010), and Huang and Huang (2012).

problem is that asset volatility is biased: all firms in the economy have an asset volatility of 24% and yet the implied asset volatility ranges from 25.9% at the 10-year horizon to 39.4% at the one-year horizon. The finding that implied asset volatility in the Merton model is too high, particular at shorter horizons has been seen as a failure of the model, but this example shows that the pattern arises mechanically from the use of a representative firm. The second problem, as panel C suggests and panel D makes clear, is that it is not possible to match the term structure of default probabilities without counterfactually changing the asset volatility.

Cremers, Driessen, and Maenhout (2008) and Zhang, Zhou, and Zhu (2009) use a representative firm to imply out asset volatility by matching long-term default rates and then use this asset volatility to calculate the term structure of credit spreads at shorter maturities. We replicate this approach by implying out the asset volatility that makes the representative firm's default probability match the average default probability for the 10-year bond in the economy and then calculate the term structure of credit spreads and default probabilities for this representative firm. The implied asset volatility is 25.9% and the term structures are in Panel D. The difference between the implied asset volatility of 25.9% and the true value of 24% reflects a moderate convexity bias at the 10-year horizon, but since the bias becomes more severe at shorter horizons, the strong downward bias in both spreads and default probabilities reappears as maturity decreases. Thus, the bias in the term structure of estimated default probabilities and credit spreads persists when using a representative firm and imputing parameters from a single maturity[30].

[30]Our results clarify those in Bhamra, Kuehn, and Strebulaev (2010). Within the framework of their structural-equilibrium model, they compare a representative firm with a cross section of firms and find that the slope of the term structure of credit spreads is flatter for the cross section of firms. In their experiment, the cross section of firms have an average default probability that is more than three times as large as the default probability of the representative firm (their Table 3, Panels B and C). Since the term structure of credit spreads becomes flatter for a representative firm at the same time as default risk increases, it is not clear if it is cross sectional variation or the rise in default probability that drives the flattening of the term structure. Since we hold the 10-year default probability fixed in Panels A and D, it is clear in our analysis

In summary, we show that it is not meaningful to calculate the term structure of credit spreads using a representative firm. If one holds parameters fixed, spreads at short maturities are much too low. If one recalibrates asset volatility to match average default probability at each bond maturity, asset volatilities required to match short-horizon default probabilities are much too high.

5. Conclusion

We test the Merton model of credit risk using data on individual firms for the period 1987-2012. In the implementation we allow for heterogeneity in firms and match historical default rates. The Merton model captures the average level of investment grade spreads; in particular the long-term BBB-AAA spread is 113 basis points in the model and 121 basis points in the data. As a further test of the model, we find that the time series of the model-implied BBB-AAA spread tracks the actual BBB-AAA spread well with a correlation between 84% and 92% depending on bond maturity.

Our results show that the credit spread puzzle in the Merton model - the perceived failure of the model to explain levels of credit spreads particularly for investment grade bonds - has less to do with deficiencies of the model than with the way in which it has been implemented. Important for our conclusions is that we calibrate the model to empirical default rates that are calculated using a long period of 92 years. In simulations we show that such a long history of default rates is essential to estimate expected default probabilities with a reasonable degree of reliability; using default rates from short periods will often lead to the conclusion that spreads in the Merton model are too low even if the Merton model is the true model. Furthermore, using a cross-section of firms in the implementation is important for matching the slope of credit spreads.

---

that the flatter term structure is driven by cross-sectional variation in leverage alone

A Firm data

To compute bond prices in the Merton model we need the issuing firm's leverage ratio, payout ratio, and asset volatility. This Appendix gives details on how we calculate these quantities using CRSP/Compustat.

Firm variables are collected in CRSP and Compustat. To do so we match a bond's CUSIP with CRSP's CUSIP. In theory the first 6 digits of the bond cusip plus the digits '10' corresponds to CRSP's CUSIP, but in practice only a small fraction of firms is matched this way. Even if there is a match we check if the issuing firm has experienced M&A activity during the life of the bond. If there is no match, we hand-match a bond cusip with firm variables in CRSP/Compustat.

Leverage ratio: Equity value is calculated on a daily basis by multiplying the number of shares outstanding with the price of shares. Debt value is calculated in Compustat as the latest quarter observation of long-term debt (DLTTQ) plus debt in current liabilities (DLCQ). Leverage ratio is calculated as $\frac{\text{Debt value}}{\text{Debt value+Equity value}}$.

Payout ratio: The total outflow to stake holders in the firm is interest payments to debt holders, dividend payments to equity holders, and net stock repurchases. Interest payments to debt holders is calculated as the previous year's total interest payments (previous fourth quarter's INTPNY). Dividend payments to equity holders is the indicated annual dividend (DVI) multiplied by the number of shares. The indicated annual dividend is updated on a daily basis and is adjusted for stock splits etc. Net stock repurchase is the previous year's total repurchase of common and preferred stock (previous fourth quarter's PRSTKCY). The payout ratio is the total outflow to stake holders divided by firm value, where firm value is equity value plus debt value. If the payout ratio is larger than 0.13, three times the median payout in the sample, we set it to 0.13.

Equity volatility: We calculate the standard deviation of daily returns (RET in CRSP) in the past three years to estimate daily volatility. We multiply the daily standard deviation with $\sqrt{255}$ to calculate annualized equity volatility. If there are no return observations on more than half the days in the three year historical window, we do not calculate equity volatility and discard any bond transactions on that day.

References

Acharya, V., Y. Amihud, and S. Bharath (2013). Liquidity risk of corporate bond returns: conditional approach. *Journal of Financial Economics 110*(2), 358–386.

Acharya, V., S. Bharath, and A. Srinivasan (2007). Does industry-wide distress affect defaulted bonds? Evidence from creditor recoveries. *Journal of Financial Economics 85*, 787–821.

Bao, J. (2009). Structural Models of Default and the Cross-Section of Corporate Bond Yield Spreads. *Working paper*.

Bao, J., J. Pan, and J. Wang (2011). The illiquidity of Corporate Bonds. *Journal of Finance 66*, 911–946.

Bhamra, H. S., L.-A. Kuehn, and I. A. Strebulaev (2010). The Levered Equity Risk Premium and Credit Spreads: A Unified Framework. *Review of Financial Studies 23*(2), 645–703.

Bohn, J., N. Arora, and I. Korablev (2005). Power and level validation of the EDF credit measure in the U.S. market. *Moody's KMV Company*.

Chen, H. (2010). Macroeconomic conditions and the puzzles of credit spreads and capital structure. *Journal of Finance 65*(6), 2171–2212.

Chen, L., P. Collin-Dufresne, and R. S. Goldstein (2009). On the relation between the credit spread puzzle and the equity premium puzzle. *Review of Financial Studies 22*, 3367–3409.

Correia, M., J. Kang, and S. Richardson (2014). Asset volatility. *Working paper*.

Cremers, M., J. Driessen, and P. Maenhout (2008). Explaining the level of credit spreads: Option-implied jump risk premia in a firm value model. *Review of Financial Studies 21*, 2209–2242.

David, A. (2008). Inflation Uncertainty, Asset Valuations, and the Credit Spread Puzzle. *Review of Financial Studies 21*(6), 2487–2534.

Davydenko, S. A. (2013). When do firms default? A Study of the Default Boundary. *Working paper*.

Dick-Nielsen, J., P. Feldhütter, and D. Lando (2012). Corporate bond liquidity before and after the onset of the subprime crisis. *Journal of Financial Economics 103*, 471–492.

Duffee, G. R. (1998). The Relation Between Treasury Yields and Corporate Bond Yield Spreads. *Journal of Finance 53*(6), 2225–2241.

Eom, Y. H., J. Helwege, and J.-Z. Huang (2004). Structural models of corporate bond pricing: An empirical analysis. *Review of Financial Studies 17*(2), 499–544.

Ericsson, J., J. Reneby, and H. Wang (2007). Can structural models price default risk? Evidence from bond and credit derivative markets. *Working Paper*.

Feldhütter, P. and D. Lando (2008). Decomposing Swap Spreads. *Journal of Financial Economics 88*, 375–405.

Hamilton, D. T. and R. Cantor (2006). Measuring Corporate Default Rates. *Special Comment. New York: Moody's Investors Services*, 1–16.

Houston, J. and C. James (1996). Bank Information Monopolies and the Mix of Private and Public Debt Claims. *Journal of Finance*, 1863–1889.

Huang, J. and M. Huang (2012). How Much of the Corporate-Treasury Yield Spread is Due to Credit Risk? *Review of Asset Pricing Studies 2*(2), 153–202.

Hull, J., M. Predescu, and A. White (2004). The Relationship Between Credit Default Swap Spreads, Bond Yields, and Credit Rating Announcements. *Journal of Banking and Finance 28*, 2789–2811.

Kealhofer, S., S. Kwok, and W. Weng (1998). Uses and Abuses of Bond Default Rates. *KMV Corporation Document*.

Keenan, S., I. Shtogrin, and J. Sobehart (1999). Historical Default Rates of Corporate Bond Issuers, 1920-1998. *Special Comment. New York: Moody's Investors Services*, 1–84.

Krishnamurthy, A. and A. Vissing-Jorgensen (2012). The Aggregate Demand for Treasury Debt. *Journal of Political Economy 120*(2), 233–267.

Leland, H. (2006). Structural Models of Corporate Financial Choice. *Princeton Lectures in Finance, Lecture 1*.

McQuade, T. J. (2013). Stochastic Volatility and Asset Pricing Puzzles. *Working Paper, Harvard University*.

Merton, R. (1974). On the Pricing of Corporate Debt: The Risk Structure of Interest Rates. *Journal of Finance 29*, 449–470.

Moody's (2013). Annual Default Study: Corporate Default and Recovery Rates, 1920-2012. *Moody's Investors Service*, 1–64.

Nagel, S. (2014). The Liquidity Premium of Near-Money Assets. *Working paper*.

Schaefer, S. and I. Strebulaev (2008). Structural models of credit risk are useful: Evidence from hedge ratios on corporate bonds. *Journal of Financial Economics 90*, 1–19.

Sundaresan, S. (2013). A Review of Merton's Model of the Firm's Capital Structure with its Wide Applications. *Annual Review of Financial Economics 5*.

Vasiček, O. (1991). Limiting Loan Loss Probability Distribution. *Technical Report, KMV Corporation*.

Zhang, B. Y., H. Zhou, and H. Zhu (2009). Explaining Credit Default Swap Spreads with the Equity Volatility and Jump Risks of Individual Firms. *Review of Financial Studies 22*(12), 5101–5131.

	#firms	Mean	10th	25th	Median	75th	90th
				Leverage ratio			
AAA	15	0.14	0.07	0.08	0.09	0.15	0.24
AA	69	0.14	0.05	0.09	0.12	0.17	0.23
A	201	0.28	0.13	0.17	0.24	0.35	0.48
BBB	228	0.38	0.19	0.26	0.36	0.47	0.59
BB	119	0.45	0.18	0.29	0.45	0.59	0.72
B	54	0.51	0.21	0.32	0.47	0.70	0.81
C	7	0.80	0.63	0.73	0.85	0.93	0.96
all	458	0.33	0.11	0.18	0.29	0.44	0.60
				Equity volatility			
AAA	15	0.20	0.16	0.17	0.19	0.23	0.26
AA	69	0.27	0.17	0.21	0.25	0.32	0.38
A	201	0.31	0.20	0.24	0.30	0.37	0.42
BBB	228	0.38	0.24	0.28	0.35	0.44	0.56
BB	119	0.46	0.25	0.31	0.42	0.54	0.71
B	54	0.50	0.29	0.33	0.46	0.68	0.79
C	7	0.69	0.31	0.64	0.72	0.76	1.01
all	458	0.35	0.21	0.25	0.32	0.41	0.54
				Asset volatility			
AAA	15	0.17	0.14	0.15	0.16	0.20	0.24
AA	69	0.23	0.14	0.18	0.23	0.27	0.32
A	201	0.23	0.14	0.17	0.22	0.28	0.34
BBB	228	0.25	0.15	0.19	0.24	0.31	0.37
BB	119	0.28	0.15	0.19	0.25	0.34	0.43
B	54	0.27	0.13	0.19	0.25	0.33	0.42
C	7	0.21	0.04	0.06	0.19	0.31	0.44
all	458	0.25	0.15	0.18	0.23	0.30	0.35
				Payout ratio			
AAA	15	0.034	0.012	0.015	0.027	0.048	0.069
AA	69	0.043	0.017	0.030	0.043	0.053	0.066
A	201	0.047	0.020	0.031	0.043	0.056	0.078
BBB	228	0.050	0.018	0.029	0.045	0.065	0.093
BB	119	0.046	0.020	0.027	0.040	0.057	0.080
B	54	0.045	0.019	0.030	0.040	0.058	0.073
C	7	0.064	0.040	0.049	0.056	0.093	0.101
all	458	0.047	0.019	0.029	0.043	0.059	0.081

Table 1 *Firm summary statistics.* For each bond yield observation, the leverage ratio, equity volatility, asset volatility, and payout ratio are calculated for the issuing firm on the day of the observation. Leverage ratio is the ratio of the book value of debt to the market value of equity plus the book value of debt. Equity volatility is the annualized volatility of daily equity returns from the last three years. Asset volatility is the unlevered equity volatility, calculated as explained in the text. Payout ratio is yearly interest payments plus dividends plus share repurchases divided by firm value. Firm variables are computed using data from CRSP and Compustat.

Short bond maturity

	AAA	AA	A	BBB	BB	B	C	all
Number of bonds	22	71	197	140	60	15	0	447
Mean number of bonds pr quarter	3.17	4.11	9.13	6.34	3.29	1.19	NaN	22
Mean number of quotes pr quarter	6.548	66.46	136.8	112.8	69.71	18.57	NaN	356.7
Age	3.13	5.70	5.58	6.79	6.57	4.34	NaN	6.07
Coupon	7.53	5.65	6.01	7.28	7.67	5.35	NaN	6.58
Amount outstanding ($mm)	287	414	374	367	314	242	NaN	367
Time-to-maturity	1.24	1.23	1.21	1.23	1.27	1.15	NaN	1.23

Medium bond maturity

	AAA	AA	A	BBB	BB	B	C	all
Number of bonds	22	85	256	202	82	22	3	587
Mean number of bonds pr quarter	4.55	4.94	12.9	10.6	5	1.63	1	33.6
Mean number of quotes pr quarter	9.895	80.53	159.9	155.1	99.1	20.81	37	480.8
Age	2.42	4.60	5.00	6.15	5.19	3.98	3.97	5.29
Coupon	7.38	5.79	6.53	7.39	7.62	8.63	11.86	6.94
Amount outstanding ($mm)	278	379	288	349	334	219	207	328
Time-to-maturity	2.98	2.90	2.95	3.03	2.99	3.07	3.44	2.98

Long bond maturity

	AAA	AA	A	BBB	BB	B	C	all
Number of bonds	39	203	666	629	226	69	8	1445
Mean number of bonds pr quarter	7.25	20.3	65.4	51	17.3	4.27	1.64	158
Mean number of quotes pr quarter	36.53	157.1	480.2	473.4	197.9	60.86	64.85	1382
Age	9.73	5.32	5.41	5.28	4.13	5.01	6.85	5.24
Coupon	7.72	6.95	7.44	7.75	7.79	9.01	10.46	7.64
Amount outstanding ($mm)	521	296	281	314	450	261	298	319
Time-to-maturity	13.32	8.32	11.25	11.10	8.34	9.33	10.75	10.43

Table 2 *Bond summary statistics.* The sample consists of noncallable bonds with fixed coupons issued by industrial firms. Short, medium, and long bond maturities are bonds with a maturity of 0-2, 2-4, and 4-30 years. This table shows summary statistics for the data set. Bond yield quotes cover the period 1987Q2-2012Q2.

	1987Q2-2012Q2			1987Q2-1997Q1			1997Q2-2007Q1			2007Q2-2012Q2		
	Bond maturity			Bond maturity			Bond maturity			Bond maturity		
	short	medium	long	short	medium	long	short	medium	long	short	medium	long
A												
Actual spread	39	32	40	16	19	22	23	28	38	103	66	74
Model spread	14	36	51	6	26	20	6	20	50	38	81	111
BBB												
Actual spread	141	117	121	57	63	64	130	93	118	311	253	230
Model spread	109	112	113	31	42	46	45	77	109	341	296	240
spec												
Actual spread	321	537	435	276	469	253	222	576	562	495	588	539
Model spread	219	306	248	86	338	162	173	144	276	476	475	358

Table 3 *Actual and Merton-model yield spreads.* This table shows actual and model-implied industrial corporate bond yield spreads. Spreads are grouped according to remaining bond maturity at the quotation date; 0-2y(short), 2-4y(medium), and 4-30y(long). 'Actual spread' is the average actual spread to the swap rate minus the average actual AAA/AA spread to the swap rate. 'Model spread' is the difference in average Merton model spreads between the bonds in a given maturity/rating bucket and bonds rated AAA or AA. The average spread is calculated by first calculating the average spread across bonds in a given month and then calculating the average of these spreads over months.

Systematic risk	mean				Quantiles			
ρ		0.005	0.025	0.25	0.5	0.75	0.975	0.995
0%	4.39%	4.02%	4.11%	4.29%	4.39%	4.49%	4.68%	4.78%
5%	4.39%	1.72%	2.16%	3.37%	4.21%	5.20%	7.62%	9.01%
10%	4.39%	1.07%	1.50%	2.91%	4.03%	5.46%	9.36%	11.87%
15%	4.41%	0.68%	1.07%	2.55%	3.85%	5.67%	10.84%	14.29%
20%	4.39%	0.45%	0.77%	2.24%	3.64%	5.73%	12.27%	16.76%
25%	**4.38%**	**0.28%**	**0.56%**	**1.94%**	**3.45%**	**5.81%**	**13.50%**	**18.80%**
30%	4.39%	0.18%	0.39%	1.69%	3.25%	5.84%	14.86%	21.17%
35%	4.38%	0.11%	0.27%	1.43%	3.02%	5.86%	16.19%	23.62%
40%	4.40%	0.06%	0.18%	1.23%	2.85%	5.88%	17.44%	26.04%
45%	4.39%	0.03%	0.12%	1.02%	2.59%	5.78%	19.02%	28.39%
50%	4.36%	0.02%	0.07%	0.82%	2.37%	5.73%	20.02%	30.20%

Table 4 *Statistical uncertainty of realized BBB default frequencies published by Moody's.* The table shows the distribution of the realized cumulative 10-year default frequency when the mean cumulative 10-year default rate is equal to 4.39%, the rate for BBB-rated issuers published by Moody's using data for 1970-1998. We assume that in year 1 there is a cohort of 1,000 firms and that in year 2 a new cohort of 1,000 firms formed. This goes on for 18 years. All firms are identical when the cohort is formed and their 10-year default probability is 4.39%. Part of firm volatility is systematic and part is idiosyncratic (see equation (7)). The degree of systematic risk is determined by ρ. For each cohort, we calculate the realized default frequency on a 10-year horizon and then calculate the average default frequency across all cohorts. Overall, the realized default rate is based on 18,000 firms over a period of 28 years. We repeat this simulation 100,000 times and the table shows the distribution of realized default rates for different levels of systematic risk ρ. The results for $\rho = 25\%$ used in the main text are highlighted.

maturity	1	2	3	4	5	6	7	8	9	10
Panel A: True economy (there is variation in leverage ratios)										
average credit spread	38	65	83	94	101	106	109	110	111	112
average default probability	0.37	1.10	1.87	2.60	3.26	3.85	4.36	4.81	5.20	5.55
asset volatility	24.0	24.0	24.0	24.0	24.0	24.0	24.0	24.0	24.0	24.0
Panel B: Representative firm (average leverage ratio used)										
credit spread	0	6	23	42	59	72	82	89	95	99
default probability	0.00	0.07	0.39	0.90	1.50	2.11	2.71	3.26	3.76	4.22
asset volatility	24.0	24.0	24.0	24.0	24.0	24.0	24.0	24.0	24.0	24.0
Panel C: Representative firm, average def. prob. at bond maturity is matched										
credit spread	44	75	94	106	113	117	120	121	122	122
default probability	0.37	1.10	1.87	2.60	3.26	3.85	4.36	4.81	5.20	5.55
implied asset volatility	38.7	32.9	30.4	29.0	28.1	27.4	26.9	26.5	26.1	25.9
Panel D: Representative firm, average def. prob. at 10-year bond maturity is matched										
credit spread	0	13	39	63	82	96	106	114	119	122
default probability	0.00	0.16	0.68	1.42	2.23	3.02	3.75	4.41	5.01	5.55
implied asset volatility	25.9	25.9	25.9	25.9	25.9	25.9	25.9	25.9	25.9	25.9

Table 5 *Convexity bias when calculating credit spreads in the Merton model using the representative firm approach.* It is common in the literature to compare average actual spreads to model-implied spreads, where model-implied spreads are calculated using average firm variables. This introduces a bias because the spread in the Merton model is a non-linear function of firm variables. This table shows the magnitude of this bias. Panel A shows, for maturities between one and 10 years, the average spread and default probability for 100,000 firms that have different leverage ratios but are otherwise identical. Their common asset volatility is 24%, payout rate 4.5%, and recovery rate 37.8%. Their leverage ratios are simulated from a normal distribution with mean 0.36 and standard deviation 0.16 (truncated at zero). The riskfree rate is 5%. Panel B shows the spread and default probability of a representative firm where the average leverage ratio of 0.36 is used. In Panel C, for each maturity – one at a time – an asset volatility is computed such that, for a representative firm with a leverage ratio of 0.36, the default probability is equal to the average default probability in the economy (given in the second row in Panel A and again in Panel C). This is done separately for each maturity. The panel shows the resulting implied asset volatility and spread. Panel D shows the results of a calculation similar to that in Panel C except here the asset volatility used to compute the spread for each maturity is the value that matches the average *10-year* default probability in the economy.

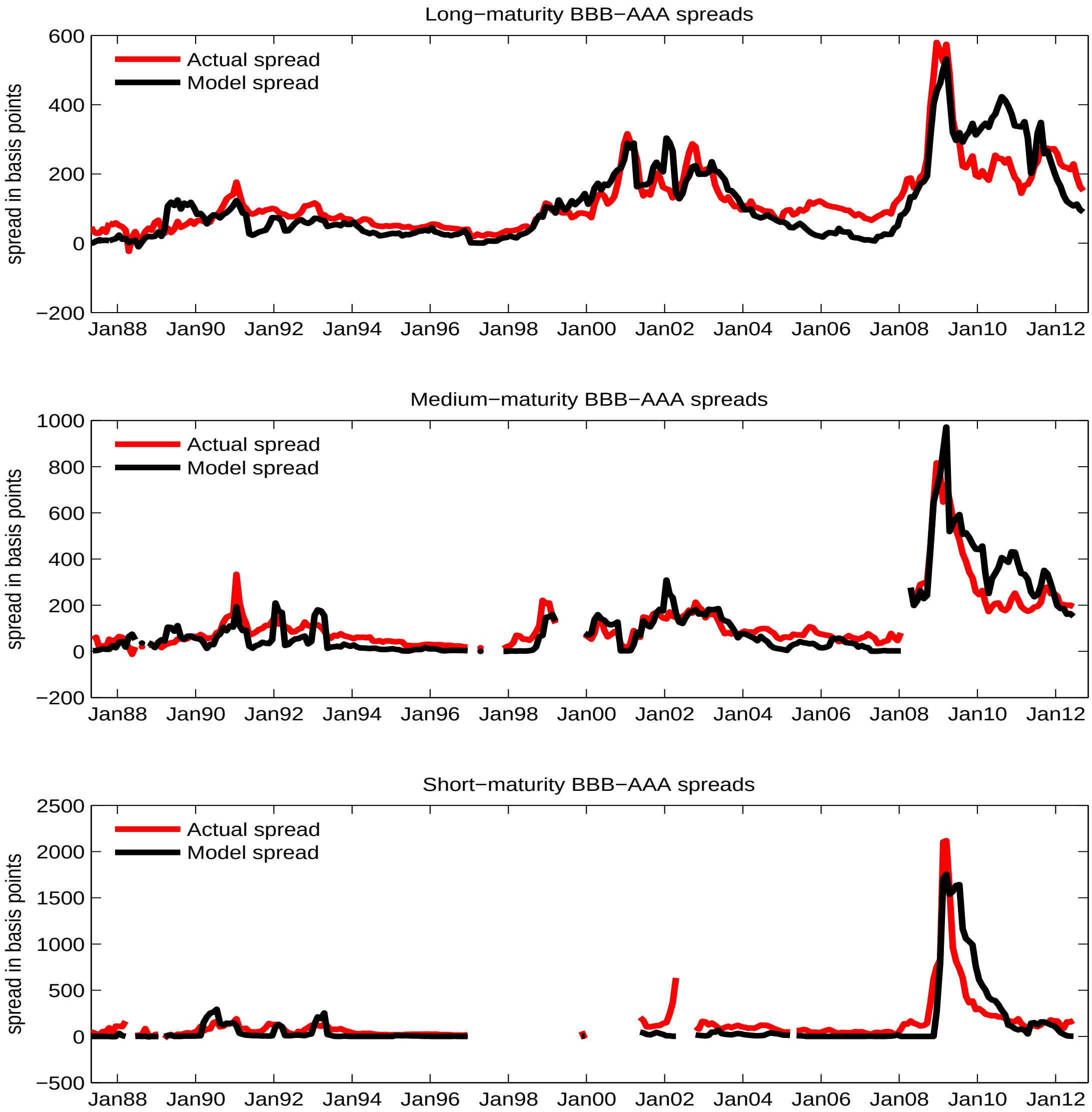

Fig. 1 *BBB-AAA corporate bond yield spreads.* This graph shows the time series of actual and model-implied BBB-AAA spreads. Each month all daily yield observations in bonds rated AAA/AA and bonds rated BBB are collected, and the graphs shows the average BBB spread to the swap rate minus the average AAA/AA spread to the swap rate. The results are shown for maturities between 4 and 30 years (long-maturity), 2 and 4 years (medium-maturity), and 0 to 2 years (short-maturity). The figure also shows the model-implied Merton spread, found by calculating the model-implied AAA-BBB spread computed in the same way as the actual spread.

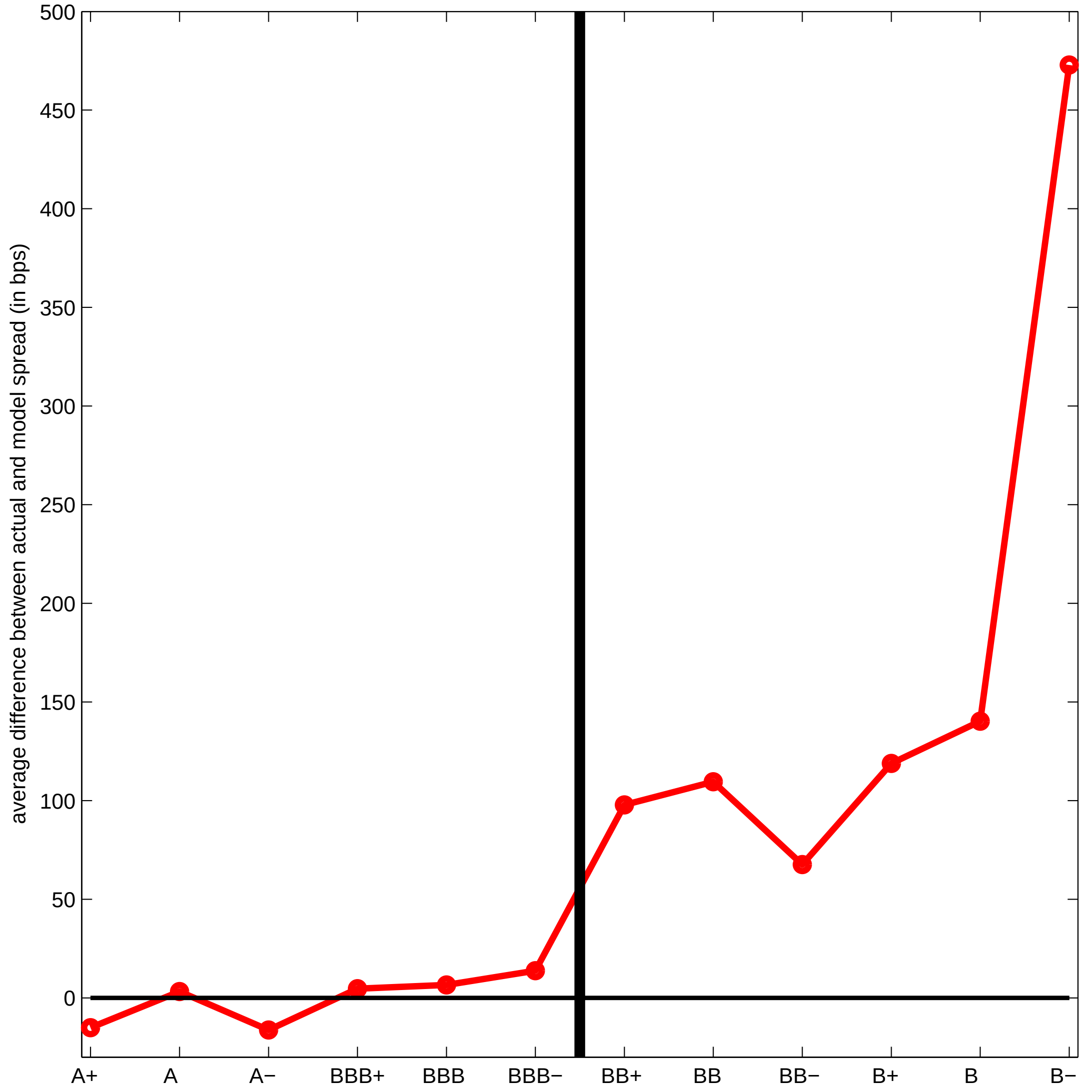

Fig. 2 *Average difference between actual and Merton-model credit spreads for rating notches.* On a monthly basis, the average actual yield spread to the swap rate for a given rating notch is calculated over all observations in that month with a maturity between 0 and 30 years and with a current rating corresponding to this rating notch. The overall average spread is calculated by calculating the average of monthly spreads. This is done for Merton model-implied spreads as well and the graph shows the difference (actual minus model) for each rating notch. Ratings to the left of the vertical line are investment grade ratings while those to the right are speculative grade.

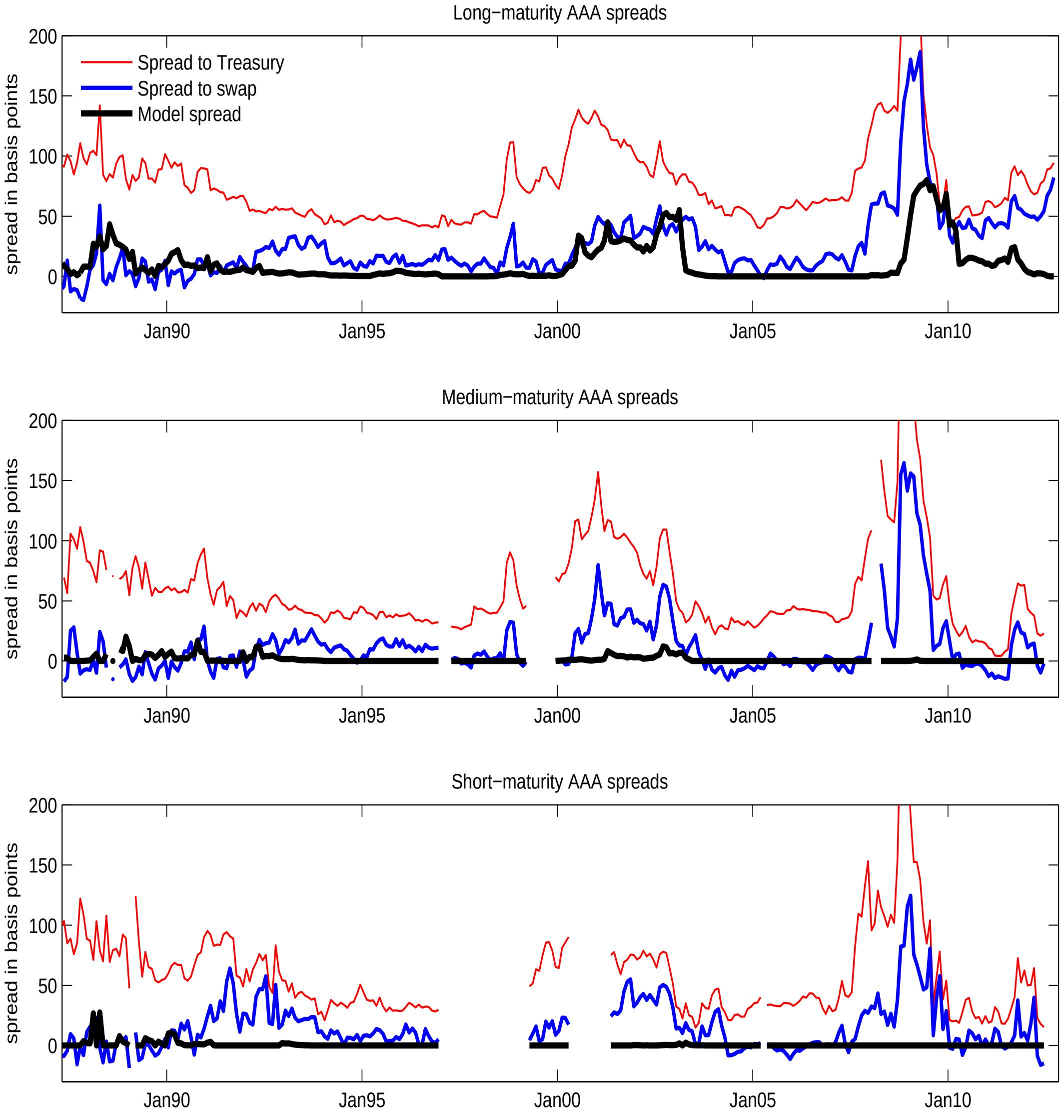

Fig. 3 *AAA corporate bond yield spreads.* This graph shows the time series of actual and model-implied AAA spreads. Each month all daily yield observations in bonds rated AAA or AA are collected (we call this the AAA yield) and the graphs shows the average AAA spread to (i) the swap rate and (ii) the Treasury yield. The results are shown for maturities between 4 and 30 years (long-maturity), 2 and 4 years (medium-maturity), and 0 to 2 years (short-maturity). The figure also shows the model-implied Merton spread.

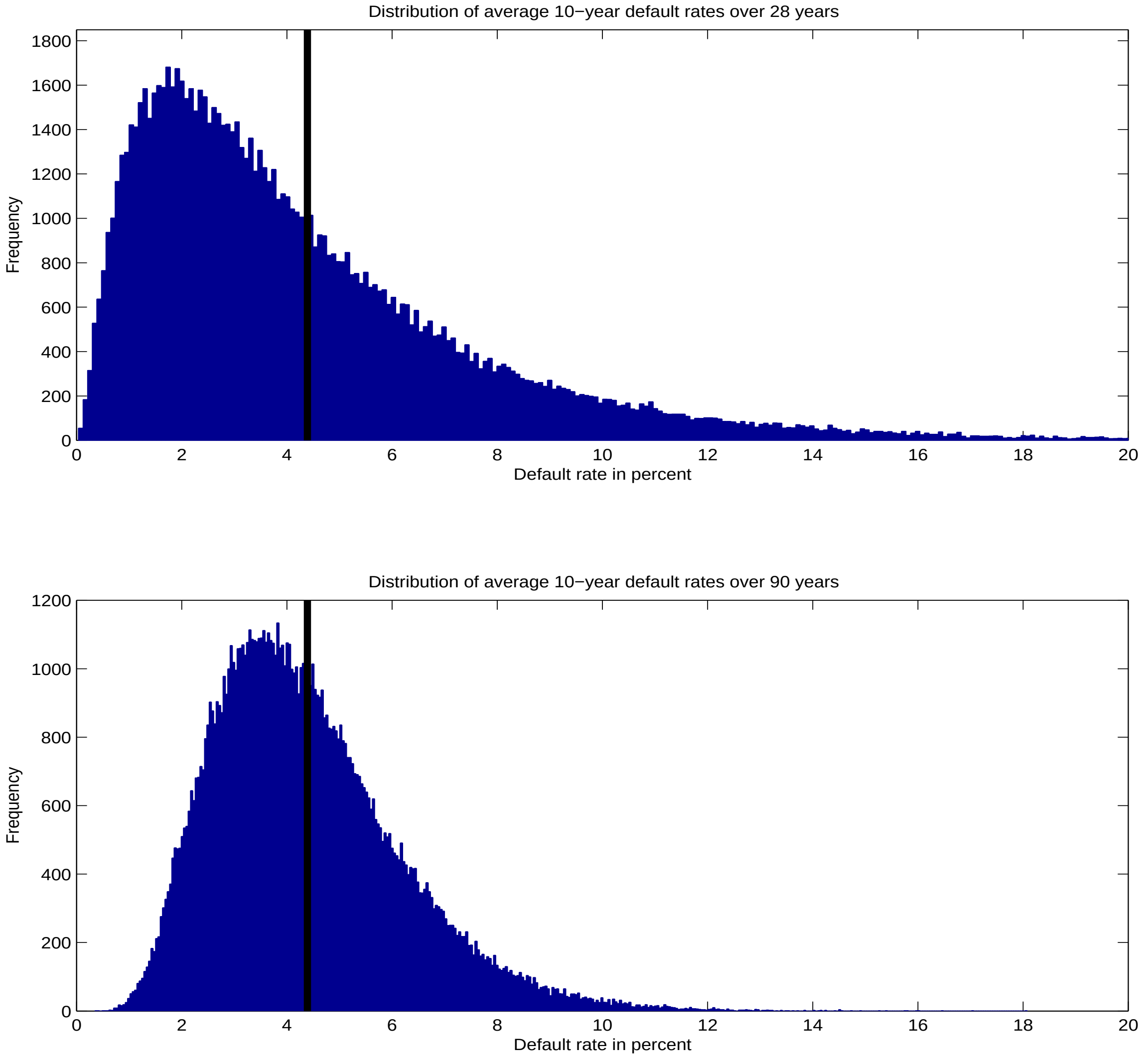

Fig. 4 *Distribution of realized default rate.* The figure shows the results of a simulation of the realized cumulative 10-year default rate when the mean cumulative 10-year default rate is equal to 4.39%, the rate for BBB-rated issuers published by Moody's using data for 1970-1998. We assume that in year 1 there is a cohort of 1,000 firms and that in year 2 a new cohort of 1,000 firms formed. This goes on for 18 years. All firms are identical when the cohort is formed and their 10-year default probability is 4.39%. Part of firm volatility is systematic and part is idiosyncratic (see equation (7)). The degree of systematic risk is determined by ρ. For each cohort, we calculate the realized default rate on a 10-year horizon and then calculate the average default rate across all cohorts. Overall, the realized default rate is based on 18,000 firms over a period of 28 years. We repeat this simulation 100,000 times and the top graph shows the distribution of realized 10-year cumulative default rates. The solid line is the ex ante default probability of 4.39%. The bottom graph shows the distribution when the realized default rate is based on 90 years (but the ex ante default probability remains equal to 4.39%.

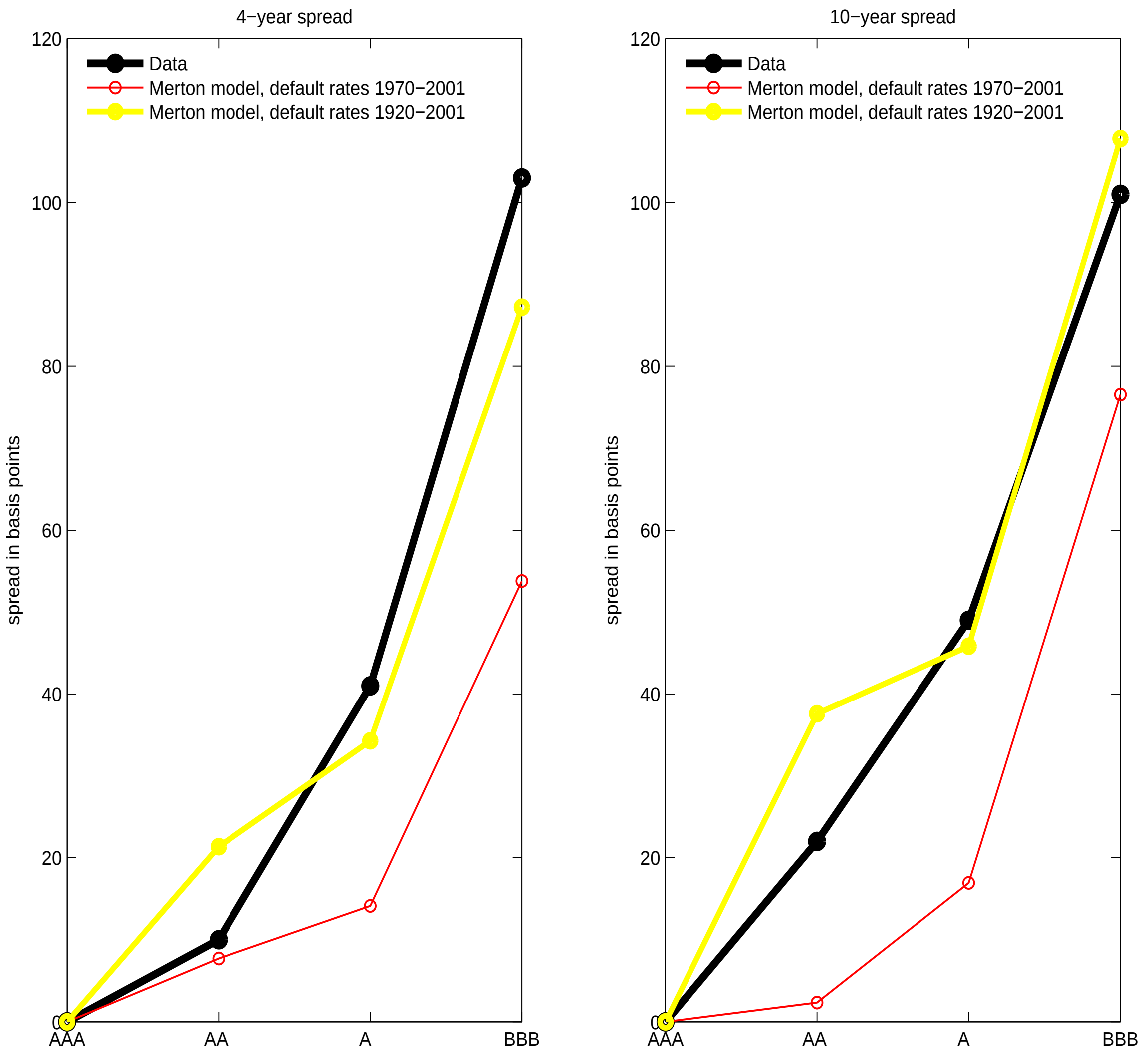

Fig. 5 *Corporate spreads to AAA-bond yields in the Merton model using default rates from different periods.* This figure shows actual and model-implied spreads to AAA yields. The thin red line shows spreads in the Merton model based on Moody's default rates from the period 1970-2001 and corresponds exactly to the calculations in Chen, Collin-Dufresne, and Goldstein(2009). The thick yellow line shows spreads in the Merton model where Moody's default rates from the period 1920-2001 are used. Actual spreads are from Duffee(1998).

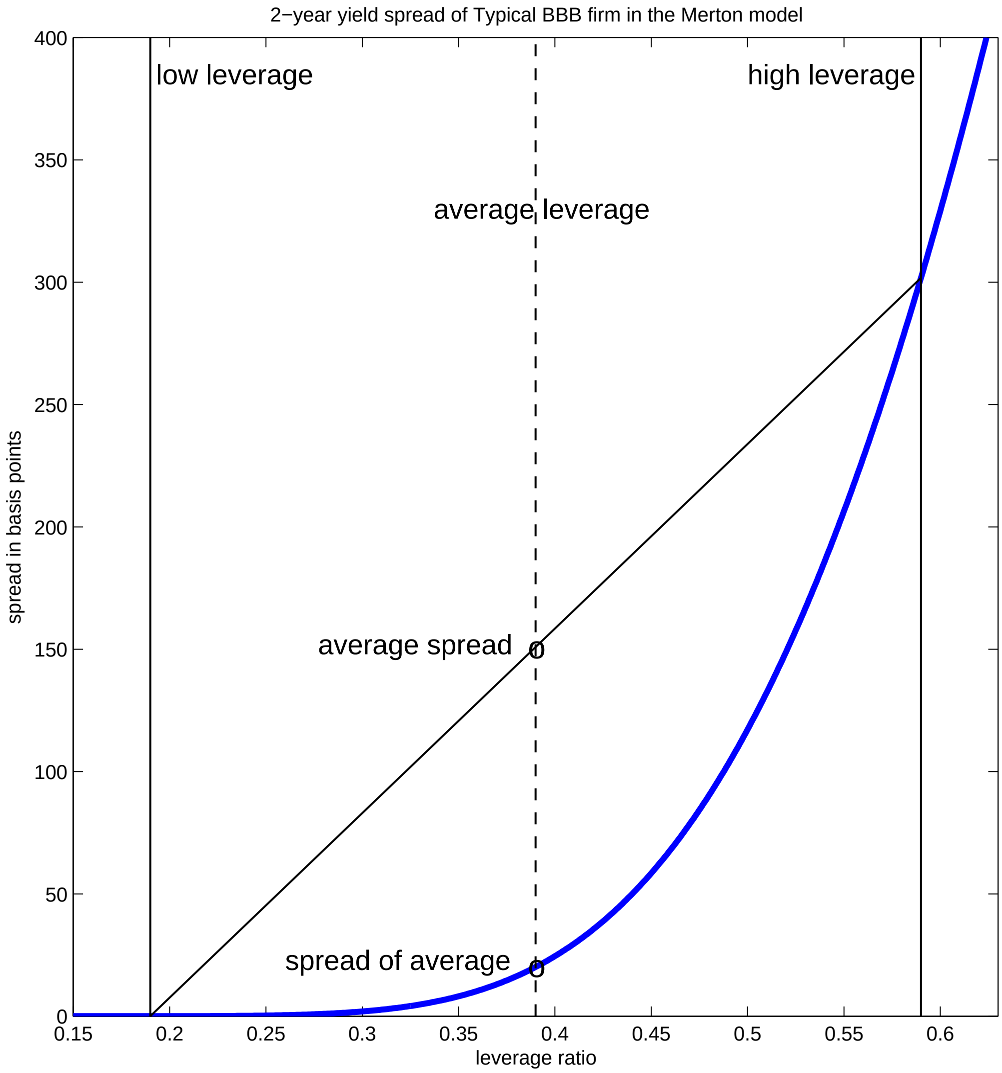

Fig. 6 *Convexity bias when calculating the spread in the Merton model using average leverage and comparing it to the average spread.* It is common in the literature to compare average actual spreads to model-implied spreads, where the latter are calculated using average firm variables. This introduces a bias because the spread in structural models is a non-linear function of firm variables. The figure illustrates the bias in case of two BBB-rated firms, one with a low leverage ratio and one with a high leverage ratio. Asset volatility is 25%, dividend yield 5.0%, recovery rate 37.8%, and riskfree rate 5%. 'Low leverage' is the 10% quantile (0.19) in the distribution of BBB leverage in our sample while 'high leverage' is the 90% quantile (0.59). The quantiles are from Table 1.