Characterizing Credit Spreads

Jeffrey R. Bohn
Haas School of Business
University of California
Berkeley, CA 94720
Tel: 415-352-1279
Fax: 415-296-9458
Email: jeff.bohn@kmv.com

April 1999
Revised: June 1999
Working Paper: Comments welcome.*

Abstract

This paper characterizes credit spreads for corporate bonds reflected in a large and comprehensive dataset from Bridge/EJV. While the data are clearly noisy, robust measures of central tendency combined with graphical analysis produce term structures of credit spreads that conform with the qualitative predictions of Black and Scholes (1973) and Merton (1974) (BSM). The theoretical prediction of BSM regarding lower credit quality firms has been controversial. Their model suggests lower credit quality firms will exhibit humped-shape or downward sloping credit spread term structures. While some empirical researchers (Sarig and Warga (1989)) confirm this feature of

*This paper is a modified version of the fifth chapter of my dissertation, *Risky Debt Valuation and Credit Spreads*, submitted to the University of California, Berkeley in Spring, 1999. I am particularly grateful to Stephen Kealhofer, Mac McQuown, Oldrich Vasicek, and Peter Crosbie for their comments and generous help in securing data. I am also indebted to Terry Marsh, Mark Rubinstein, Tom Rothenberg, and Francis Lim for their insights. Any errors are, of course, my own.

the BSM modeling framework, other researchers (Helwege and Turner (1998)) dispute this feature arguing it is an artifact of the method of constructing the term structures. Comparing results using both agency ratings and EDF^{TM} (expected default frequency: a measure provided by KMV Corporation which provides an estimate of a firm's expected default probability over a specific time horizon), I present a resolution to this controversy. Properly controlling for credit quality in the cross-section, I show that low credit quality issuers tend to exhibit humped-shape and downward sloping credit spread term structures.

1 Introduction

> *"Bonds are different. Although the roughly \$350 billion of bonds traded each day in the U.S. dwarfs the \$50 billion of stocks that change hands, almost all bond trading is done 'over the counter.' Someone who wants to buy or sell a bond calls a broker and asks for a price quote. The broker is free to name virtually any price, because there is no effective reporting mechanism that an investor can turn to for information about the latest trade in a particular bond."* Zuckerman (September 10, 1998)

Corporate bond pricing presents some intriguing and irritating challenges. Pricing data come either in large quantities of dubious quality or small quantities of only reasonable quality. The fact that many corporate bonds do not trade much results in a large number of "matrix" prices where dealers use simplistic algorithms or matrices to price an issue that did not actually trade. While we have many theories for valuing corporate bonds, we have few empirical results to guide us toward truth in valuation. (See Bohn (1999) for an overview.) This paper reports results gleaned from a large database maintained by Bridge/EJV. While transacted prices are mixed with matrix prices, I have gone to great lengths to dampen the impact of the spurious prices. The filtering techniques will be discussed in this paper. Initially, I will characterize the data to demonstrate general patterns in credit spreads as well as demonstrate the noisiness in the data. This paper primarily describes data. Further analysis will be presented in a subsequent paper testing a particular specification of a model for risky debt valuation.

Section 2 describes previous corporate bond data used by other financial researchers. Section 3 presents a description of the EJV sample used to conduct the research reported in this paper. Next, section 4 presents a first look at the risky term structure. Sections 3 and 4 explain the essence of the patterns apparent in these data. Section 5 specifically addresses the question of a downward or humped slope for the credit spread term structure of non-investment grade issuers. Section 6 summarizes the results from this paper.

2 Previous Data Descriptions

Until the last decade, data on corporate bonds were generally not available in any quantity. Jones, Mason, and Rosenfeld (1984) (JMR) produced a dataset of 27 firms covering each firm's publicly traded debt monthly from January, 1975 to January, 1981. They find a contingent-claims model in the spirit of Black and Scholes (1973) and Merton (1974) (BSM) does not explain actual prices well. The predicted spreads tend to be substantially less than the actual spreads. In the last ten years, a handful of papers have appeared describing new corporate bond datasets. Unfortunately, these data have raised more questions then they have answered. Sarig and Warga (1989) (SW) present a picture of the term structure of credit spreads based on monthly zero-coupon prices collected from Lehman Brothers (LB: see Arthur Warga's website for more information at www.uh.edu/˜awarga) for the period February, 1985 through September, 1987. Their dataset covers 137 corporate issues from 42 different companies. They use U.S. treasury strips to derive a reference curve proxying for the risk-free term structure to calculate credit spreads. Averaging the spreads across bonds (cross-sectionally) each month and then averaging the spreads across time results in a term structure qualitatively similar to the one predicted by BSM. (See Lee (1981) and Pitts and Selby (1983) for refined, graphical presentations of BSM's theoretical term structure of credit spreads.) SW demonstrate a key feature of the BSM framework with the humped and downward sloping curves for low credit quality issues. The small sample size prevented SW from presenting any conclusive statistical evidence. The other difficulty with this sample is the focus on zero-coupon bonds. While many credit spread calculation issues are resolved (e.g. duration equals maturity for zero-coupon bonds eliminating potential problems surrounding tenor) with zero-coupon bonds, they are not representative of the typical bond issued by most corporations. Consequently, this subset of the corporate bond universe may exhibit biases related to the type of firm that issues zero-coupon bonds. Nonetheless, SW provided a response to JMR and revived the applicability of the BSM framework to risky debt pricing.

As counterpoint to SW, Helwege and Turner (1998) (HT) find upward sloping term structures when looking at 64 offerings of 163 non-investment grade bonds issued between 1977 and 1985. They use data from both Securities Data Company (SDC) and LB. Their contribution lies in holding credit quality constant by analyzing multiple bonds of the same company with the same seniority. They argue that downward sloping term structures result from sample selection bias where bonds within a particular credit grade are not actually of equal credit quality. Said differently, a particular non-investment grade, say B, contains "good" Bs and "bad" Bs. The good Bs issue longer dated debt (because they are more creditworthy) while the bad Bs can only issue short-dated debt. The result is a downward bias for a particular non-investment grade category. They consider (and reject) two possible explanations that can be used to reconcile their results and the BSM model. Recall that the parameter in the BSM framework driving the slope of the credit spread term structure is d, also known as the quasi-leverage parameter. If we define F as the face value of the firm's debt, r as the risk-free rate, T as the debt's maturity, and V_A as the market value of the firm's assets, quasi-leverage is calculated as follows:

$$d = \frac{Fe^{-rT}}{V_A}$$

For the slope of the credit-spread term structure to be negative, d must be equal to, or greater than, one. The bonds in HT's sample may not have been sufficiently risky in terms of leverage. They report typical leverage ratios (using book leverage as a proxy for quasi-market leverage) less than one for their sample of non-investment grade bonds. One explanation of their results is that firms with quasi-market leverage exceeding one do not issue publicly traded debt. Their second explanation extends this argument surrounding leverage to the question of relative leverage. They concede that a few firms do exhibit downward sloping term structures. If leverage is the driver of term structure shapes, then firms with downward sloping term structures should, in fact, have higher leverage than firms with upward sloping term structures. They look at the few firms in their sample that do exhibit downward sloping credit term structures to see if their leverage is significantly larger than that of the other firms in the sample. They find that the firms with downward sloping term structures typically have less leverage than the other firms in their sample.

While other modeling frameworks (e.g. Jarrow, Lando, and Turnbull (1997)) can generate downward sloping term structures for low credit quality firms without relying on leverage as the trigger, HT present important counter-evidence to the picture presented by SW. Again, small sample sizes coupled with the ambiguity concerning rating agencies' accuracy (ratings may not be an accurate or consistent measure of credit risk) prevents strong acceptance of these results. However, these results cannot be ignored when testing other samples.

3 Data Description

The data for this analysis were provided by Bridge. While Bridge distributes this data, the original source is a partnership of fixed income dealers and brokers, called EJV (Electronic Joint Venture). Bridge assembles and maintains the fixed income database generated by EJV. The data cover provisions in bond indentures, current and historical credit ratings, industry classification for the bond issuers, price and yield history, amount outstanding, and relevant information about bond issuance. My analysis focused on bonds with the following criteria:

1. Issued by a corporation without any convertibility provisions.

2. Denominated in U.S. currency (but not necessarily issued by a U.S. company).

3. Not part of a unit (e.g. a bond and warrant sold together).

4. Rated by Standard and Poors.

Using these criteria, I obtained a sample of 24,465 bonds issued by 1,749 firms. These issues were tracked monthly from June, 1992 to January, 1999. Each bond had between 1 and 80 observations resulting in a sample size of almost 600,000. Each observation consisted of an indicative price and the issue's worst-yield (minimum of yield-to-worst[1] and yield-to-best[2]). EJV describes the source of these prices as follows:

> These prices reflect where the market closed as of 3PM for each trading day on the bid side. Our pricing process takes in as input price 'runs' from our 'partner' firms, i.e. Salomon Brothers, Goldman Sachs, CS First Boston, Lehman Brothers, Liberty Brokerage, Morgan Stanley...etc.[3]

Unfortunately, many of the data were matrix prices (rather than transacted or quoted prices) received from dealers who regularly trade the issue being priced. Because matrix prices are not traded prices (dealers use proprietary algorithms to generate matrix prices which are only theoretically the price at which the issue would trade on that day), some bias can appear due to factors unrelated to the issuer. Because I looked at medians, I did not worry too much about the presence of these matrix prices. Taking medians in these cross-sections will eliminate much of the irrelevant information injected with matrix prices. While this approach will work for characterizing large quantities of data, future work with individual data will need to include more elaborate measures to eliminate biases introduced by matrix pricing. Roughly half the issues in the sample were from financial firms while well over half the firms represented were non-financial.

3.1 Distributions of the Sample

Given that I temper the noise in the data by analyzing medians, the distribution of data across dimensions relevant to risky debt valuation becomes essential. This section reviews the distribution of this sample in this light.

[1]Yield-to-worst is a proxy for the option adjusted yield for a callable bond. The possibility that the bond may be called makes the actual maturity of the issue uncertain. Yield-to-worst is calculated by taking the minimum of all the possible yields (each possible yield corresponds with the maturity equal to the different call dates). The assumption behind this calculation is the buyer of the option (in this case the borrower) will optimally exercise the option.

[2]Yield-to-best applies to putable bonds. In this case, the assumption is the owner of the issue will optimally exercise the put option and find the best yield. Yield-to-best is the maximum of the possible yields calculated by setting the maturity equal to the different possible put dates.

[3]This information is taken from the documentation accompanying the EJV data.

The implicit assumption woven in to the fabric of this research is credit quality primarily drives a debt issue's credit spread. Consequently, the sample needs to have reasonable data density across each credit class. Let us turn to a number of figures illustrating this distribution. Figure 1 presents the sample of firms stratified by S&P credit rating using the December, 1998 rating. This distribution resembles the distribution found on any particular date of the sample. The seven general rating classes are defined as follows:

1. AAA
2. AA+, AA, AA-
3. A+, A, A-
4. BBB+, BBB, BBB-
5. BB+, BB, BB-
6. B+, B, B-
7. CCC+, CCC, CCC-

Using the December, 1998 rating for each issue, figure 2 shows the distribution of the issues in the sample across rating classes divided between bonds issued by financial and non-financial firms. Rating class three has the largest number of issues followed by rating class four. As would be expected, the non-financial group has a relatively more even distribution across rating classes while the group of financial firms have a relatively small number of issues in the lowest rating classes. Earlier attempts at fitting these data relied on stratification of S&P credit class producing reasonable results. The results in this paper now benefit from the laborious task (accomplished by KMV) of mapping the issues by CUSIP into an identifier allowing me to match an issue with its term structure of expected default frequencies (EDF^{TM}). [4] The EDF measure is an expected default probability over a particular time horizon (e.g. one year, two years, etc.) for the issuer. This mapping facilitates the construction of credit classes that better group together homogeneous (in terms of credit quality) issuers. A subsequent section will discuss this point

[4]See appendix A for a description of the KMV EDF. This measure reflects an estimated default probability for the issuer. An EDF is available for time horizons one through five. EDFs for other horizons are interpolated.

in greater detail. Evidence supporting the strength of EDF as a measure of credit quality will also be presented. For now, if one accepts that EDF accurately reflects credit quality, the distribution of issues by one-year EDF provides a more even distribution of data density across credit classes. The sample was divided into nine credit classes on each date as follows:

1. .02%
2. .02% to .04%
3. .04% to .08%
4. .08% to .16%
5. .16% to .32%
6. .32% to .64%
7. .64% to 1.28%
8. 1.28% to 20%
9. 20%

KMV truncates the probabilities at .02% and 20% due to estimation difficulties at the extremes of the distribution. Issuers at the extremes were placed in their own classes to minimize their influence on the medians in the nearby credit classes. Figure 4 presents this distribution. (For reference, figure 3 presents the distribution of issuers.) Notice that a large number of issues fall into the first class. Despite this potential bias, the larger numbers in the lower credit classes (high EDFs) makes this distribution more amenable to fitting a model across credit classes. A drawback of using one-year EDF to classify credit quality lies in the exclusion of information regarding expectations about how a firm's EDF might evolve. Since KMV publishes EDFs out to five years, I looked also at the distribution of the sample using the geometric mean of the five EDFs (one-year EDF through five-year EDF) available from KMV. Figure 6 presents the distribution of the sample issues using the geometric mean. The ranges for the average (geometric mean) EDFs (AEDF) are as follows:

1. Under .04%

2. .04% to .10%
3. .10% to .18%
4. .18% to .25%
5. .25% to .35%
6. .35% to .60%
7. .60% to 2%
8. 2% to 20%

Classification by AEDF provides a more inclusive measure of credit quality (i.e. includes some information about the expected evolution EDF) that is also more disperse. In other words, the AEDF does not have undue concentration in any particular range. Specifically, the one-year EDF stratification results in a large percentage of issues placed in the lowest EDF ranges (highest credit quality). This concentration is a consequence primarily of using only the one-year horizon in classifying the issue. By definition, firms with little debt have nothing to default on. The resulting EDF will be lower than the actual ability of the firm to carry new debt. Said differently, the firm's sensitivity to new debt may be quite high. The one-year EDF will not necessarily reflect this characteristic of the firms. Classifying by AEDF creates classes that have firms of greater similarity in terms of credit quality when looking out over a longer time horizon. Figure 5 shows the distribution of firms issuing these bonds. Partitioned across credit classes, the non-financial firms appear to be more uniformly distributed than the financial firms. It should be noted that I have removed the financial subsidiaries (e.g. GMAC) given the difficulty of assigning a default probability to these entitities.[5] A noticeable characteristic of this sample (and of the corporate bond market more generally) is the small number of firms who issue the bulk of traded corporate debt. Over one-third of the issues in the sample were issued by 35 firms. Table 1 lists these firms and the number of their issues in the sample. Looking at these distributions, the sample appears to characterize a broad cross-section of the corporate bond market (denominated in U.S. dollars) with no systematic bias along any particular aspect of a bond issue.

[5] I am indebted to Nancy Wallace for pointing out the problems with analyzing financial subsidiaries in the same framework as typical corporates.

After partitioning the sample into credit classes, I considered the questions of term and coupon. Ideally, one would want to use the maturity of each coupon bond and accurately match a similar default-free security to arrive at a credit spread. While some modeling approaches allow for cutting up a coupon bond into a portfolio of mini-discount bonds, complications arise when trying to look at groups of bonds in particular buckets (date, credit class, time to maturity). In this paper, my objective is to find general patterns of central tendency so I chose a simpler approach to grouping bonds across the term dimension. Macauley duration was calculated for each issue essentially transforming the bond into a discount bond with time-to-maturity equal to duration. Note that in the case of callable and putable bonds, a worst-yield Macauley duration was calculated where the maturity used in the calculation was adjusted to the worst-yield maturity. In this way, I reduced the complexity of the analysis while retaining the essence of the information regarding the tenor of the security.

3.2 Options

The next issue to tackle concerns the presence of options. Figure 7 presents the distribution of issues in the sample segregated by option embedded and non-option embedded bonds. Notice that two-thirds of the issues in the lowest credit class (class 8) have options. By focusing only on non-option bonds, the cost is underrepresentation in the lowest credit classes. This bias will likely affect the results. To avoid this bias, I have chosen to use worst-yield spreads as a proxy for option adjusted spreads (OAS). To minimize the problem of having potential specification problems with both the valuation model and the OAS model, I am focusing on worst-yield spreads. In this way, I eliminate a source of modeling uncertainty. Again, the fact I work with medians resolves many of the issues associated with extreme outliers (e.g. a long-dated bond with a call option expiring within one or two-years will potentially result in a worst-yield spread that is significantly off the mark) when using a measure like worst-yield spread.

Along the term dimension, duration is calculated with the maturity date equal to the worst-yield maturity. In other words, a bond with an option that expires prior to actual maturity will have duration calculated as if its maturity is equal to the option's expiration date (assuming that date is associated with the worst-yield). Bonds without options will have their duration equal to their worst-yield duration. Again, I could have used option adjusted duration (OAD) and have, in fact, done some work with this measure. However, I avoid the added layer of modeling uncertainty by using the worst-yield duration. Qualitatively, the results do not change much when I substitute suitable models for OAS and OAD instead of worst-yield spread and worst-yield duration (the results using OAS and OAD are available upon request).

Because the bond spreads behaved erratically for issues with duration less than a year (probably affected significantly by the presence of call and put options written into some indentures), the term structure analyzed began at year one. The first bucket contained bonds with durations between one and two years. At the long end of the term structure, lack of data prevented reasonable inference. Consequently, the term structure studied begins at one year and ends at 12 years. The final bucket contained bonds with durations between 11 and 12 years.

3.3 Noise

This research effort focused on subsetting the data into homogeneous groups and calculating medians to overcome the presence of noise in the data. As mentioned by Chairman Levitt, price data for bonds are less than transparent. The inclusion of matrix or evaluated prices in the construction of a sample to test pricing models introduces noise arising from modeling issues, measurement issues, and synchroneity problems (i.e. the recorded price may not actually be the price on the day or time associated with the price in the dataset; this problem can then spill over into properly matching the dates of the default probabilities and the price.) To demonstrate the extent to which noise permeates these data, I constructed histograms and box and whisker plots across dimensions that group together reasonably homogeneous bonds. The dimensions to consider are as follows (listed in order of importance):

1. Credit Quality
2. Duration
3. Priority in the Capital Structure
4. Options

Since I am focused on worst-yield spreads, I have mostly controlled for the impact of options. The exceptions lie in bonds with duration less than one year. These issues' spreads tend to behave erratically so I have eliminated them from the sample. It should be noted also that standard BSM models have difficulty generating non-zero spreads for short duration bonds given that the uncertainty derives from the asset value following a diffusion process. (Diffusion processes cannot move very far over short time horizons reducing the probability of default almost to zero.) This weakness of the BSM modeling framework has given rise to models incorporating jump processes. (See Bohn (1999) for a survey of risky debt models.) This extension of including a jump process is beyond the scope of my research so focusing on issues with duration greater than one year makes sense from this perspective as well.

Consider next priority in the capital structure: I am interested in understanding the impact on loss given default (LGD) and then by extension the impact on credit spreads. Unfortunately, little is known about LGD. In theory it would seem to be an important determinant of credit spreads; in practice it is difficult to assess its impact. Altman and Kishore (1996) looked at 700 defaults (occurring between 1978 and 1995). They documented LGD across a number of dimensions including seniority and industrial classification. They find LGD unrelated to the original rating of the bond once seniority is taken into account. They find also that LGD differs across a few industries and differs across seniority. LGD ranges from 45% for senior secured issues to 68% for subordinated issues. The standard deviations, however, are quite large. For example, senior unsecured bonds experience an average LGD of 52% with a standard deviation of 31%. One other study completed by Carty, Keenan, and Shtogrin (1998) confirms the averages reported by Altman and Kishore (1996) as well as the dispersion in results. They report high interquartile ranges for each level of seniority. This volatility in LGDs complicates the picture as investors cannot easily determine, ex-ante, an accurate LGD. The average LGD, in this case, can be misleading. Herein lies another source of noise in the credit spread data. As far as this sample is concerned, the distribution across priority is nearly uniform. The number of senior issues approximately equals the the number of subordinated issues. One difficulty with the data lies in the large number of issues of undetermined priority. For example, a large number of the issues are classified as unsecured notes. No reference is made to these issues' priority. For most tests in this paper I have simplified the LGD specifications and consider either one rate of 45% or three rates of 45%, 50%, and 65% (corresponds with senior, undetermined, and subordinated respectively). This area of research deserves considerably more attention. The difficulty (as always) is lack of data.

Let us now consider some broad brush measures of dispersion within these classes enumerated above. Figure 8 illustrates the dispersion of data at different points of time for a reasonably homogeneous class of issues. (In this case the class includes senior bond issues from AEDF class 2 and duration class 3 which reflects durations from 3 to 4 years.) As might be expected, the data are more disperse for the lower credit grades than for the higher credit grades. Compare the previous graph with figure 10 to see the increasing dispersion within the grouping that is reasonably homogeneous. As has been discussed before, the heterogeneity in the data results from the lack of transparency and disclosure in the corporate bond market.

Surprisingly, figures 9 and 11 demonstrate that subordinated issues in this sample tend to be more homogeneous than senior issues. The reason for this counterintuitive result likely lies in the stratification. Using AEDF to stratify the sample creates more homogeneous groups in the lower credit classes and slightly less homogeneous groups in the higher credit classes. See table 2 for a schedule of the average standard deviation across buckets in each credit class for both AEDF stratification and S&P stratification. A bucket is created for each date by including issues of similar credit quality and similar duration. The standard deviation of the worst-yield spreads in the bucket is calculated for each date. The average over all buckets for a particular credit class is recorded. Notice that one column in the table reflects the results of using S&P ratings as the criteria for assigning credit class while the other column reflects using AEDF. Since the AEDF measure reflects the information in the traded equity market, these results are not surprising. Investment grade debt trades less frequently and is more likely to reflect S&P ratings. High-yield debt (i.e. low credit quality), on the other hand, trades more frequently making it more likely to reflect pricing consistent with observations in the prices of the firm's traded equity. Figures 12, 13, 14, 15, and 16 present the histograms for the standard deviations averaged in table 2. The differences in the dispersion between S&P stratification and AEDF stratification can be seen in these figures.

4 A First Look at the Risky Term Structure

Before turning to the estimation of this paper's model, an overview of the term structure implied in the data will be useful. Sarig and Warga (1989) used zero-coupon bonds to calculate a term structure of credit spreads by taking the average spread cross-sectionally in each credit class and then averaging the time-series of cross-sectional averages. Using the EJV data, I generated a similar picture looking at both cross-sectional averages and cross-sectional medians averaged over time. Both graphs are strikingly similar to the one generated by Sarig and Warig. Figure17 presents the cross-sectional medians averaged over time. Note that this graph presents the data stratified by S&P rating. As reported by Sarig and Warga, the qualitative shape of the credit spread curves match the predictions made by Merton (1974). The use of more data sharpens the results and sets the stage for successful implementation of a BSM-type model. Figure 18 presents the data stratified by AEDF. The creation of more homogeneous credit quality groups using the EDF measure results in a smoother term structures. Figures 19, 20, 21, and 22[6] present two snapshots from the dataset: cross-sectional medians for 3/31/98 and 10/30/98. The number of observations in each bucket used to determine medians highlights the importance of data density in generating the term structures predicted in the contingent-claims framework. As the number of observations decreases (especially in the lowest credit classes– 6 to 8), the picture becomes fuzzier and choppier as outliers have more impact. This snapshot highlights also the noise evident in the choppy curves and the occasional crossings (i.e. lower credit quality issues paying a lower spread then higher credit quality issues.) The cause of these differences may include measurement error, reporting error, market inefficiency, or something else we still have not considered. Regardless of the cause, these anomalies wreak havoc on many estimation approaches. These data problems are likely the cause of the lack of empirical research in this area. My approach to estimation attempts to minimize these problems while leaving most of the information in the data intact. The results, however, can not tell us much about these individual anomalies. For the time being, I assume they are– for the most part– idiosyncratic and unimportant to the general pricing trends for credit risk in the economy. In later research, I plan to return to analysis of individual bonds. Despite the noise, this cross-sectional picture still demonstrates the qualitative predictions of the Merton (1974) model– upward sloping curves for high credit quality and humped to downward sloping curves for low credit quality.

[6]These snapshots are representative of the data density and the general appearance of the raw corporate spread curves in this sample.

Looking at averages and medians, we find an intuitive picture that suggests the contingent-claims modeling approach is on the right track. However, we still need to be aware of the noisiness in this data and the difficulty this noise creates when looking at individual issues. Take for example, three companies (IBM, Texas Instruments, and Xerox) in the technology industry who all are rated A. These bonds periodically exhibit divergent behavior in their respective credit spreads. Note that the bonds sampled were standard corporate bonds without call provisions, sinking funds, or put options. Theoretically, at a particular date and duration level, the spreads should be equal. As indicated before, the likely problem is noisiness in the data and problems with assessment of credit quality. In all likelihood, these companies are not that similar in terms of credit quality (In fact, their individual EDFs differ significantly.). Combing through these data, we can uncover a number of anomalies that cannot be explained by options written into the bond indentures (the options magnify the differences). I highlight this problem to motivate my methodology of using medians in large cohorts (buckets) of bonds. Small sample size coupled with the heterogeneity of corporate bonds have hindered empirical work in the past. My approach is designed to make use of as much data as possible while minimizing the impact of the anomalous features of individual bond issues. While individual outliers exist, grouping together similar bonds and taking medians appears to filter out much of this noise.

5 The Downward Slope

Earlier in this paper, I summarized recent recent research casting doubt on the empirical evidence of downward or humped-shaped credit spread term structures. Before Helwege and Turner (1998), the evidence of a downwardly shaped credit spread term structure for low credit quality issues seemed irrefutable. The insight emphasized by Helwege and Turner (1998) focused on the mixing of different issuers when constructing the credit spread term structure. This section responds to their challenge and presents a larger sample of low credit quality issues. The characteristic differentiating this sample is the added measure of credit quality embodied in each issuer's EDF or average EDF.

First, I will present a histogram supporting the conclusions of Helwege and Turner (1998). Figure 23 reflects the distribution of differences between the worst-yield spreads for longer duration issues and the worst-yield spread for the 4-year duration issue from the same issuer. In this way, I avoid any sample selection bias from mixing different issuers in the construction of the term structure. Essentially, these slopes reflect the slope of each individual issuer's unique credit spread term structure. The choice of 4 years as the starting point allows me to focus on a region where the slope of the credit spread structure is likely negative across a large number of issuers. A number of these issuers will have humped-shaped (rising then falling) term structures and the rising portion tends to be in the range less than 4 years. For example, a particular firm may have 4 bond issues outstanding. Each issue has a different duration. I identify the issue with duration closest to 4 years without falling below 4 years. Suppose in this example the duration of the first issue is 3.8 years. Suppose also the remaining three issues have durations equal to 4.4, 5, and 6.5 years, respectively. Most issuers reflected in this sample contribute only one datapoint; however, a few firms will contribute multiple datapoints depending on the number of issues with durations that exceed 4 years. For this histogram, I would ignore the first issue and calculate the difference in spreads between the third and the second issue as well as the fourth and the second issue. The key in this histogram is that each datapoint reflects a difference for the same issuer on the same date. Both bond issues also have similar priority (I have created three classes: Senior, Subordinated, and Undetermined.). The important control in this exercise concerns testing the difference in spreads for bonds issues from the same issue. Said differently, I guarantee that I hold credit quality constant. In this way, I avoid the potential selection bias arising from reporting results where the long-dated low credit quality bond issues may, in fact, be of better quality than the short-dated low credit quality issues. As the figure shows, the median difference (or slope of the credit spread term structure between those two points on the curve) is positive. In fact, three-quarters of the sample exhibits a positive slope. These results confirm the results of Helwege and Turner (1998) with a deeper (in terms of including more low credit quality issues) and larger sample. For completeness, figure 26 shows a similar histogram for the investment grade issuers (i.e. issuers with S&P rating BBB or better.) As expected, these issuers also exhibit positively sloped terms structures.

Now let us turn to a histogram where we control for credit quality by looking only at firms with an AEDF greater than 3%. As I have demonstrated before, the credit quality within an S&P credit class changes over time. The

AEDF measure reflects the same credit quality regardless of the time period. Moreover, I can construct cohorts of bonds (especially in the sub-investment grade classes) of a more homogeneous nature. Figure 24 shows the distribution of spread differences on a sub-sample of the sub-investment grade class analyzed before. In this case, I isolate the issuers truly in a credit quality class low enough to exhibit the behavior predicted in the contingent-claims framework. Now the median slope of the term structure for each of these issuers is negative. Moving the bar a little higher and isolating issuers with an AEDF greater than 5% (i.e. 1 out of 20 of these borrowers are expected to default over the next year) generates more convincing results. As can be seen in figure 25, almost three-quarters of the sample exhibits a negative difference or downward sloping term structure. Given the presence of humped-shaped term structures that peak at durations greater than 4 years and the difficulty in identifying the exact AEDF threshhold at which firms begin reflecting the negatively sloped term structure, a few positive slopes are expected. The preponderance of negative slopes strongly confirms the theoretical predictions of the contingent-claims framework.

6 Summary

In summary, the credit spreads are noisy even after controlling possible causes of heterogeneity of prices (e.g. duration, options, and priority). Credit spreads coincide a little more closely with S&P ratings in the highest credit quality classes while middle to low credit quality classes coincide much more closely with AEDF. The qualitative shapes of the credit spread term structures conform with predictions of standard contingent-claims model of risky debt value. Even after controlling for the presences of sample selection bias, low credit quality issuers exhibit a downward sloping credit spread term structure.

Table 1: 35 Largest Issuers of Debt in Tested Bridge/EJV Sample

Company Name	**# of Issues**
Travelers Group Inc.	1347
General Electric Company	887
AMR Corporation	646
Sears Roebuck and Co.	646
Chrysler Corporation	520
Consumers Energy Company	441
International Lease Finance Corporation	422
AT&T Capital Corporation	407
Delta Airlines Inc.	406
US Airways Inc.	390
Citicorp	364
Household International Inc.	363
Beneficial Corporation	319
American General Corporation	312
Union Pacific Railroad Company	310
Caterpillar Inc.	278
Barnett Banks Inc.	273
Kroger Co.	269
CSX Corporation	252
Lehman Brothers Inc.	244
Southern New England Telephone Co.	241
Merrill Lynch & Co.	212
Bear Stearns Companies Inc.	211
Kraft General Foods Corporation	205
Norfolk and Western Railway Company	186
Paccar Financial Corp.	185
Transamerica Corporation	174
Boeing Company	163
Occidental Petroleum Corporation	159
Morgan Stanley Dean Witter & Co.	156
U S West Inc.	141
Pacific Gas and Electric Company	139
Quaker Oats Company	129
Burlington Northern Inc.	128
Ryder System Inc.	127

Table 2: Mean Standard Deviation of Worst-yield Spreads Calculated on Credit Class, Duration, and Seniority Buckets (Monthly, June, 1992 to January, 1999)

Credit Class	**Mean SD**		**Mean SD**		**Mean SD**	
(Basis Points)	**AEDF**	**S&P**	**AEDF**	**S&P**	**AEDF**	**S&P**
	All	All	Senior	Senior	Subordinated	Subordinated
1	27	25	28	11	19	10
2	38	31	49	15	23	25
3	45	27	51	26	29	21
4	50	59	58	64	29	33
5	64	84	70	84	38	70
6	75	135	81	142	47	75
7	110	329	105	333	90	N/A

Figure 1: Distribution of Firms within Financial and Non-Financial Issuer Groups across S&P Credit Class (Using December, 1998 Ratings.)

(1=AAA; 2=AA; 3=A; 4=BBB; 5=BB; 6=B; 7=CCC)

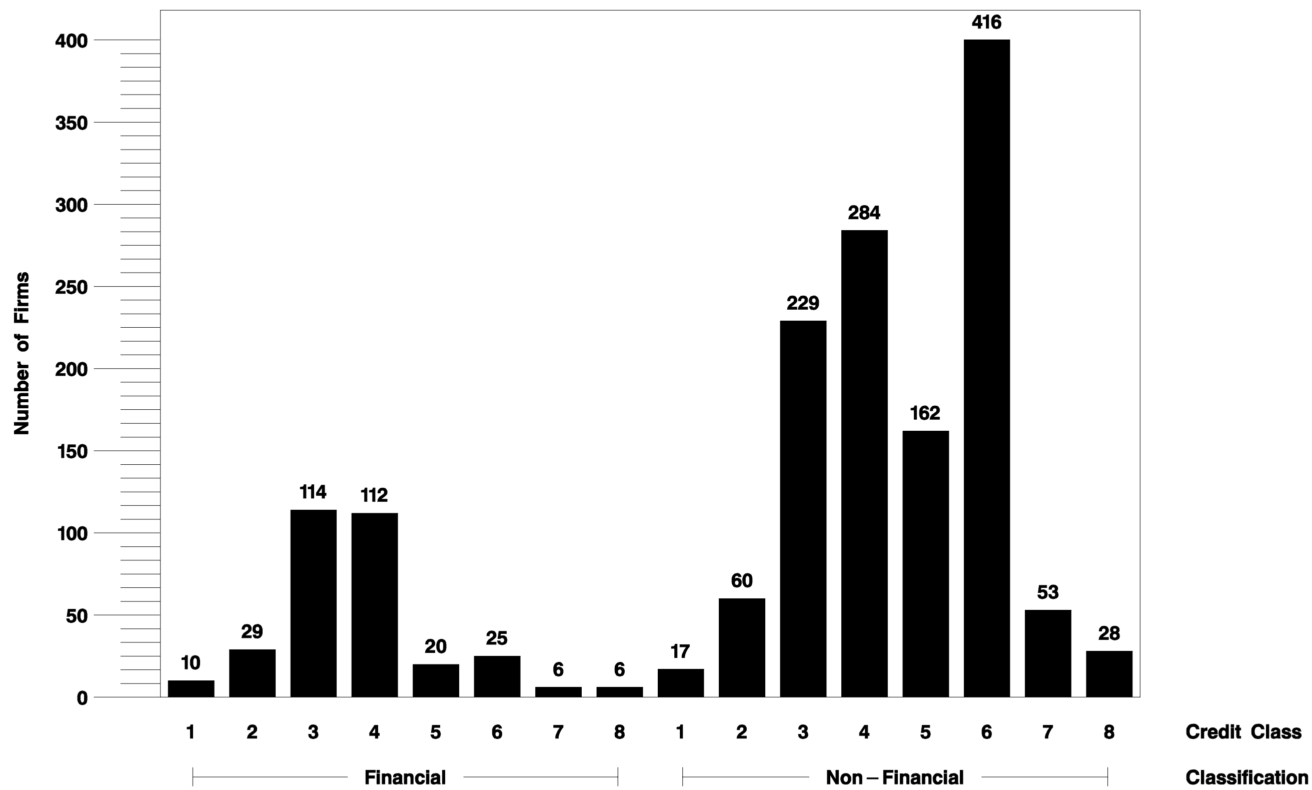


Figure 2: Distribution of Issues within Financial and Non-Financial Issuer Groups across S&P Classes (Using December, 1998 Ratings.)

(1=AAA; 2=AA; 3=A; 4=BBB; 5=BB; 6=B; 7=CCC)

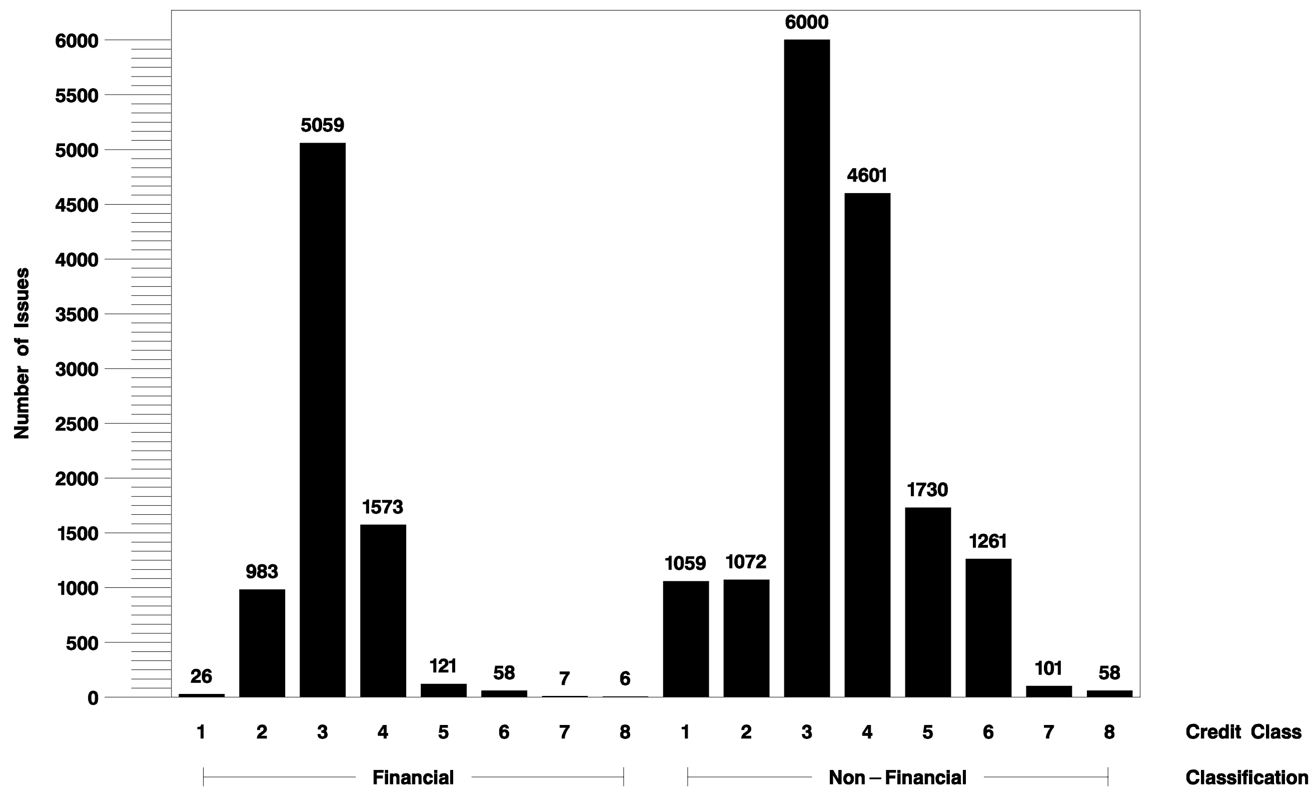


Figure 3: Distribution of Firms within Financial and Non-Financial Issuer Groups across One-year EDF Classes (Using December, 1998 One-year EDFs.)
(1=.02%; 2=.02%to.04%; 3=.04%to.08%; 4=.08%to.16%; 5=.16%to.32%; 6=.32%to.64%; 7=.64%to1.28%; 8=1.28%to20%; 9=20%)

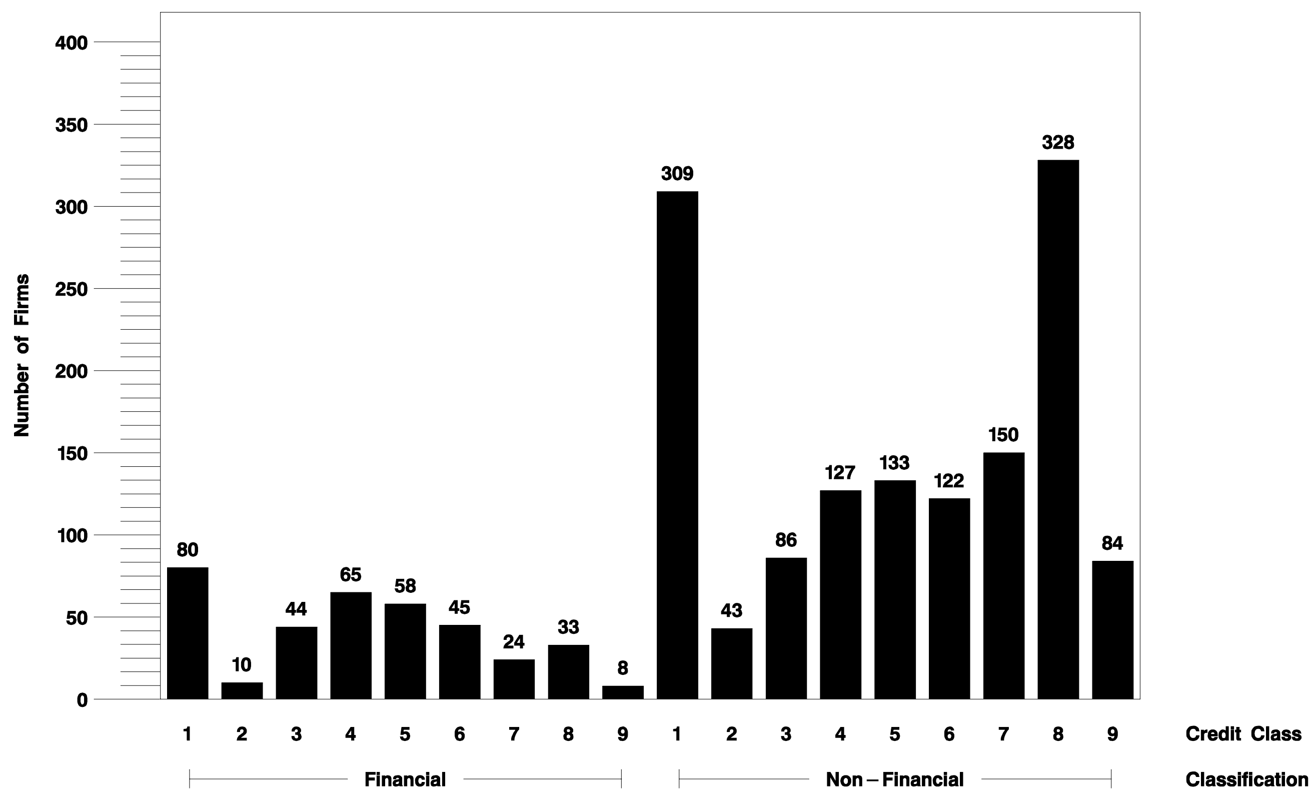


Figure 4: Distribution of Issues within Financial and Non-Financial Issuer Groups across One-year EDF Classes (Using December, 1998 One-year EDFs.
(1=.02%; 2=.02%to.04%; 3=.04%to.08%; 4=.08%to.16%; 5=.16%to.32%; 6=.32%to.64%; 7=.64%to1.28%; 8=1.28%to20%; 9=20%)

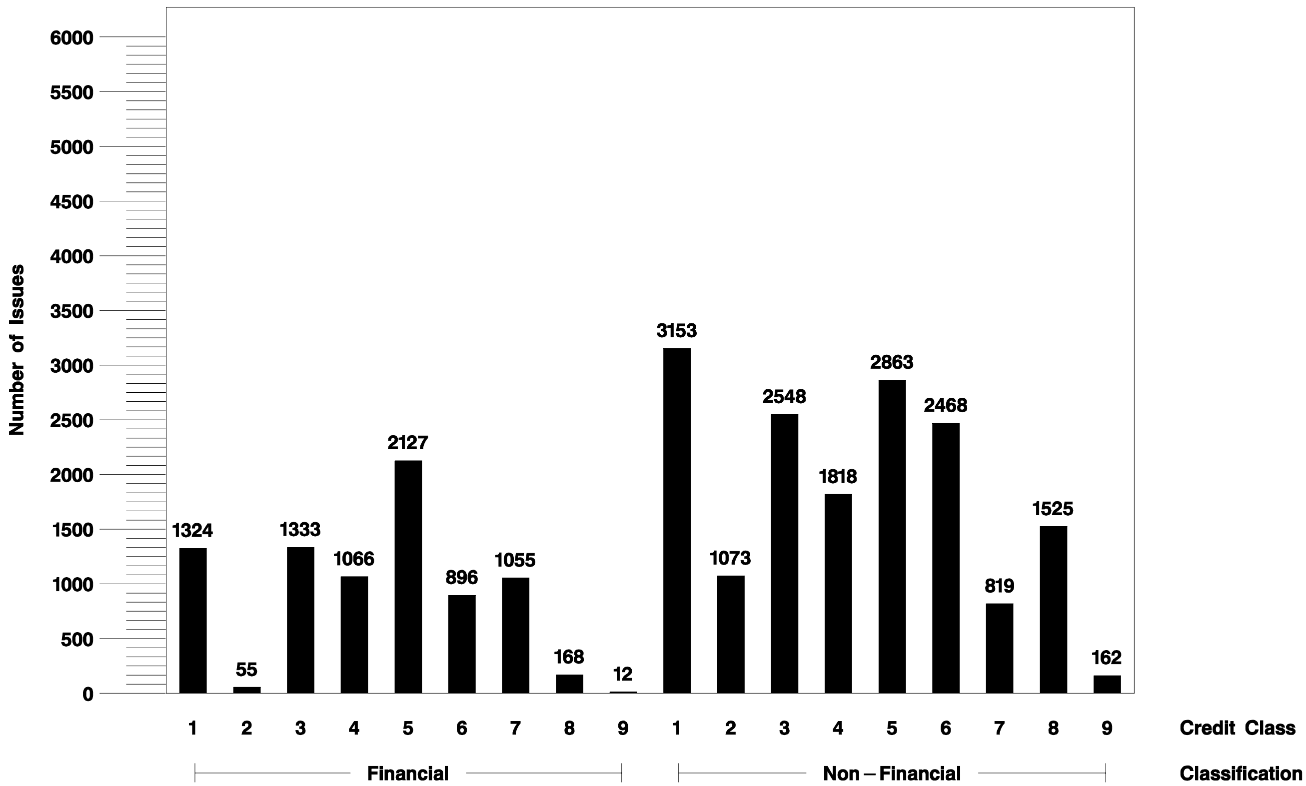


Figure 5: Distribution of Firms within Financial and Non-Financial Issuer Groups across Average EDF Classes (Using Geometric Mean of December, 1998 One through Five-Year EDFs.)
(1=< .04%; 2=.04%to.10%; 3=.10%to.18%; 4=.18%to.25%; 5=.25%to.35%; 6=.35%to.60%; 7=.60%to2.00%; 8=> 2.00%)

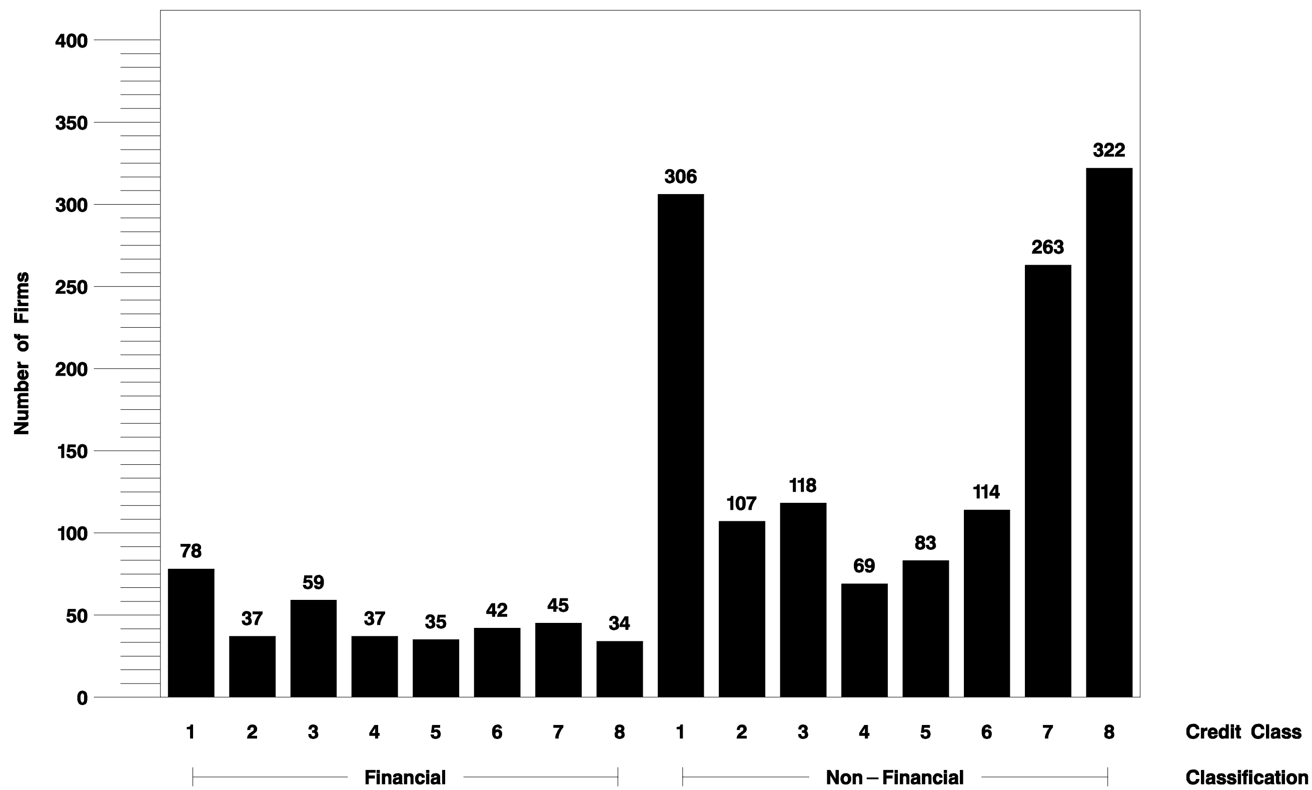


Figure 6: Distribution of Issues within Financial and Non-Financial Issuer Groups across Average EDF Classes (Using Geometric Mean of December, 1998 One through Five-Year EDFs.)

(1=< .04%; 2=.04%to.10%; 3=.10%to.18%; 4=.18%to.25%; 5=.25%to.35%; 6=.35%to.60%; 7=.60%to2.00%; 8=> 2.00%)

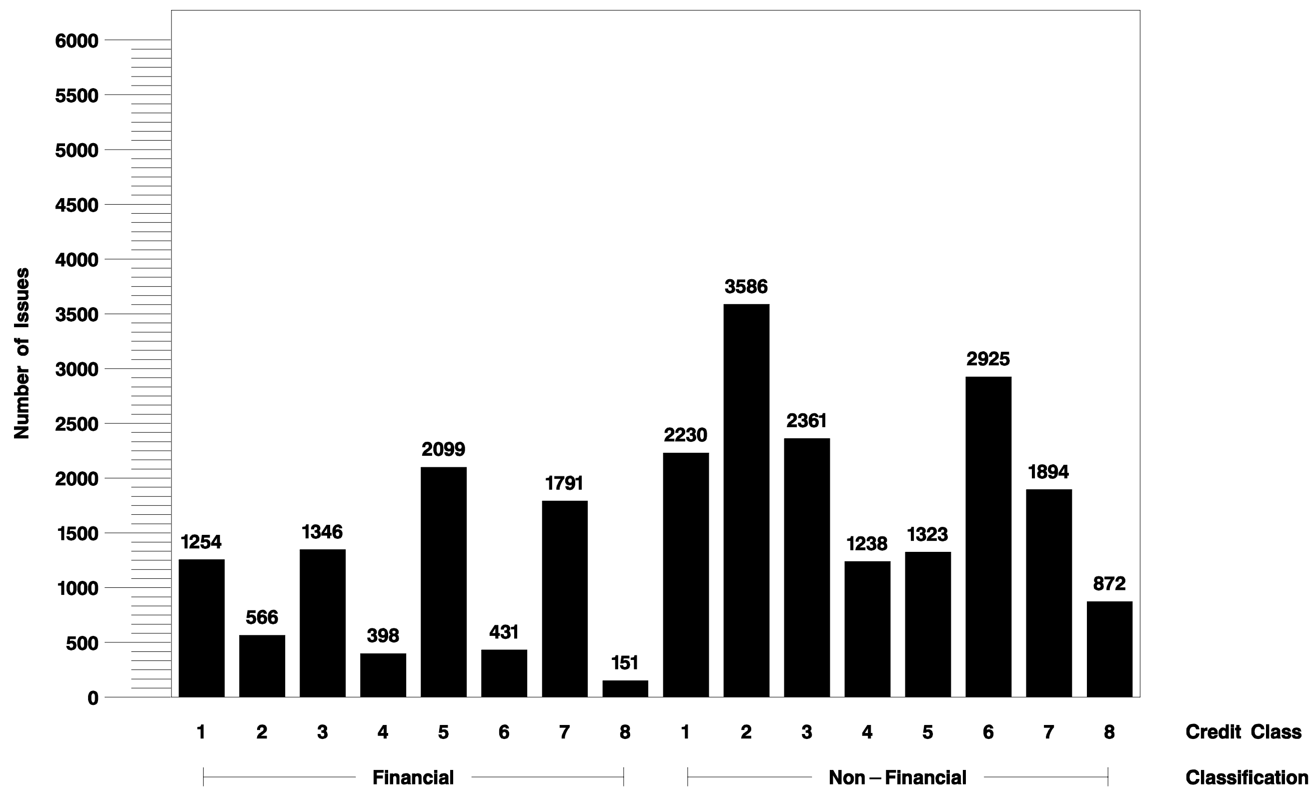


Figure 7: Distribution of Issues within Option (i.e. Issues have options– call, put, sink– attached) and Non-Option Issue Groups across Average EDF Classes (Using Geometric Mean December, 1998 One through Five-Year EDFs.)

(1=< .04%; 2=.04%to.10%; 3=.10%to.18%; 4=.18%to.25%; 5=.25%to.35%; 6=.35%to.60%; 7=.60%to2.00%; 8=> 2.00%)

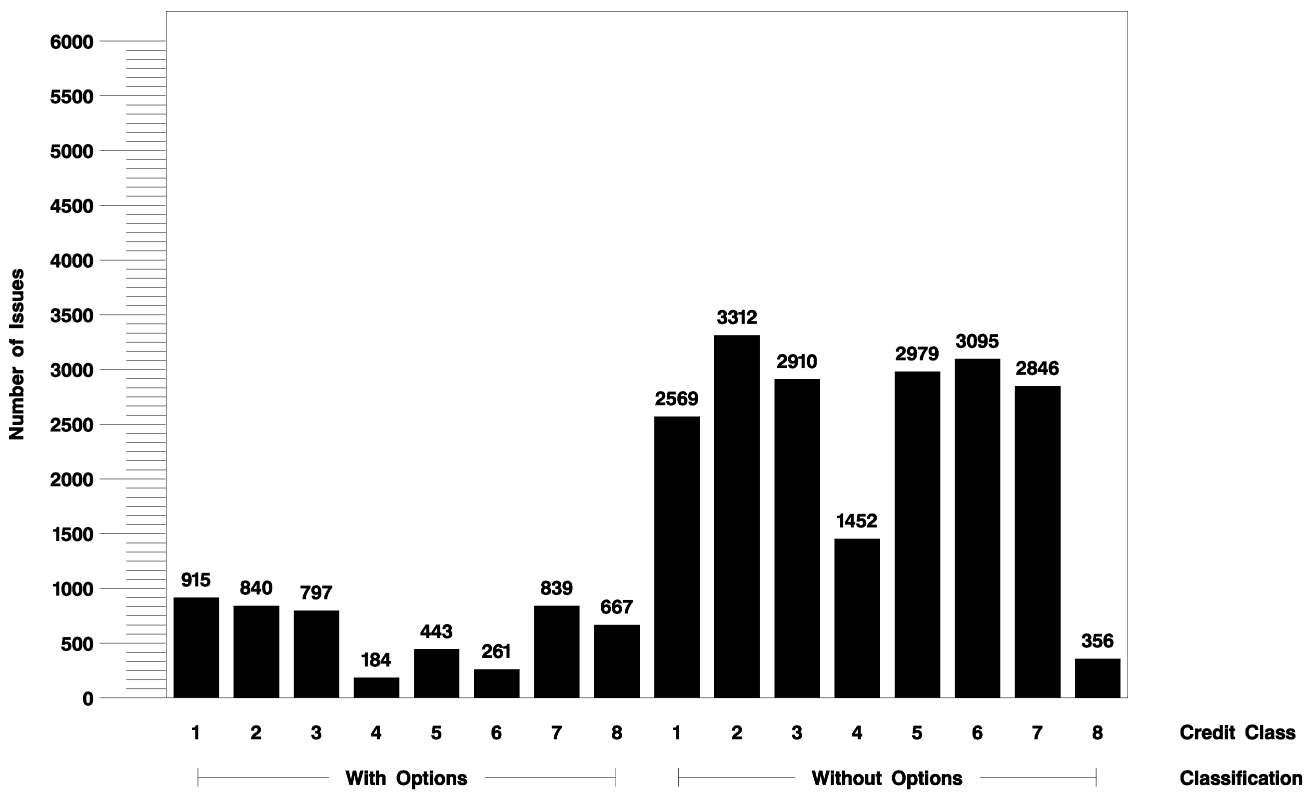


Figure 8: Distribution of Worst-Yield Spreads Monthly, June, 1992 to January, 1999 for AEDF Class 2, Duration Class 3, Senior Bonds.
(Top whisker is 95th%. Bottom whisker is 5th%. Top of box is 75th%. Bottom of box is 25th%. Bar in box is median. Dots are outliers.)

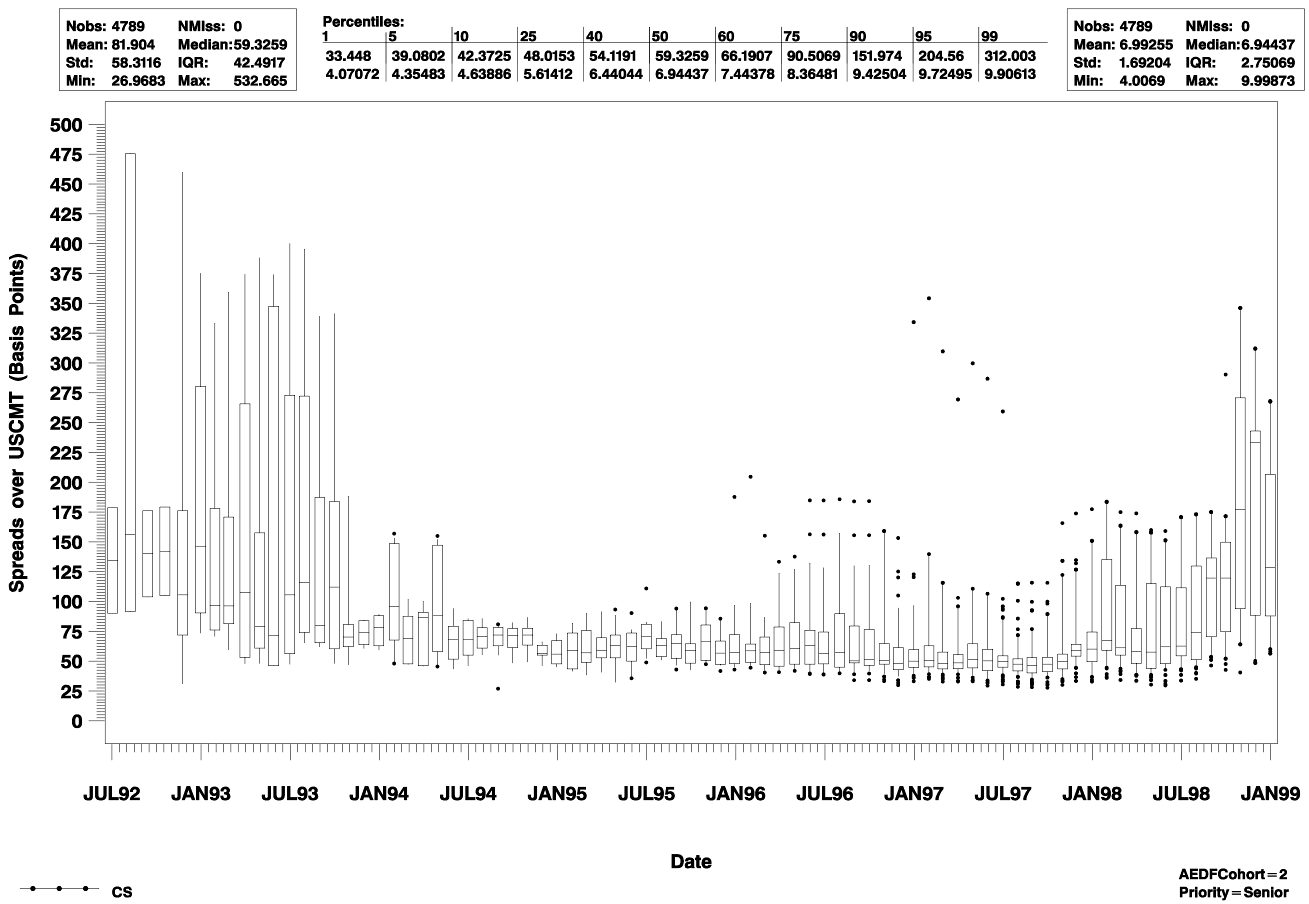


Figure 9: Distribution of Worst-Yield Spreads Monthly, June, 1992 to January, 1999 for AEDF Class 2, Duration Class 3, Subordinated Bonds.
(Top whisker is 95th%. Bottom whisker is 5th%. Top of box is 75th%. Bottom of box is 25th%. Bar in box is median. Dots are outliers.)

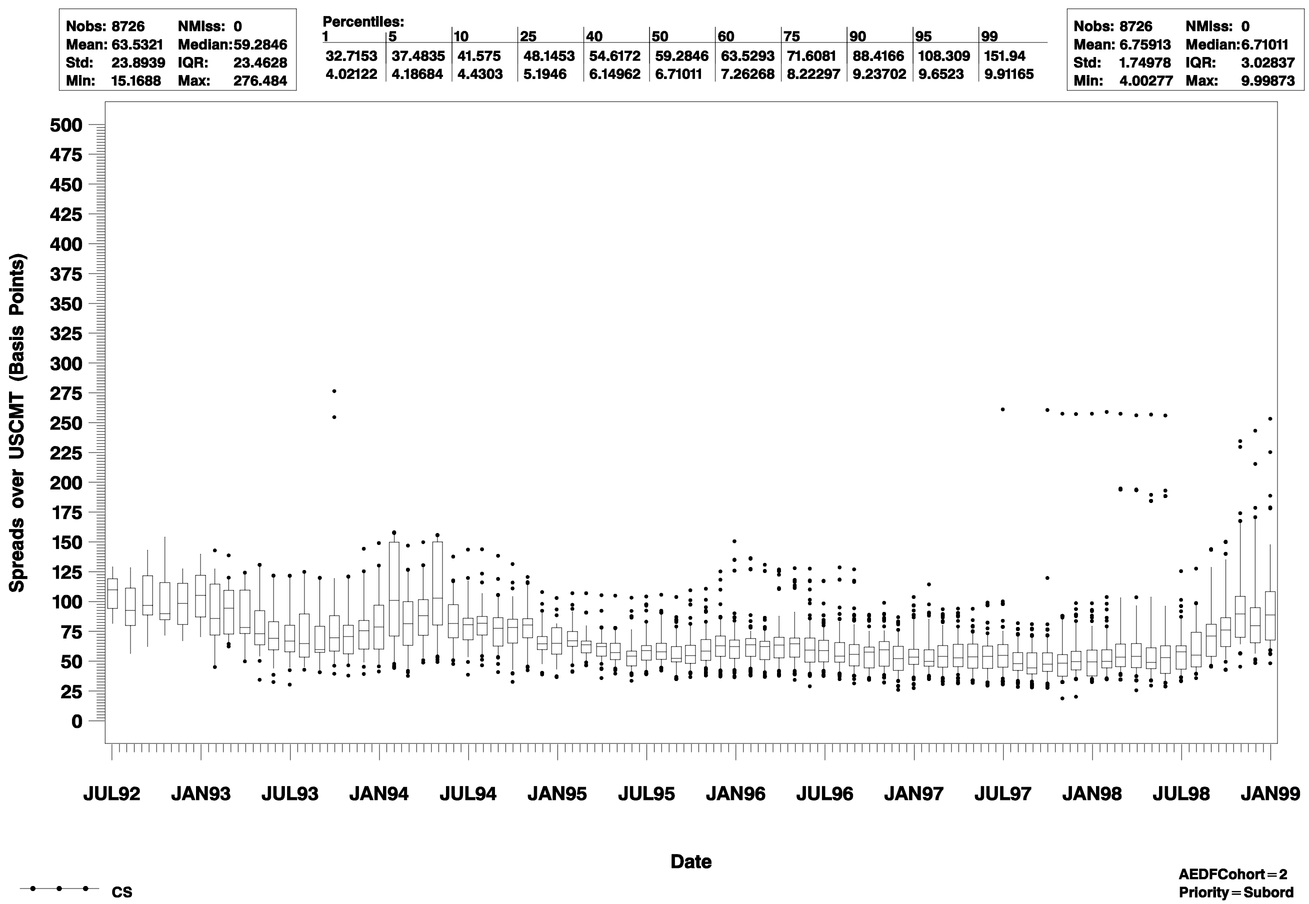


Figure 10: Distribution of Worst-Yield Spreads Monthly, June, 1992 to January, 1999 for AEDF Class 4, Duration Class 3, Senior Bonds.
(Top whisker is 95th%. Bottom whisker is 5th%. Top of box is 75th%. Bottom of box is 25th%. Bar in box is median. Dots are outliers.)

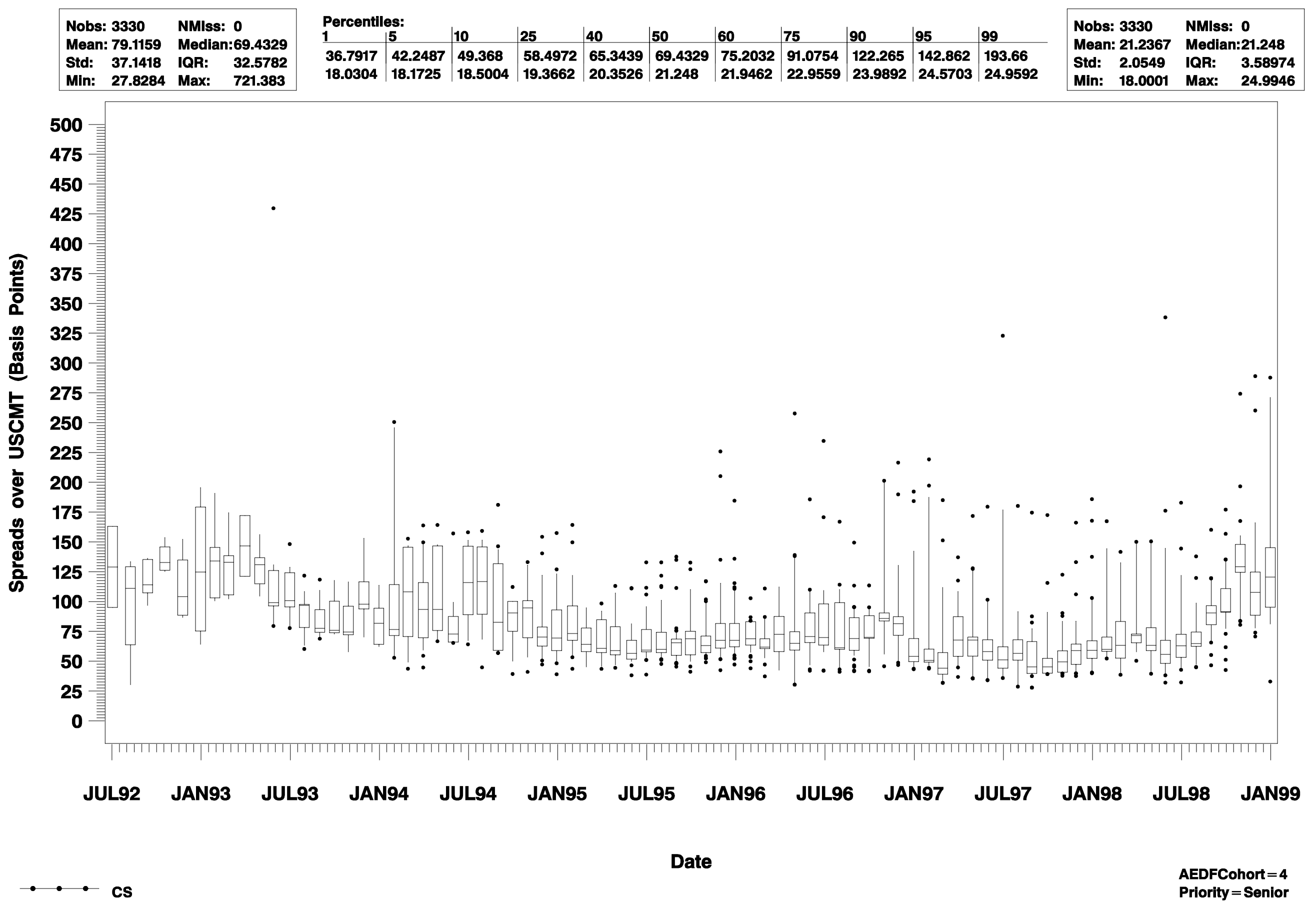


Figure 11: Distribution of Worst-Yield Spreads Monthly, June, 1992 to January, 1999 for AEDF Class 4, Duration Class 3, Subordinated Bonds.
(Top whisker is 95th%. Bottom whisker is 5th%. Top of box is 75th%. Bottom of box is 25th%. Bar in box is median. Dots are outliers.)

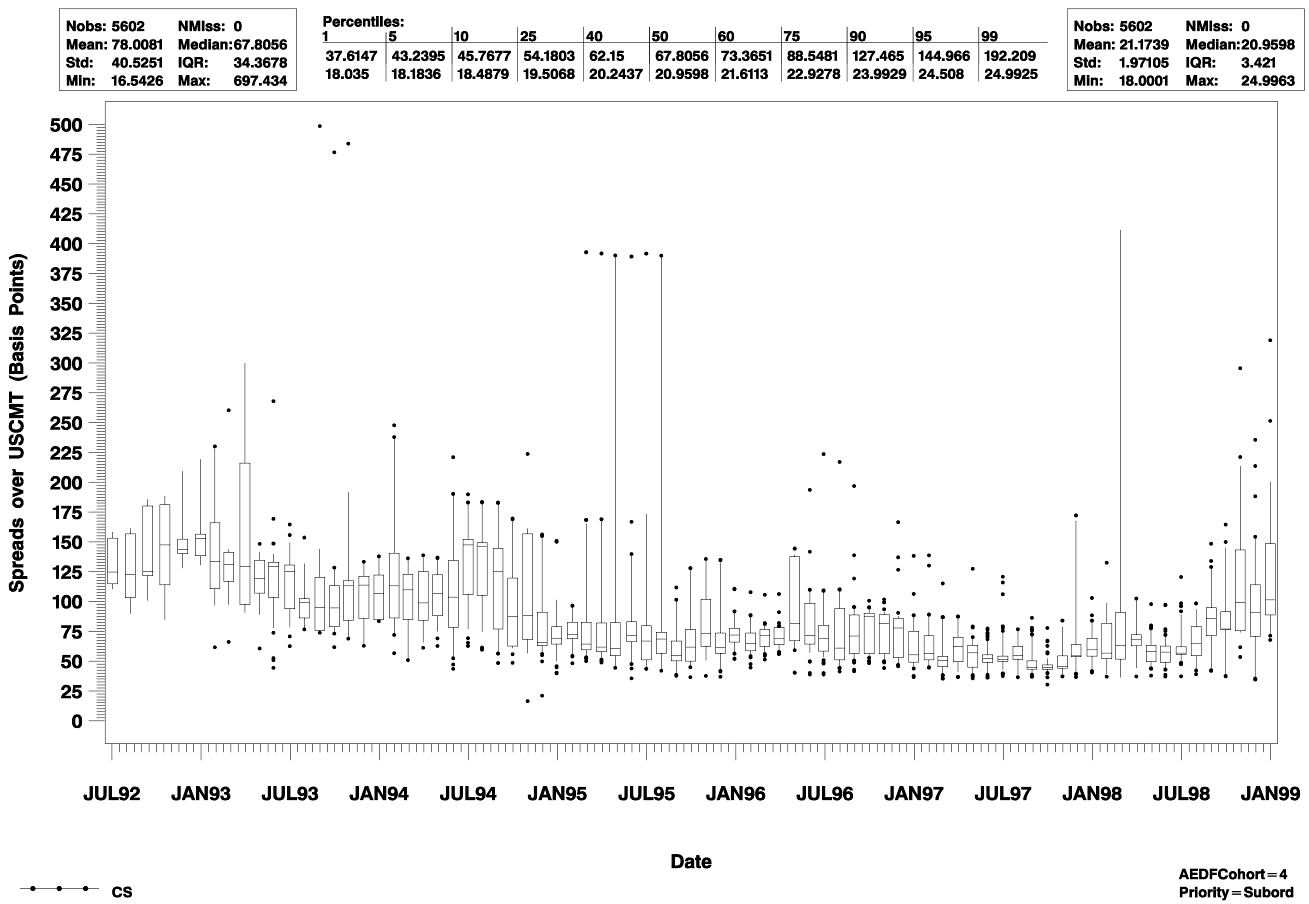


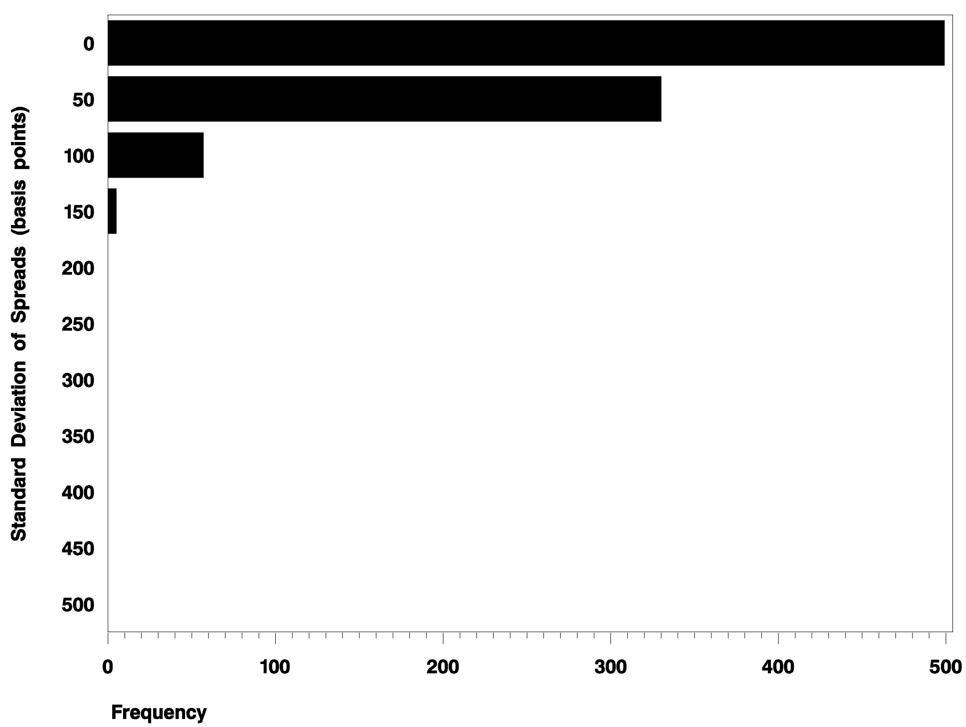
Figure 12: Distribution of Standard Deviations of Credit Spreads for Observations Monthly June, 1992 to January, 1999 within each Duration Class for S&P Class 2 (AA.)

Nobs: 891	NMiss: 0
Mean: 30.94	Median: 23.33
Std: 22.486	IQR: 17.5
Min: 6.07	Max: 157.51

Percentiles:

1	5	10	25	40	50	60	75	90	95	99
9.07	11.37	13.32	16.86	20.64	23.33	26.15	34.36	63.7	82.29	112.42

S&P Cohort = 2



FREQ.	CUM. FREQ.	PCT.	CUM. PCT.
499	499	56.00	56.00
330	829	37.04	93.04
57	886	6.40	99.44
5	891	0.56	100.00
0	891	0.00	100.00
0	891	0.00	100.00
0	891	0.00	100.00
0	891	0.00	100.00
0	891	0.00	100.00
0	891	0.00	100.00
0	891	0.00	100.00

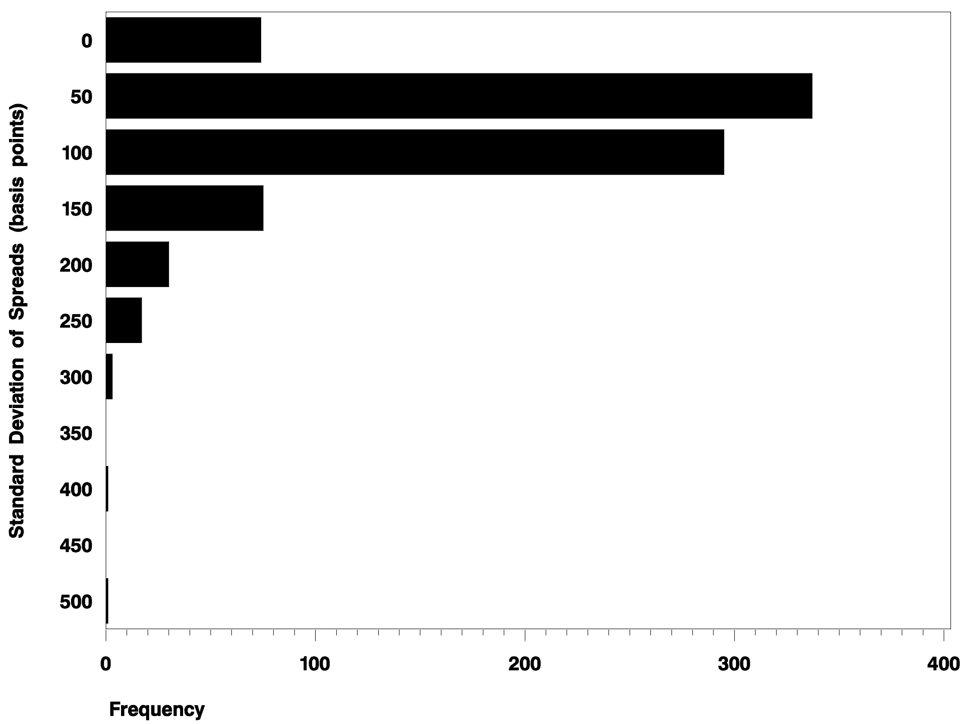
Figure 13: Distribution of Standard Deviations of Credit Spreads for Observations Monthly June, 1992 to January, 1999 within each Duration Class for S&P Class 5 (BB.)

Nobs:	833	NMiss:	14
Mean:	83.691	Median:	75.42
Std:	55.418	IQR:	61.28
Min:	0.2	Max:	584.34

Percentiles:

1	5	10	25	40	50	60	75	90	95	99
4.81	15.13	27.72	45.78	60.99	75.42	89.31	107.06	146.69	181.48	271.19

S&P Cohort=5



FREQ.	CUM. FREQ.	PCT.	CUM. PCT.
74	74	8.88	8.88
337	411	40.46	49.34
295	706	35.41	84.75
75	781	9.00	93.76
30	811	3.60	97.36
17	828	2.04	99.40
3	831	0.36	99.76
0	831	0.00	99.76
1	832	0.12	99.88
0	832	0.00	99.88
1	833	0.12	100.00

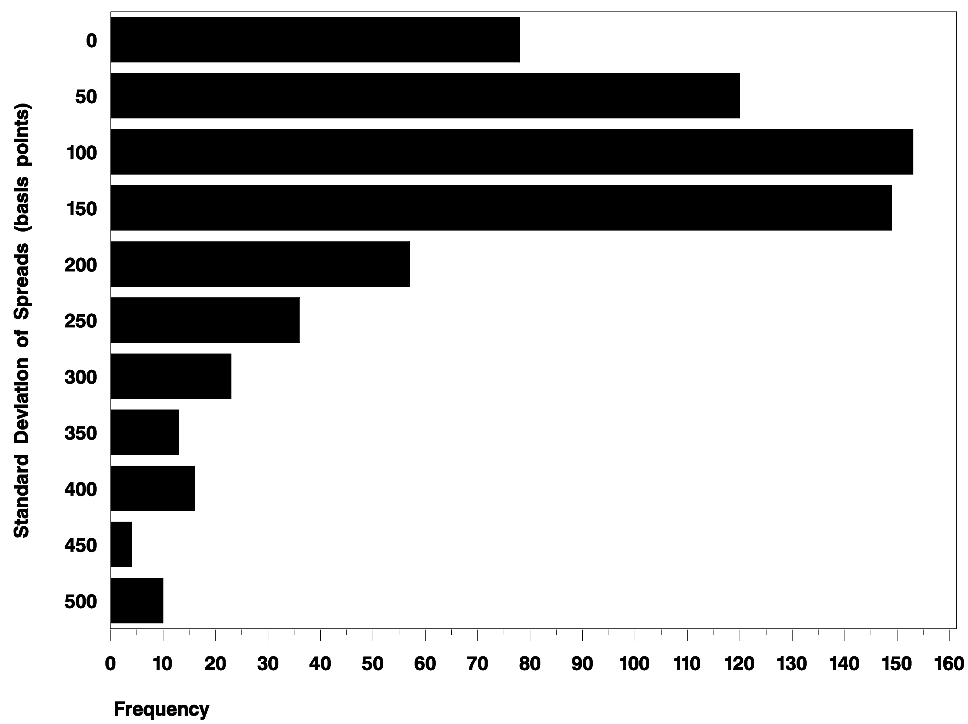
Figure 14: Distribution of Standard Deviations of Credit Spreads for Observations Monthly June, 1992 to January, 1999 within each Duration Class for S&P Class 6 (B.)

Nobs:	659	NMiss:	72
Mean:	135.08	Median:	117.73
Std:	106.54	IQR:	111.4
Min:	0	Max:	623.41

Percentiles:

1	5	10	25	40	50	60	75	90	95	99
0.02	6.03	21.9	61.49	98.73	117.73	133.18	172.89	275.09	368.54	510.96

S&P Cohort=6



FREQ.	CUM. FREQ.	PCT.	CUM. PCT.
78	78	11.84	11.84
120	198	18.21	30.05
153	351	23.22	53.26
149	500	22.61	75.87
57	557	8.65	84.52
36	593	5.46	89.98
23	616	3.49	93.47
13	629	1.97	95.45
16	645	2.43	97.88
4	649	0.61	98.48
10	659	1.52	100.00

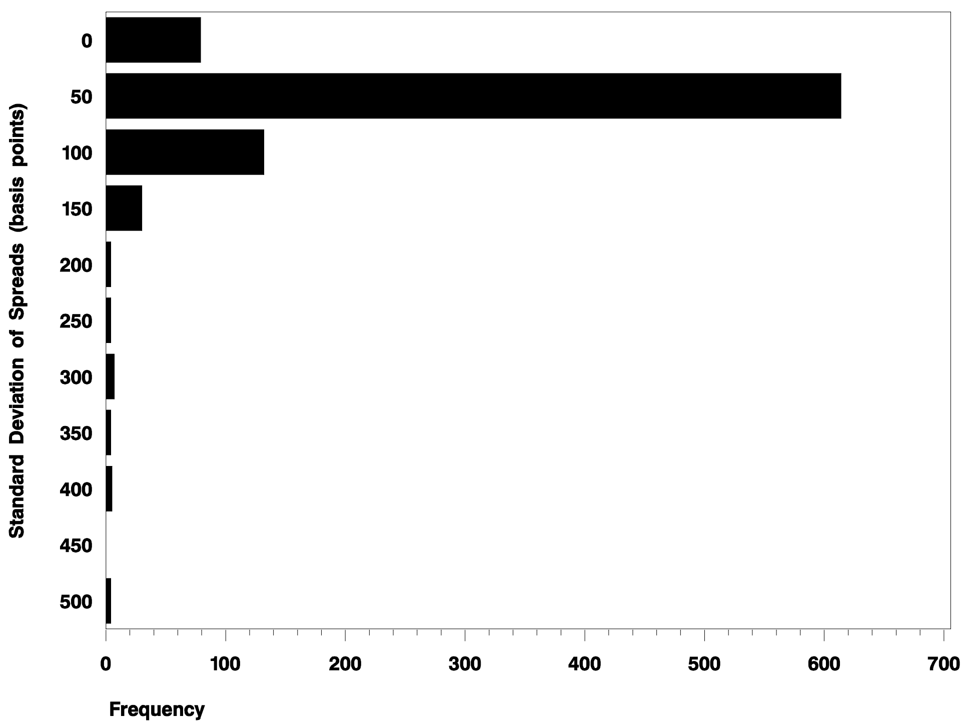
Figure 15: Distribution of Standard Deviations of Credit Spreads for Observations Monthly June, 1992 to January, 1999 within each Duration Class for AEDF Class 5 (25bps to 35bps.)

Nobs:	883	NMiss:	4
Mean:	64.031	Median:	50.75
Std:	58.631	IQR:	35.74
Min:	0.17	Max:	536.69

Percentiles:

1	5	10	25	40	50	60	75	90	95	99
10.59	19.12	26.08	35.41	44.46	50.75	56.63	71.15	102.32	142.56	376.55

AEDF Cohort=5



	FREQ.	CUM. FREQ.	PCT.	CUM. PCT.
0	79	79	8.95	8.95
50	614	693	69.54	78.48
100	132	825	14.95	93.43
150	30	855	3.40	96.83
200	4	859	0.45	97.28
250	4	863	0.45	97.73
300	7	870	0.79	98.53
350	4	874	0.45	98.98
400	5	879	0.57	99.55
450	0	879	0.00	99.55
500	4	883	0.45	100.00

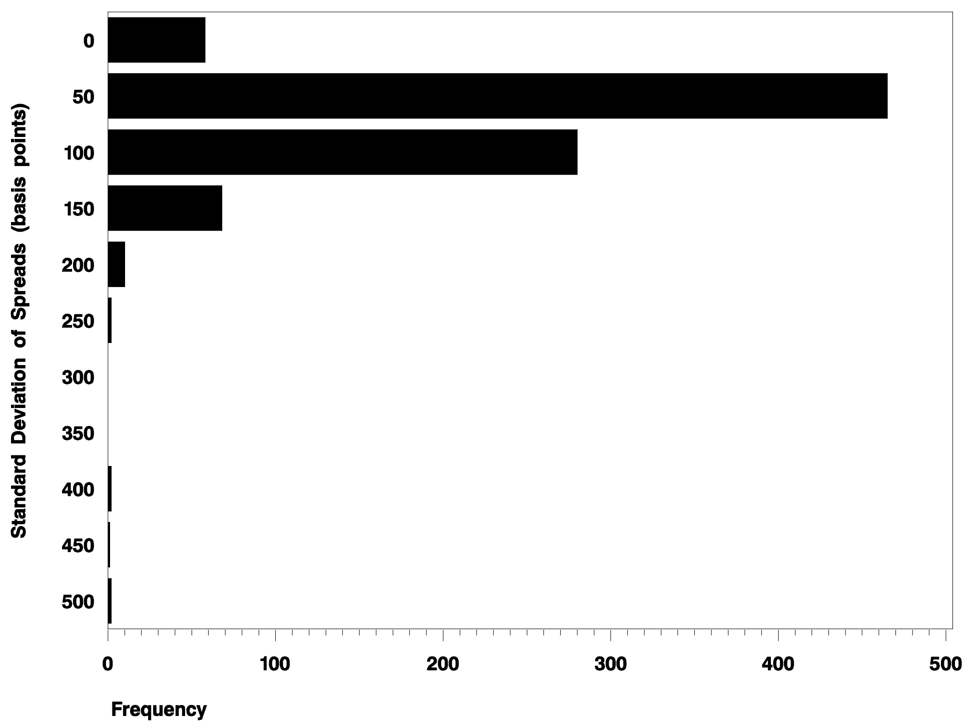
Figure 16: Distribution of Standard Deviations of Credit Spreads for Observations Monthly June, 1992 to January, 1999 within each Duration Class for AEDF Class 6 (35bps to 60bps.)

Nobs:	888	NMiss:	2
Mean:	74.609	Median:	68.11
Std:	46.887	IQR:	43.11
Min:	0.38	Max:	525.86

Percentiles:

1	5	10	25	40	50	60	75	90	95	99
12.42	22.49	30.06	48.19	60.82	68.11	76.01	91.3	124.02	144.05	208.6

AEDF Cohort=6



FREQ.	CUM. FREQ.	PCT.	CUM. PCT.
58	58	6.53	6.53
465	523	52.36	58.90
280	803	31.53	90.43
68	871	7.66	98.09
10	881	1.13	99.21
2	883	0.23	99.44
0	883	0.00	99.44
0	883	0.00	99.44
2	885	0.23	99.66
1	886	0.11	99.77
2	888	0.23	100.00

Figure 17: Term Structure of Worst-Yield Spreads (median in cross-section, average over time, monthly June, 1992 to January, 1999) over USCMT for S&P Classes.

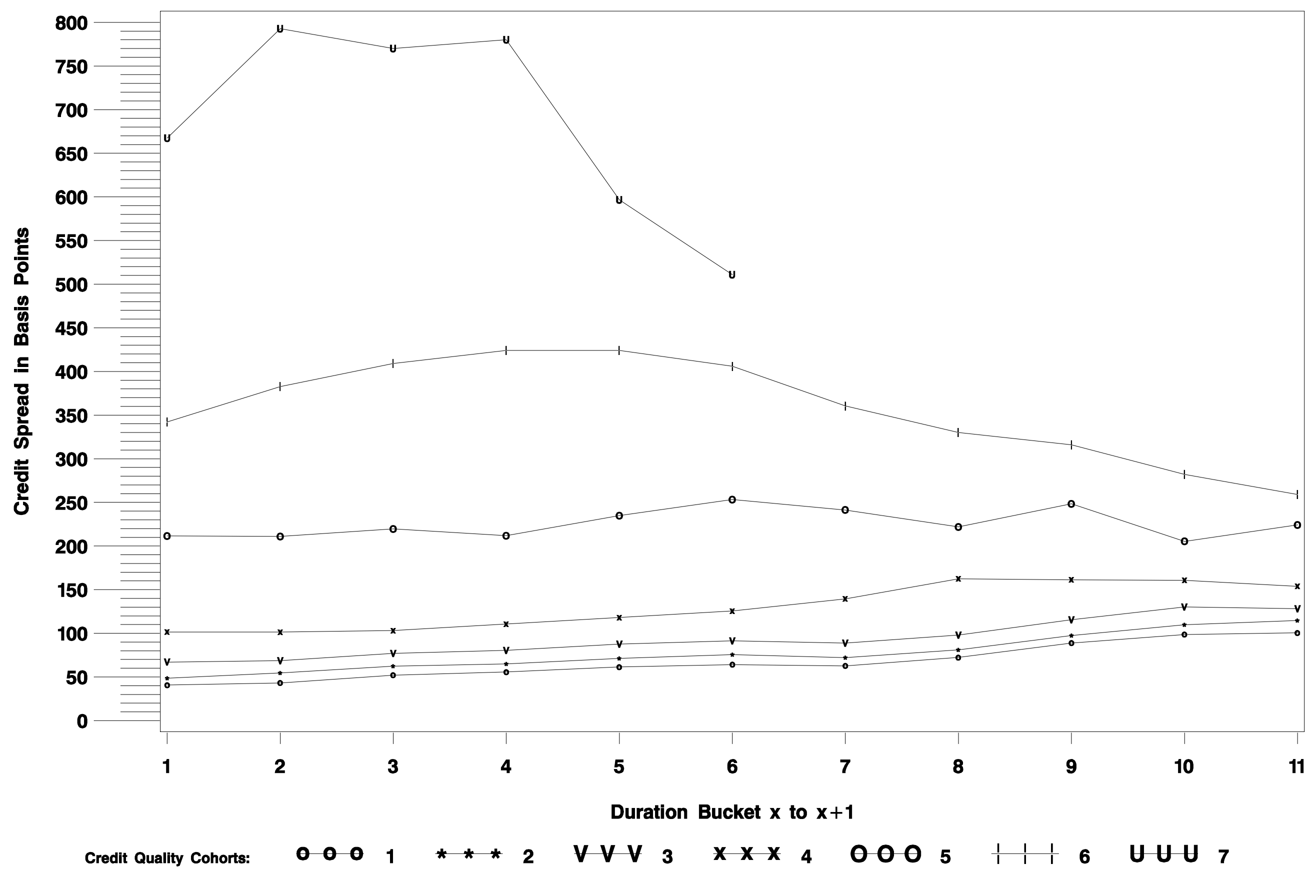


Figure 18: Term Structure of Worst-Yield Spreads (median in cross-section, average over time, monthly June, 1992 to January, 1999) over USCMT for AEDF Classes.

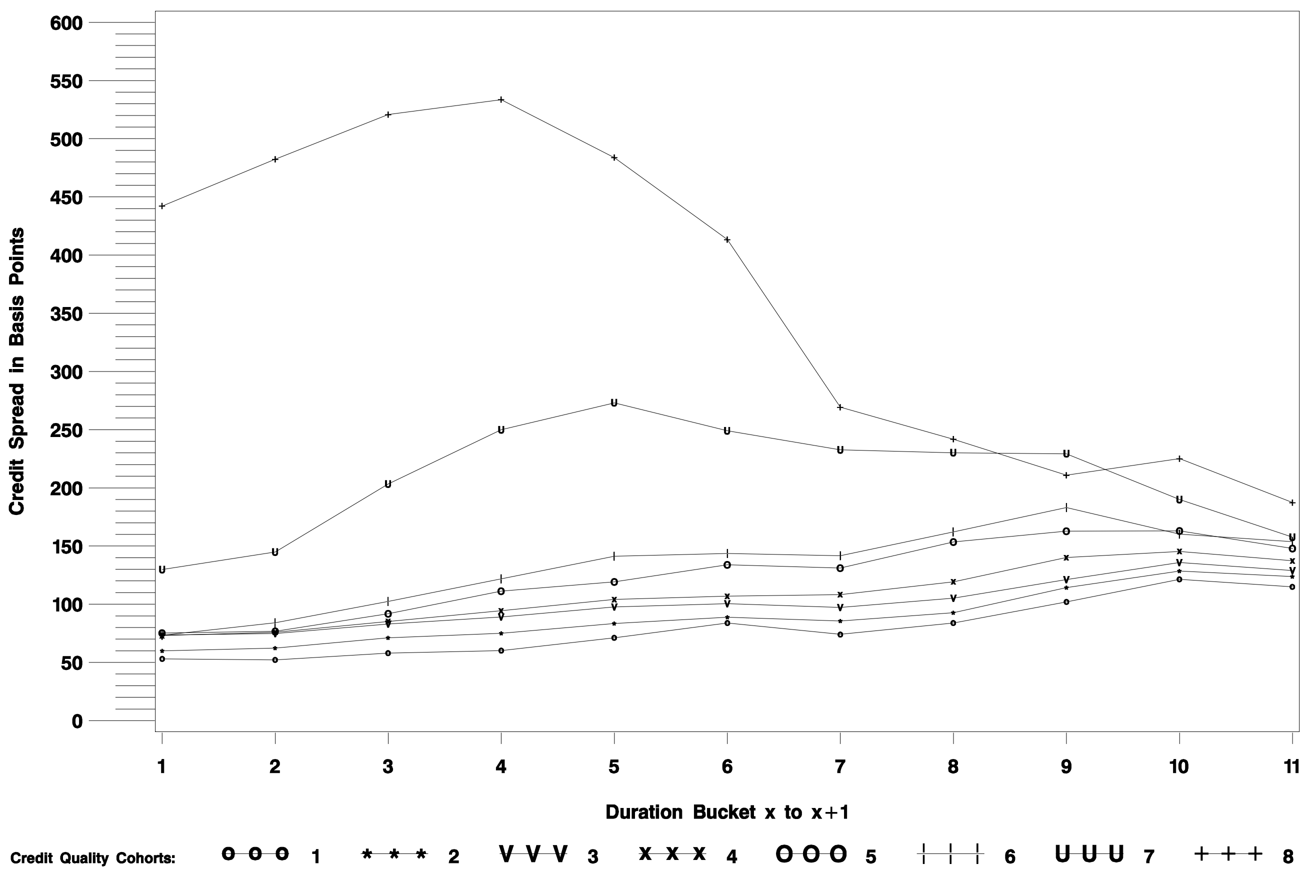


Figure 19: March, 1998 Term Structure of Median Worst-Yield Spreads over USCMT for AEDF Classes 1 to 4.

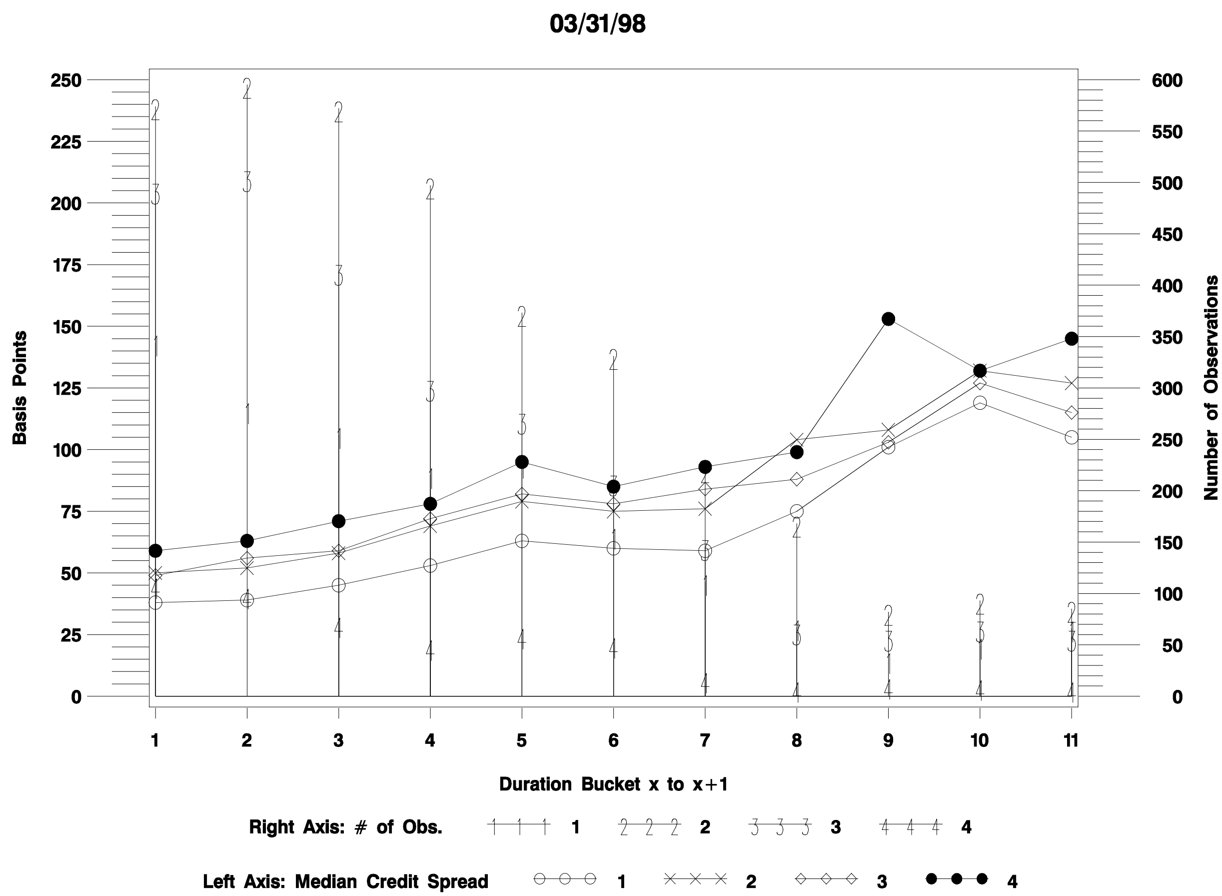


Figure 20: March, 1998 Term Structure of Median Worst-Yield Spreads over USCMT for AEDF Classes 5 to 8.

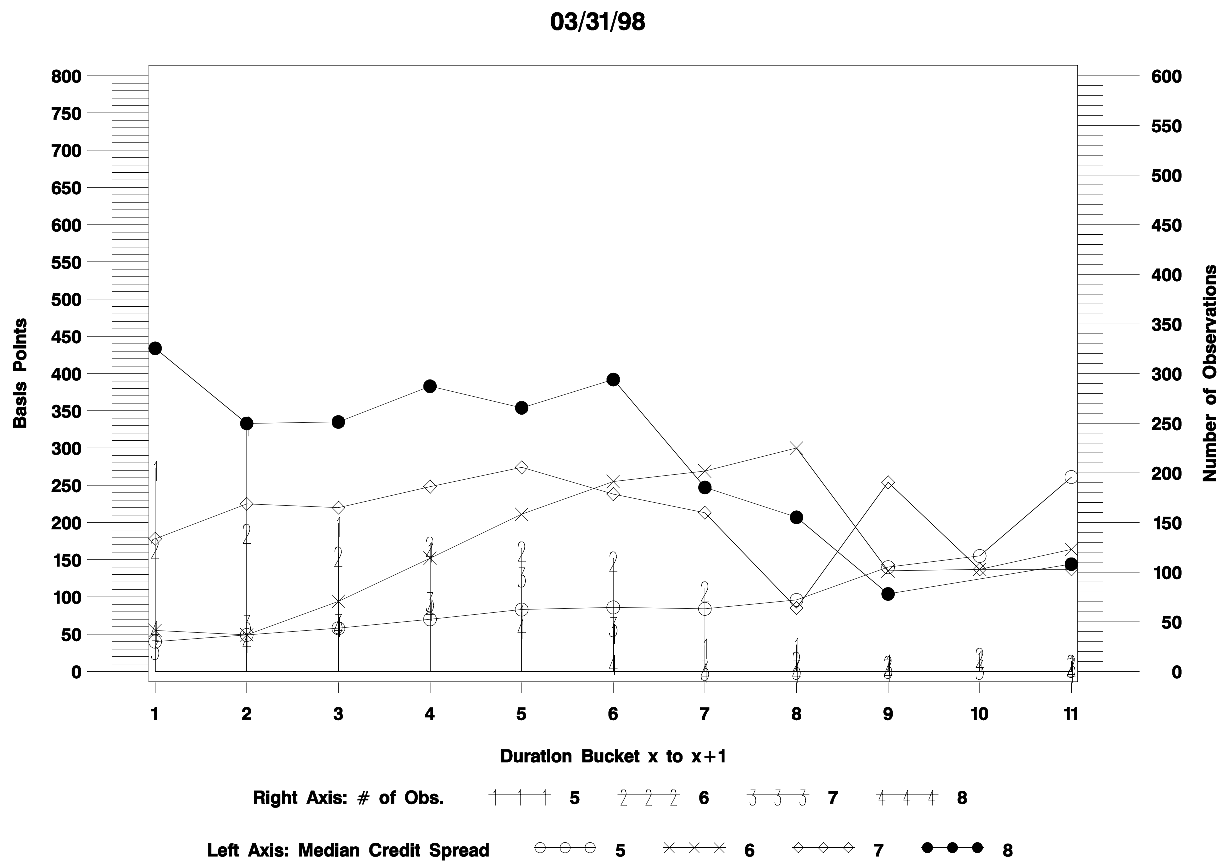


Figure 21: October, 1998 Term Structure of Median Worst-Yield Spreads over USCMT for AEDF Classes 1 to 4.

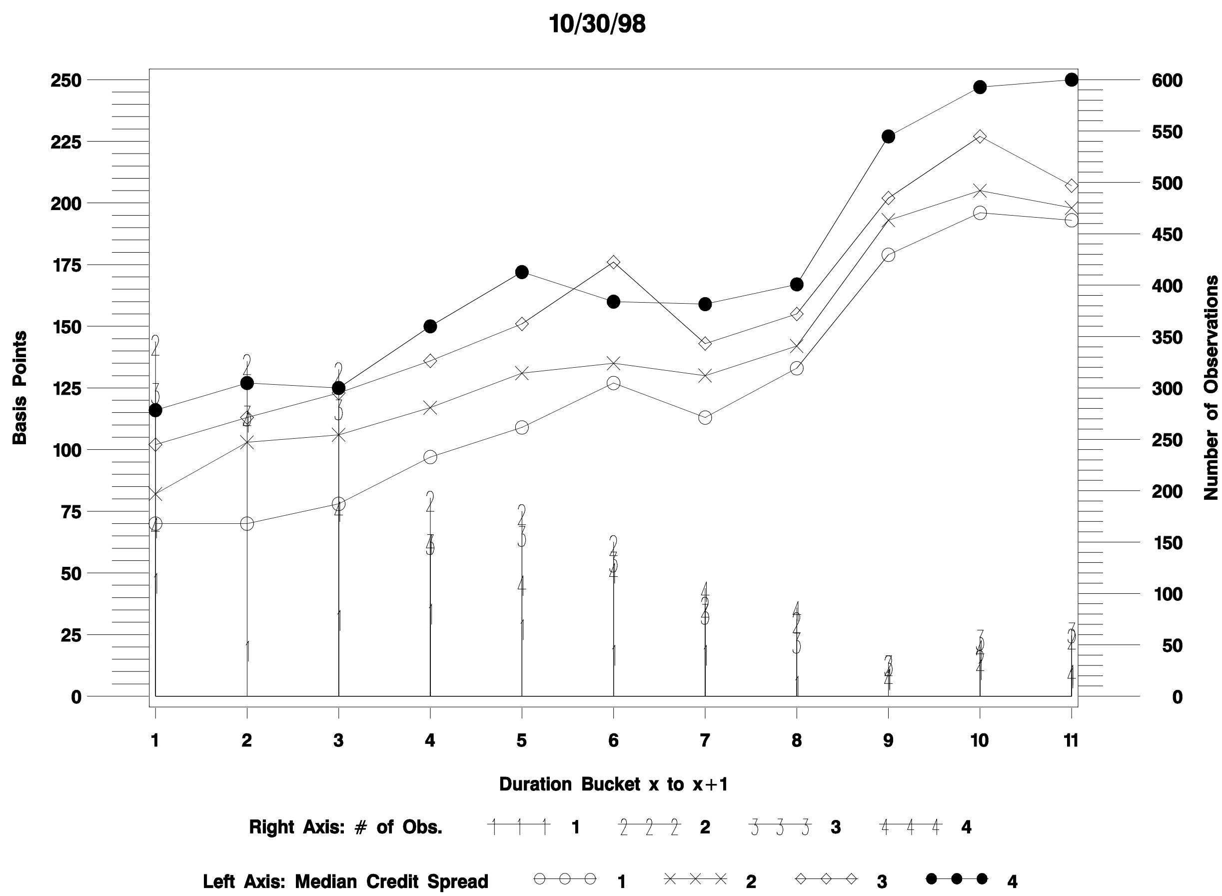


Figure 22: October, 1998 Term Structure of Median Worst-Yield Spreads over USCMT for AEDF Classes 5 to 8.

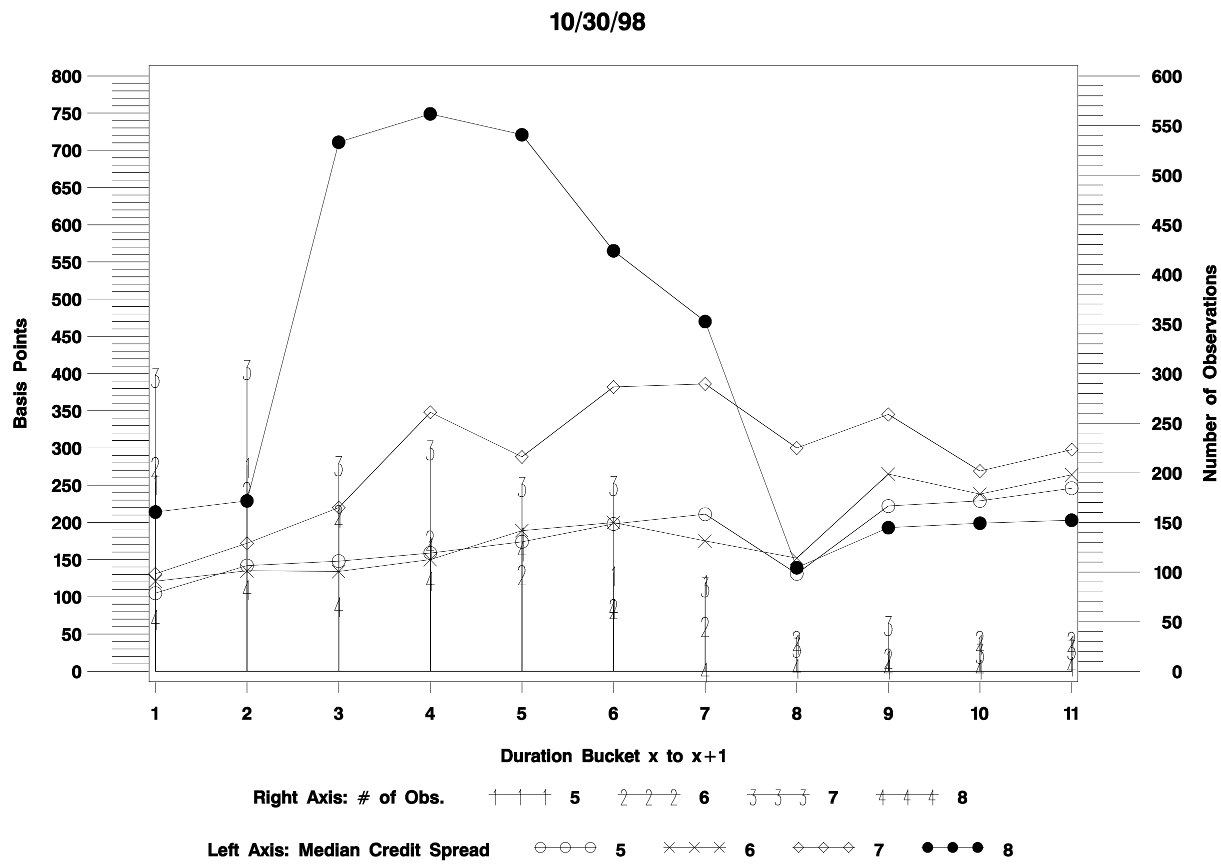


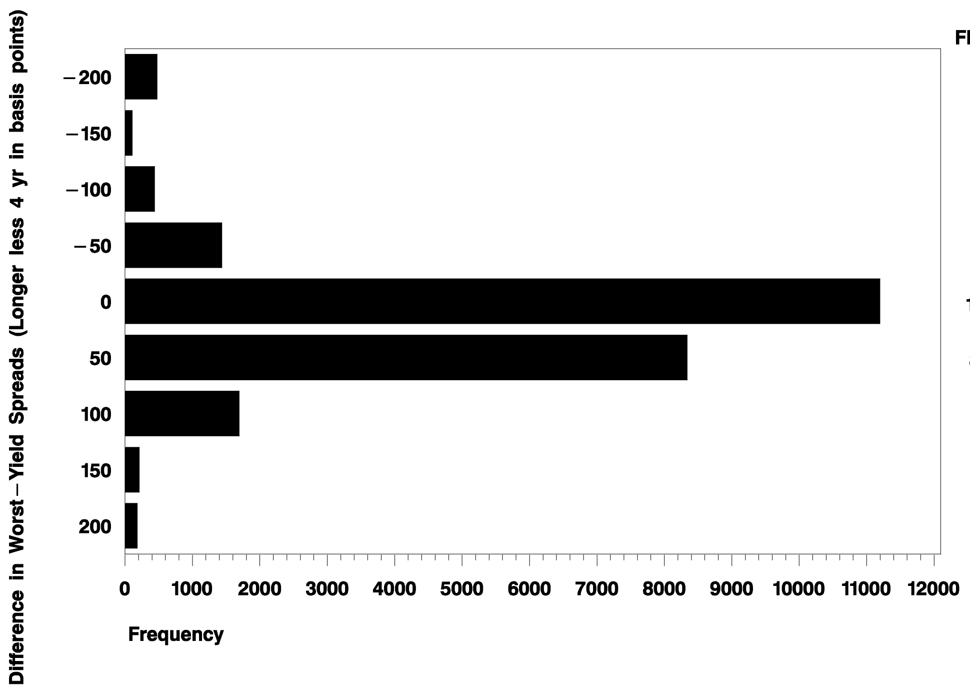
Figure 23: Distribution of Differences between Worst-Yield Spreads of Longer Duration Issues and 4-year Duration Issue of Same Sub-investment Grade Issuer.

(Sub-investment grade reflects S&P rating of BB, B, or CCC. Each datapoint represents a difference calculated on the same date.)

Nobs:	24076	NMiss:	0
Mean:	13.363	Median:	16
Std:	86.303	IQR:	50
Min:	−1328	Max:	900

Percentiles:

1	5	10	25	40	50	60	75	90	95	99
−468	−65	−27	−1	7	16	27	49	71	89	157



FREQ.	CUM. FREQ.	PCT.	CUM. PCT.
476	476	1.98	1.98
107	583	0.44	2.42
438	1021	1.82	4.24
1435	2456	5.96	10.20
11196	13652	46.50	56.70
8338	21990	34.63	91.34
1693	23683	7.03	98.37
211	23894	0.88	99.24
182	24076	0.76	100.00

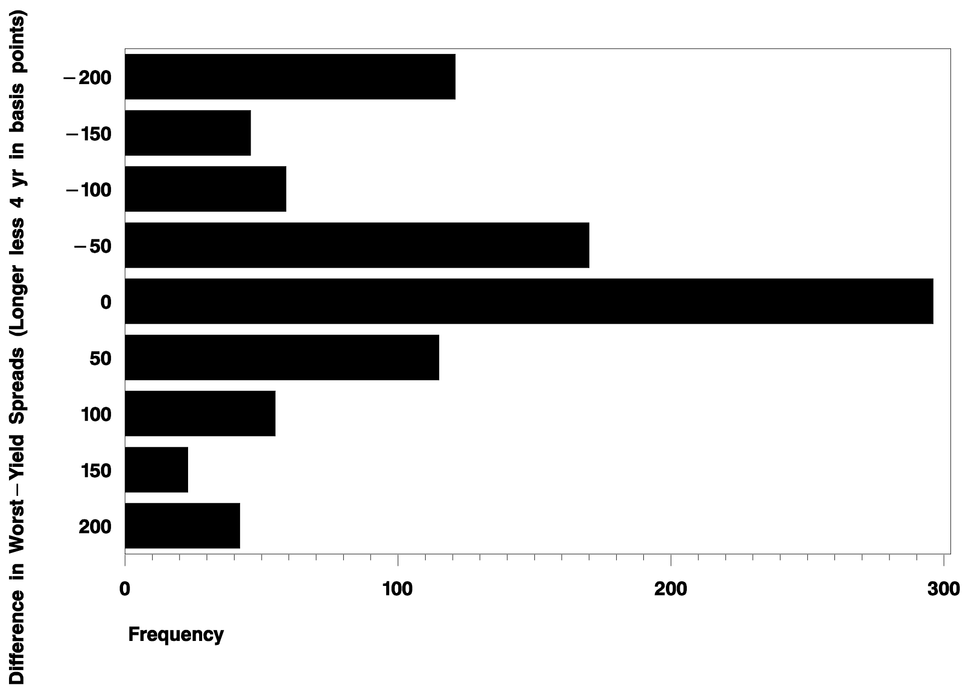
Figure 24: Distribution of Differences between Worst-Yield Spreads of Longer Duration Issues and 4-year Duration Issue of Same Sub-investment Grade Issuer where AEDF$>3\%$.

(Sub-investment grade reflects S&P rating of BB, B, or CCC. Each datapoint represents a difference calculated on the same date.)

Nobs:	927	NMiss:	0
Mean:	−62.88	Median:	−15
Std:	239.14	IQR:	101
Min:	−1328	Max:	900

Percentiles:

1	5	10	25	40	50	60	75	90	95	99
−1118	−507	−258	−72	−31	−15	−5	29	94	164	430



	FREQ.	CUM. FREQ.	PCT.	CUM. PCT.
−200	121	121	13.05	13.05
−150	46	167	4.96	18.02
−100	59	226	6.36	24.38
−50	170	396	18.34	42.72
0	296	692	31.93	74.65
50	115	807	12.41	87.06
100	55	862	5.93	92.99
150	23	885	2.48	95.47
200	42	927	4.53	100.00

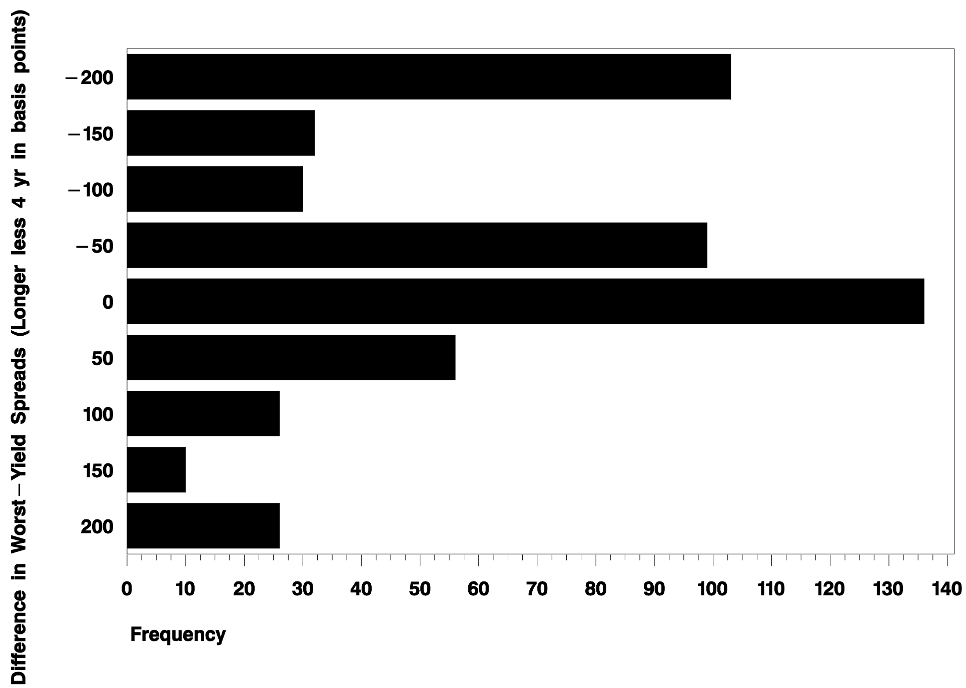
Figure 25: Distribution of Differences between Worst-Yield Spreads of Longer Duration Issues and 4-year Duration Issue of Same Sub-investment Grade Issuer where AEDF$>5\%$.

(Sub-investment grade reflects S&P rating of BB, B, or CCC. Each datapoint represents a difference calculated on the same date.)

Nobs:	518	NMiss:	0
Mean:	−106.8	Median:	−27
Std:	299.36	IQR:	140
Min:	−1328	Max:	900

Percentiles:

1	5	10	25	40	50	60	75	90	95	99
−1211	−829	−450	−129	−51	−27	−15	11	81	175	619



FREQ.	CUM. FREQ.	PCT.	CUM. PCT.
103	103	19.88	19.88
32	135	6.18	26.06
30	165	5.79	31.85
99	264	19.11	50.97
136	400	26.25	77.22
56	456	10.81	88.03
26	482	5.02	93.05
10	492	1.93	94.98
26	518	5.02	100.00

References

Altman, E. I., and V. M. Kishore (1996): "Almost Everything You Wanted to Know about Recoveries on Defaulted Bonds," *Financial Analysts Journal*, November/December, 57–64.

Black, F., and M. Scholes (1973): "The Pricing of Options and Corporate Liabilities," *Journal of Political Economy*, 81, 637–659.

Bohn, J. R. (1999): "Survey of Contingent-Claims Approaches to Risky Debt Valuation," *University of California, Berkeley*, Working Paper.

Carty, L. V., S. C. Keenan, and I. Shtogrin (1998): "Historical Default Rates of Corporate Bond Issuers, 1920-1997," *Moody's Investors Service, Global Credit Research*, February.

Helwege, J., and C. M. Turner (1998): "The Slope of the Credit Yield Curve for Speculative-Grade Issuers," *Working Paper*.

Jarrow, R. A., D. Lando, and S. Turnbull (1997): "A Markov Model for the Term Structure of Credit Risk Spreads," *Review of Financial Studies*, 10(2), 481–523.

Jones, P. E., S. P. Mason, and E. Rosenfeld (1984): "Contingent Claims Analysis of Corporate Capital Structures: An Empirical Investigation," *Journal of Finance*, 39, 611–625.

Lee, C. J. (1981): "The Pricing of Corporate Debt: A Note," *Journal of Finance*, 36(5), 1187–1189.

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

Pitts, C. G. C., and M. J. P. Selby (1983): "The Pricing of Corporate Debt: A Further Note," *Journal of Finance*, 38, 1311–1313.

Sarig, O., and A. Warga (1989): "Some Empirical Estimates of the Risk Structure of Interest Rates," *Journal of Finance*, 44, 1351–1360.

Zuckerman, G. (September 10, 1998): "Mysterious Bond Pricing Will Get Some Scrutiny," *Wall Street Journal, Interactive Edition.*

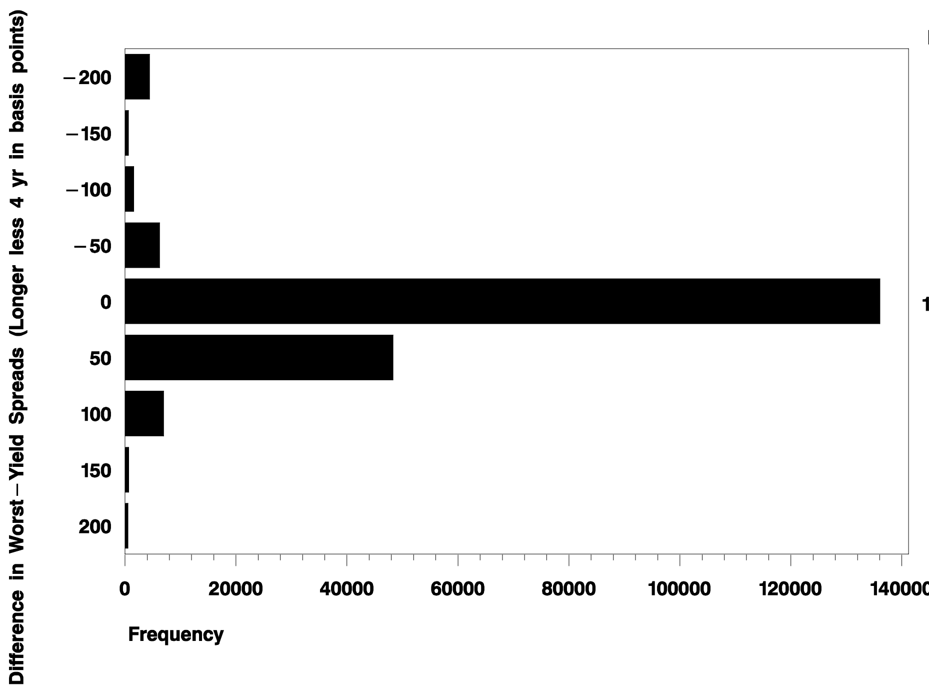
Figure 26: Distribution of Differences between Worst-Yield Spreads of Longer Duration Issues and 4-year Duration Issue of Same Investment Grade Issuer.

(Investment grade reflects S&P rating of AAA, AA, A, or BBB. Each datapoint represents a difference calculated on the same date.)

Nobs:	205332	NMiss:	0
Mean:	−5.697	Median:	10
Std:	183.56	IQR:	27
Min:	−1880	Max:	1859

Percentiles:

1	5	10	25	40	50	60	75	90	95	99
−1099	−37	−10	0	6	10	15	27	52	69	109



	FREQ.	CUM. FREQ.	PCT.	CUM. PCT.
−200	4432	4432	2.16	2.16
−150	609	5041	0.30	2.46
−100	1568	6609	0.76	3.22
−50	6231	12840	3.03	6.25
0	136032	148872	66.25	72.50
50	48279	197151	23.51	96.02
100	6959	204110	3.39	99.40
150	671	204781	0.33	99.73
200	551	205332	0.27	100.00

A EDF Calculation

The expected default frequency or EDF™ is a forward-looking measure of the actual probability of default. Using traded equity prices, KMV corporation uses option pricing technology to calculate the EDF. In simple terms the steps for calculation are as follows:

1. The price of a firm's exchanged-traded equity is combined with an estimate of the equity's volatility to determine the firm's asset value and volatility. Similar to Merton (1974), equity is viewed as a call option on the underlying assets of the firm with an exercise price equal to the face value of the debt. With a suitable model to price this option, the firm's asset value and asset volatility can be calculated by inverting the function relating asset value and equity value. KMV uses a proprietary option pricing model to determine asset value and asset volatility.

2. Next, the empirical default distribution is estimated using a large historical default database. The key relationship concerns the probability of default given that the firm's asset value is a certain number of standard deviations away from its default point. This point can be calculated in a number of different ways; however, extensive testing at KMV determined that current liabilities plus one-half of long-term liabilities does as well as more complicated characterizations of the default point.

3. Based on the empirical distribution of default and the current number of standard deviations (recall the standard deviation or asset volatility is calculated simultaneously with the firm's asset value) the firm's asset value is away from its default point, an expected default frequency can be calculated. In other words, EDF measures the probability of reaching the default point given the firm's current asset value, its asset value, and its capital structure.

EDFs are part of KMV's software package called *Credit Monitor*. They are updated monthly and available on a subscription basis. Currently, EDFs are calculated for one to five years.