Modelling Extreme Credit Events

Presented by:

Parit Jakhria

To:

The Staple Inn Actuarial Society

06 September 2011

MODELLING EXTREME CREDIT EVENTS

A report from the Extreme Events Working Party

By R. Frankland, A. Holtham, P. Jakhria, J. Kingdom, O. Lockwood , A. Smith, E. Varnell

1. INTRODUCTION

1.1 As part of their capital modelling, insurance companies need to project extreme market events, such as falls in equity markets or changes in credit markets. These market stresses are particularly important in the context of FSA's ICAS regime and in the forthcoming Solvency II regime. Under both regimes, insurers are required to hold enough capital to be able to withstand a '1-in-200' year event over a one-year period. That is, over a one-year horizon, the probability that an insurer's 'own funds' become negative is at most 0.5%[1].

1.2 This report focuses on credit risk and aims to illustrate possible methods to model a distribution of corporate bond spreads over a one-year time horizon with the aim of coming to a view on extreme outcomes over a one year period.

2. PROPERTIES / ISSUES OF CORPORATE BONDS

2.1 Taking a step back, it is interesting to draw parallels with equity risk, where there is a large degree of consistency between equity benchmarks amongst insurers (e.g. most UK insurers are very likely to have either the FTSE All Share or MSCI UK as the benchmark for UK Equity portfolio). However, this does not appear to be the case for fixed interest, where there is a wide range of credit portfolios benchmarks held by insurers.

2.2 This is reflected in the risk modelling and capital calculations within the insurance industry. For example the CEIOPS consultation papers on the draft advice for the level 2 implementing measures on Solvency II published a single 'standard formula' stress for global equities, whereas the credit stresses were split over credit ratings, durations and different structures within credit[2].

2.3 This is partly the result of multi-dimensionality apparent within credit risk, as the portfolios can be differentiated by duration, credit rating and type of structure in addition to the sector and geographical splits that exist within equity space.

---

[1] This is a very broad simplification, as there are many technical differences in the interpretation. For example the interpretation of how the 'liquidity premium' or 'matching premium' on corporate bonds spreads is treated is differently under each regime.

[2] It should be noted that most major insurance firms will be aiming to use an internal model under Solvency II. However, we would expect at least the same level of granularity under an internal model.

2.4 As such, there is a lot of debate on the required level of granularity of corporate bond modelling, particularly for the purposes of defining stresses. On the one hand, we would like to be able to precisely capture the different dimensions of credit risk. However on the other hand, as we shall see in the next section, we are somewhat constrained on the quality of conclusions by the availability of long term historical data and need to avoid spurious accuracy.

2.5 We also note that, in general, corporate bond returns can be decomposed into three elements; interest rate movements, spread movements and defaults / migrations[3].

2.6 It is often convenient to separate the interest rate risk element from credit risk as it affects both sides of the balance sheet – movements in risk-free interest rates cause changes in asset values and changes to the discount rates applied to liabilities. This is largely the case in UK. Also, the interest rate derivatives market is extremely large and liquid, and hedging via interest rate derivatives is relatively cheap and well understood within the insurance industry. As a result, the interest rate risk in corporate bond portfolios is often hedged to a large extent, especially within annuity business.

2.7 The richness of long term data available for migrations and defaults can vary considerably. While aggregated annual transition matrices are available in the public domain, more detailed data can be obtained, at a fee, from other sources like the three major rating agencies (Moody's, S&P and Fitch). Depending on the product purchased, this data can be quite extensive – for example, the Moody's Default and Recovery Database can offer data on the default and migration experience of individual issuers from the 1920s onwards (although questions exist on the quality and relevance of data of such age).

2.8 Various methods exist to analyse the risk from credit migrations and defaults, but we do not explore these in this paper. One potential way of looking at this risk would be through scenario testing, based on data from previous stressed credit events (we know for example that the ABI has commissioned such studies for its members in the past).

2.9 Consequently, we have focused the bulk of our modelling and analysis towards coming up with extreme (1 in 200) credit spread stresses over a period of 1 year.

2.10 An alternative approach which we partly investigate in this paper would be to work with total return indices which should capture all elements of credit risk, depending on the details of how the index is constructed. The only drawback of this approach is that the conclusions are constrained by the lack of a suitably long period of historical data.

---

[3] It should be noted that this list is not exhaustive, as we can also have other credit events, such as credit restructurings, exercising of call options, etc. (although these may still be classified as defaults by rating agencies using different recovery values compared to more conventional defaults)

3. CHOICE OF DATASET

3.1 There were three main datasets that we considered, summarised below:

Moody's / Federal Reserve

This comprises of spread (over treasuries) and yield data of long-dated investment grade bonds. The data is based on Seasoned Bonds with Remaining Maturities of at Least 20 Years. This is derived from pricing data on a regularly-replenished population of nearly 90 seasoned corporate bonds in the US market, each with current outstanding over $100 million.

The bonds have maturities as close as possible to 30 years and they are dropped from the list if their remaining life falls below 20 years or if their credit ratings change. Bonds with deep discounts or steep premiums to par are generally excluded. All yields are yield-to-maturity calculated on a semi-annual compounding basis. Each observation is an unweighted average, with Average Corporate Yields representing the unweighted average of the corresponding Average Industrial and Average Public Utility observations.

The indices are available monthly from 1919, and daily from 1997.

Merrill Lynch - Broad Market Index (UK00)

The Broad Market Index tracks the performance of investment grade public debt of Sovereign, Quasi-Government and Corporate issuers. It includes collateralized, securitized and unsecured investment grade bonds having at least one year remaining term to maturity, a fixed coupon schedule and a stated minimum amount outstanding (for example GBP 500 million for Gilts and GBP 100 million for all other securities.)

Bonds must be rated investment grade based on a composite of Moody's and S&P. In addition to their own rating requirements, qualifying issuers (other than Supranationals) must be domiciled in a country having an investment grade foreign currency long-term debt rating (composite of Moody's and S&P).

The index is re-balanced on the last calendar day of the month. Issues that meet the qualifying criteria are included in the index for the following month. Issues that no longer meet the criteria during the course of the month remain in the index until the next month-end re-balancing at which point they are dropped from the index.

A wide range of indices are available, including USD, EUR and GBP denominations. Additional sub-indices are available that segment the Index by maturity, sector and rating. The inception date of the Index is December 31, 1996 for GBP, 31 December 1995 for EUR and 31 December 1988 for USD.

iBoxx

Markit iBoxx indices cover the cash bond market. They comprise liquid investment grade issues. Underlying bond prices and indices are available in real time for EUR and GBP and end of day for USD and Asia. The indices are sorted by type of issuer, maturity band, credit rating and sector. Most Markit iBoxx bond indices are rebalanced monthly.

Markit iBoxx indices are rules-based to ensure they are objective and replicable (please see appendix 1). The selection criteria used to determine which bonds are included, mean that the indices represent the part of the market that is tradable and thus available to investors and asset managers.

The iBoxx GBP benchmark indices comprise an overall and two major index sub-groups for the gilts and non-gilts sectors. The non-gilts group is detailed further into sub-groups for Sovereigns & Sub-Sovereigns, Collateralized and Corporates.

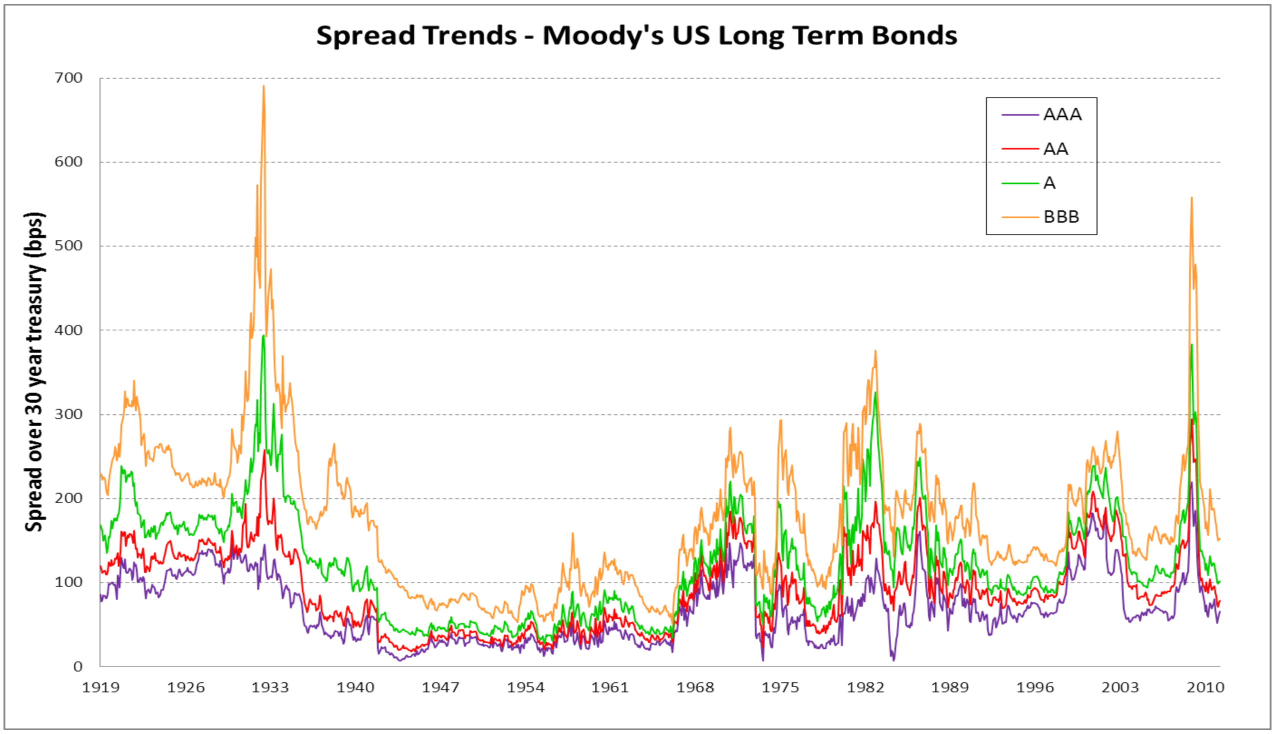
The Corporates index includes rating and sector sub-indices. A further split is made into financial and non-financial sectors (including the economic and market sectors, rating and maturity indices) with senior and subordinated debt indices being calculated for Financials and Non-Financials and for each financial sub-sector. Maturity indices are published for most index sub-groups.

The inception date for all the main iBoxx corporate bond indices is 11/04/2002 (although some indices start from 01/01/1998).

3.2 Given the main purpose of our research, the longest dataset was our favoured choice as we felt that it was by far the most credible with which to elicit an extreme event over one year. It also covered all of the major global economic events over the past century, including the oil price shocks in the 1970s, the 'Great Depression' and two world wars in addition to the more recent crises of the 21st century that are captured in the iBoxx and Merrill Lynch datasets.

3.3. The only drawback with the Moody's data is in terms of granularity as the data is not available for different durations. For that reason, we have also spent some time analysing the iBoxx dataset.

4. EMPERICAL OBSERVATIONS



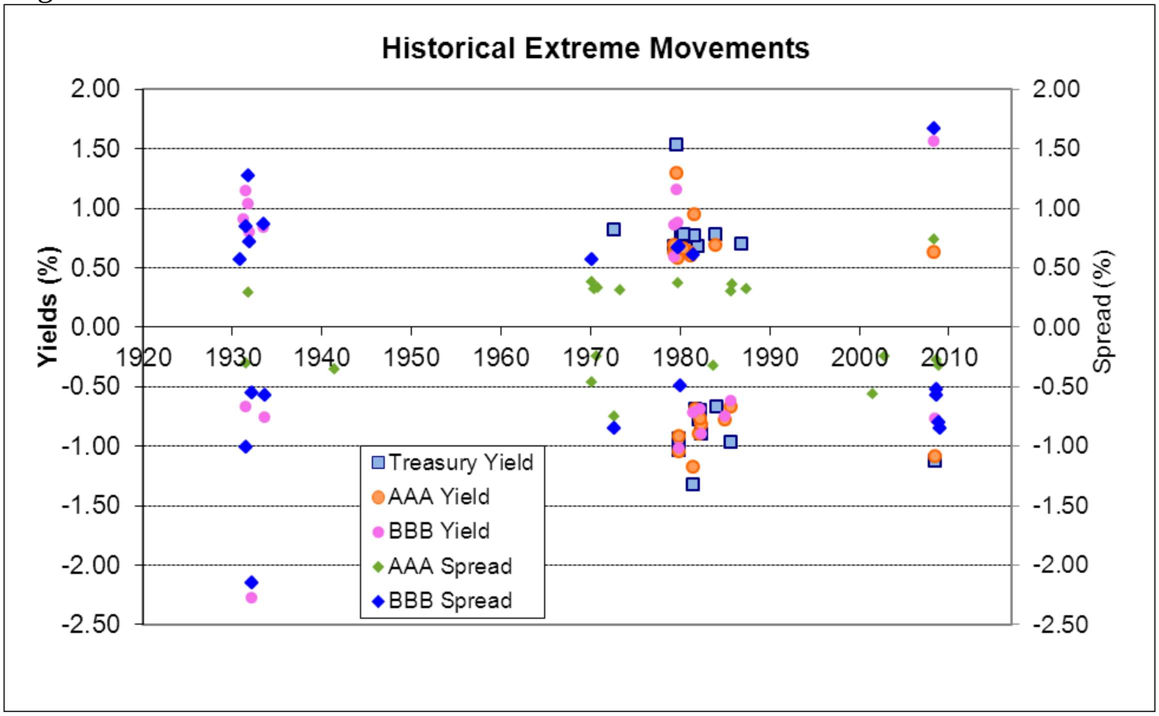
4.1 It is worth spending some time analysing the salient features of this graphical time series, which upon inspection appears to exhibit some stylized properties:

- No obvious trend is observed in the dataset.
- The spreads of different ratings behave largely in the same way; they are highly, but not perfectly correlated.
- Also the relative sizes of the spreads are largely maintained (i.e. BBB > A > AA > AAA) over the data period. Thus it appears that the ratings are good indicators of relative[4] credit quality.
- However, we may also notice the apparent existence of different credit regimes. For example the 20-year period post 1945 seems to have fundamentally different properties to the periods before and after.
- It is not immediately obvious, but the size of spread movements does appear to be linked to the spread levels.
- Finally, the occurrence of two extreme events stands out in the dataset; the Great Depression of 1929, and the more recent 'Credit Crunch' of 2008.

4.2 Another useful exercise is to simply look at the worst historical movements for the different credit ratings. The graph below shows the 10 most extreme movements in both directions for each of the data series, in spread space as well as yield space. Analysing the

[4] However, the absolute volatility appears to vary, as explained in the next point

data in yield space means that the 1930s and 2008 are less dominant, due to a number of large movements in the 1970s and 1980s.



4.3 We can extend this concept to a one year timeframe by considering the worst historical movements over (non-overlapping) 12 month and 18 month periods. This also helps to put the 2008 spread movements into context.

Max Annual Changes	AAA		AA		A		BBB	
1st	124	2008	156	2008	224	2008	346	2008
2nd	114	1986	109	1980	158	1932	271	1932
3rd	69	1998	109	1986	138	1982	193	1975

Max Changes over 18 months	AAA		AA		A		BBB	
1st	161	2008	197	2008	270	2008	409	2008
2nd	98	1986	98	1980	187	1932	340	1932
3rd	84	1975	84	1932	156	1980	202	1975

4.4 If we consider the combined impact on all investment grade credit, the credit crisis constitutes the worst period over the 90 years of spread data!

4.5 Analysing the figures above, it would be tempting to conclude that the 2008 credit crunch arguably constitutes the worst period for credit in over 90 years of data. However, we need to be aware of two important elements:

(a) The data above excludes the impact of credit defaults and downgrades. The resultant impact on defaults in the recent crisis has been less onerous, so far, compared to the decade following the 1930s.

(b) The data above relates to bonds with US long maturities only There were significantly greater spread stresses for shorted dated bonds in the recent crises. Spreads on shorter-dated bonds may have increased by more or less in the 1930s compared to the recent credit crunch period, although as we do not have the data to conclude either way.

(c) Sectors – The Moody's spread values represent the average of the Industrial and Public Utility components. Thus the shock to the railroad sector would have been fully included for the Great Depression, but not reflect the full impact of the shock to the financial sector in the recent crisis.

5. MODELLING

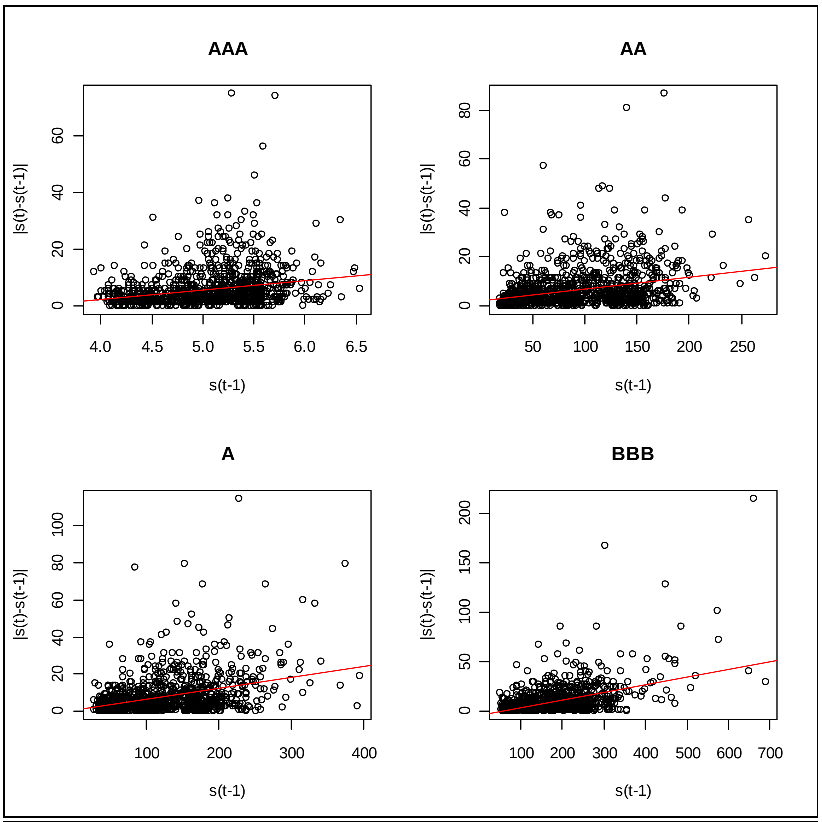
5.1 The next step was to address each of the stylized attributes, discussed in 4.1, and define a model. This would then be used together with Monte Carlo simulation to derive approximate spread stresses for different percentiles. A summary of the attributes is given below

- No obvious trend
- Lack of normality
- Volatility is related to spread
- Existence of different regimes
- Spreads of different ratings are connected

5.2 A naïve approach to estimating percentiles would be to choose the frequency so as to start off with a sufficiently large dataset (i.e. monthly in our case would give us approx 1000 points), empirically calculate the required percentile of the first differences and finally annualise them using an approximation, the most common of which is multiplication by $\sqrt{t}$. However this method makes fundamental assumptions on the properties of the first differences, in particular that they are independent, identically distributed and Gaussian. It can be shown that the spread time series violates *all* of these fundamental assumptions!

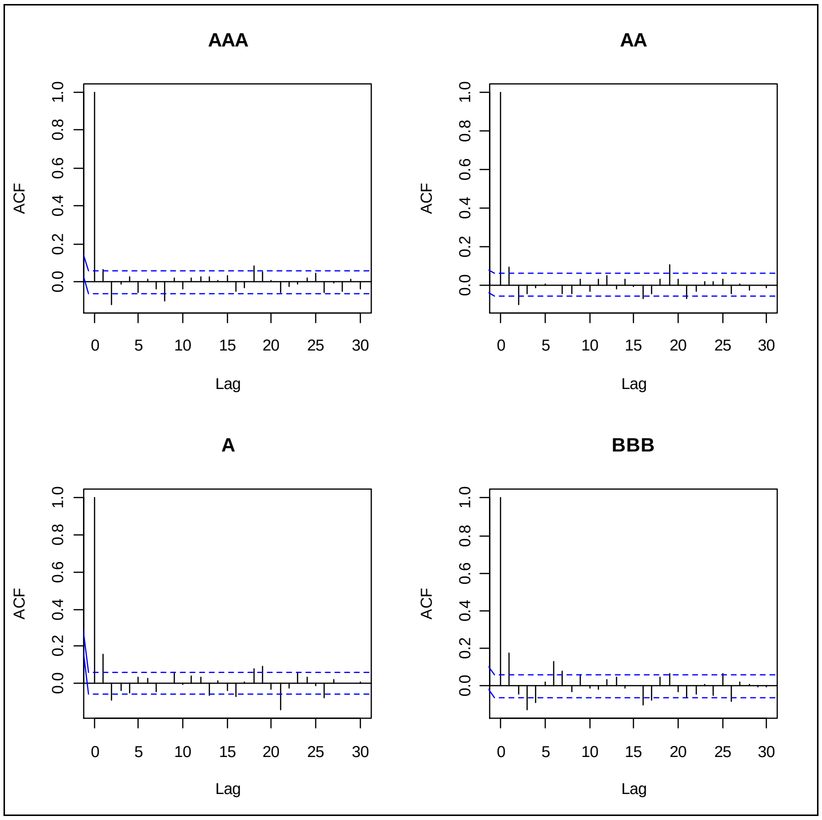
5.3 *Testing the 'identically distributed' assumption*:

This is a crucial property to satisfy when analysing time series, as it effectively defines how much of the dataset is valid for ex-ante estimates of spread stresses. Unfortunately, we can show with historical spread data that assuming spread changes are identically distributed is also fundamentally flawed. One way of depicting this is to plot the (absolute) spread changes against the level of the spread. If the spread changes were identically distributed, we wouldn't expect to find any relationship with the spread level. However, there is a clear indication that the absolute change increases with the level of the initial spread level, as can be seen from the graphs below. This is equivalent to the observation in 4.1 that the volatility of spread appears to have a clear link to the level of spread.



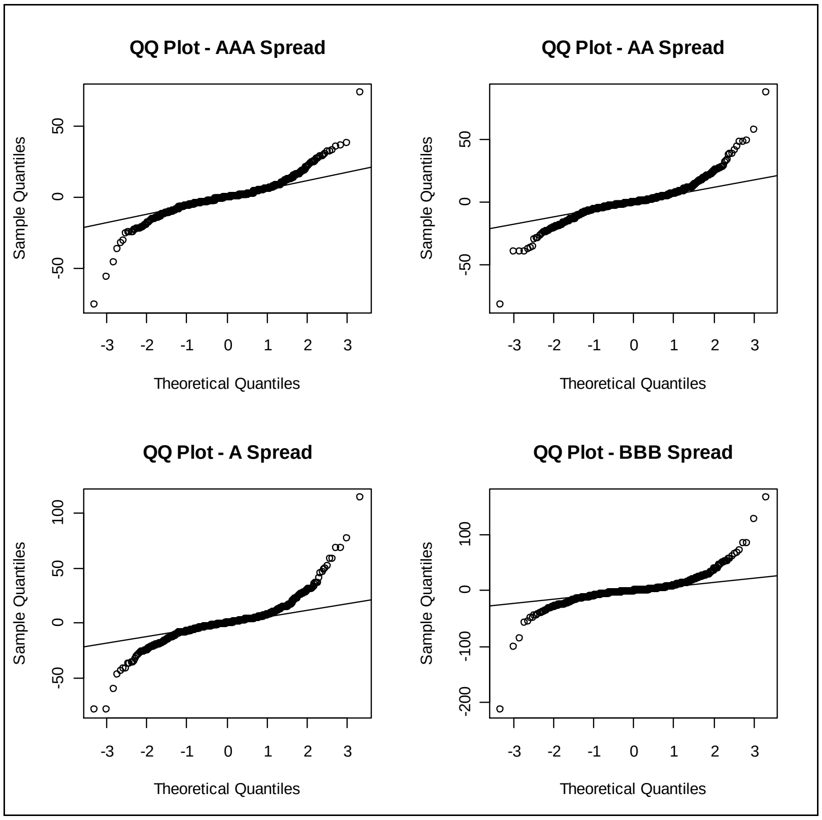
5.4 *Testing the independence assumption*:

We can also disprove the independence assumption by considering the autocorrelation of returns. Although there is no obvious autocorrelation of spread changes across the whole data set, this becomes significant when we consider absolute or squared changes instead, as can be seen in the graphs below. This means that although we may not be able to have a view on the direction of change based on recent history, it does give us information on the expected magnitude of the changes (i.e. that a large absolute change in the current period would tend to imply a large absolute change in the next period, and vice versa). This is evidenced in the following graph, which shows the ACF (Auto Correlation Function) with respect to absolute changes:



5.5 *Testing the normality assumption*:

The Q-Q plots below clearly illustrate the existence of fat tails, and the results of the Shapiro Wilkie normality test tabulated below also provide fairly conclusive evidence of non-normality



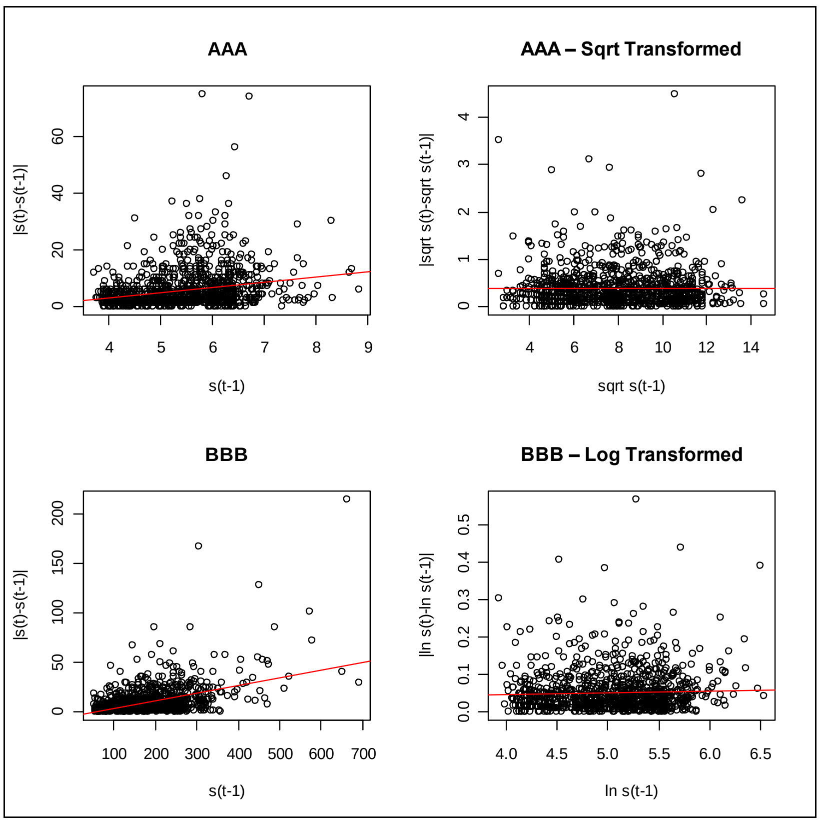
Results from the Shapiro-Wilkie normality test:

	AAA	AA	A	BBB
P.value of S-W Test	2.82E-28	7.19E-29	8.98E-32	7.69E-38

5.6 The intuitive idea for the modelling is, starting from first principles, to use the smallest number of transformations to correct the elements identified above and reduce the time series to independent, identically distributed random errors. This would then be assembled back up into a model to derive the spread stresses.

5.7 Converting the time-series into identically distributed elements is probably the most important as well as challenging task. The plots on the left-hand side of the graphic below carry out a least squares regression of the absolute change in spread verses the prevailing level of spreads. We can see a clear positive relationship in the un-transformed data, which tells us that the residuals are not identically distributed.

5.8 It appears that the best approach to nullify this relationship between the change in spread and the level of spread is to take an appropriate transformation of the time series. Initial investigations appeared to show that the most appropriate transformation depended on the credit rating, and gradually changed as we move up the credit ratings, resulting in stronger transformations (lognormal) for the lower credit quality (BBB). Although the concept of the transformation is interesting phenomenon, it requires more thought to explain on an intuitive level and is certainly something which may be worth further investigation.



5.9 A nice by-product of the transformation is that the resulting distribution of error terms has a greatly improved symmetry as well. This means that the autocorrelation observed in (5.4) can be allowed for in a relatively straightforward manner. A simple approach is to separately model the magnitude and the direction, which enables us to allow for correlation of the magnitude but not direction in the Monte Carlo simulation.

5.10 The normality can also be allowed for by using the simulated uniform random numbers to sample from a bespoke cumulative distribution function. These can be created by using a combination of empirical error terms for the body and results from extreme value theory for the tails. Another allowance, given enough data would be to simply sample the error terms from the residuals of the initial model fitting process.

5.11 The transformations do not fully allow for the existence of different regimes (i.e. looking at historical data suggests that the past can be broadly split into three distinct regimes). For this reason, combined with the doubts on the validity and robustness of very early data, our initial investigations were based on the past 50 years of data.

5.12 Finally, we also need to make a decision on the frequency of data. Quarterly data provided a good trade-off between the size of the dataset and the extra approximations required when grossing up to annual stresses.

5.13 An example set of results, based on simply using the 50-year dataset together with the methodology described above are shown below. Although the results would vary depending on the exact historical time period and methodology chosen, this gives a reasonably good overall picture of the dataset. The transformations used are identity, square root, cube root and log respectively.

	AAA	AA	A	BBB
Identity	121	140	176	242
Square Root	155	162	200	249
Cube Root	186	179	218	260
Log	404	255	293	318

5.14 It should be noted that, although some of the less visible parameters (i.e. autocorrelation of the direction and magnitude terms) should be informed by historical data, it is important to have an intuitive understanding for the values and apply a judgemental overlay. This may be relevant if we are deliberately aiming for a certain part of the distribution.

5.15 There is also scope for some judgemental overlay regarding the current spread levels in relation to longer term equilibrium spread levels. Thus, for example, if the in-house economist function believes that current spreads are significantly above or below equilibrium levels and there is likely to be some mean reversion, this can explicitly be incorporated into the methodology.

5.16 Finally, it should be noted that given we are using less than 200 years of data, there is still likely to be an element of spuriousness within the stresses, particularly when we consider individual credit ratings. As such there may be need for further smoothing across different percentiles and credit ratings. Also, it is advisable to run the results past sense checks from the in-house economists, who may spot an outlier result due to the spurious element described above.

5.17 Thus far, we have allowed for the existence of different regimes by deciding on the current regime and then using the appropriate subset of the historical data to calibrate our model. This may be a valid approximation in the current environment given that we are arguably in a volatile environment and focused on looking at upward spread shocks. However, this method is unlikely to have provided realistic shocks if we had used the historical data of a benign regime to calibrate shocks in early 2007.

5.18 There are different possibilities for incorporating the concept of 'regimes' into the modelling. The obvious intuitive way is to have a finite set of regimes, with their own set of parameters and probabilities of transitioning from one regime to another. However, you can see that the calibration problem becomes exponentially complex and arguably spurious as one increases the number or regimes. Thus we should stick to a small and intuitive number of regimes.

6. FURTHER ANALYSIS

6.1 We also analysed this problem using other datasets (iBoxx) as well as a number of other more complex time series models, in particular looking at the ARCH and GARCH family of models.

6.2 ARCH, which stands for Autoregressive Conditional Heteroskedasticity, is a family of models that considers the variance of the current error terms to be a function of the variance of the previous error term(s). The working party did some exploratory work on the GARCH series of models.

6.3 GARCH stands for ‘generalised autoregressive with conditional heteroscedasticity’. The GARCH(1,1) model for a time series (y_t) is defined by:

$$y_t = \mu + \sigma_t \varepsilon_t,$$
$$\sigma_t^2 = \omega + \alpha_1 (y_{t-1} - \mu)^2 + \beta_1 \sigma_{t-1}^2,$$

where the ε_t are i.i.d. random variables of unit variance and μ, ω, α_1 and β_1 are parameters to be determined, with ω, α_1 and β_1 non-negative.

6.4 We have fitted this model to both the iBoxx daily log return data for gilts, corporates and corporates in excess of gilts (Appendix A) and the absolute month-on-month changes in the Moody’s spread data for AAA and BBB. The model parameters have been estimated by maximum likelihood. The ε_t were assumed to be normally distributed. The following are the resulting parameter estimates and estimates for the 99.5% VaR over one year, based on 2007 YE, 2008 YE and 2010 YE levels:

Moody’s AAA Spreads	**2007 YE**	**2008 YE**	**2010 YE**
α_1	0.222	0.223	0.223
β_1	0.821	0.821	0.821
ω	8.04 E-09	7.803 E-09	7.944 E-09
VaR	178	230	234

Moody’s BBB Spreads	**2007 YE**	**2008 YE**	**2010 YE**
α_1	0.214	0.240	0.232
β_1	0.819	0.809	0.81
ω	9.475 E-09	8.261 E-09	9.391 E-09
VaR (bps)	232	662	259

6.5 These VaR estimates with a GARCH model are naturally very sensitive to the recent variance. In a normal / benign regime (e.g. 2007), the VaR estimates appear to be on the low side, whereas the estimates in an extreme circumstance (e.g. 2008 YE) would appear to be quite large – looking at the 662bps 1 in 200 stress for BBB spreads. This is unintuitive if one believes there is significant mean reversion to equilibrium spread levels. These also have the effect of being pro-cyclical and may cause unintended consequences for insurers. Finally, the estimate for iBoxx corporates (Appendix 1) in excess of gilts is excessively high and casts doubt on the appropriateness of this model.

6.6 It can be shown that the condition for a GARCH(1,1) model to be covariance stationary is that the sum of the parameters α_1 and β_1 should be less than 1. If this condition is not satisfied, then the absolute values of the y_t will in general increase without limit over time t. It could be argued that this is intuitively unreasonable where the y_t are log returns or credit spread changes. In fact, we can see from the table above that we have estimated $\alpha_1 + \beta_1 > 1$ for the iBoxx corporates in excess of gilts and for both the Moody's data sets, indicating that alternative models need to be considered to achieve covariance stationarity. For iBoxx gilts and corporates, $\alpha_1 + \beta_1$ is very close to 1, and so stationarity is only achieved over very long timescales.

6.7 Another potential issue is the serial correlation still inherent in the autocorrelation of the residuals for GARCH models. This also indicates that more work may need to be done to fully specify the model. This would need to be done in a way that doesn't inadvertently overfit the model by greatly increasing the number of parameters.

6.8 The serial correlation, relative instability (across time) and the 'unit root' issue (i.e. $\alpha_1 + \beta_1 > 1$) inherent in the GARCH model means that, after a fair amount of exploratory work, we did not consider GARCH to be a fruitful avenue in designing credit spread models. However, we have included useful pointers in the Appendix to comment on some of the approaches that we tried, and some potential further avenues to explore.

7. OTHER CONSIDERATIONS

7.1 There are various ways in which stress tests could be expressed. For instance they could be: a) a fixed addition to spread in basis points; b) a fixed multiple of the spread; c) a fixed % loss in market value etc. The appropriate structure will depend on prior beliefs. For instance b) implies that (on the same bond) if a spread of 50bps should be stressed to 100bps, then a spread of 100bps should be stressed to 200bps; some find this implausible. This has been investigated to some extent in the modelling by choosing the appropriate transformations (section 5.8), but a choice would still need to be made when communicating and disseminating the stresses.

7.2 Stresses could vary by term. We have not explored this so far. The Moody's US dataset is not split by term. It would be possible to split the iBoxx data by term buckets although obviously the number of issuers in each data cell would be diminished. Note that expressing credit stresses as a spread widening or as a reduction in market value both make implicit (and very different) assumptions about how stresses vary with term.

7.3 Spreads can be measured relative to government bond yields or swap rates. In our analysis we have used government bond yields. Swap rates history does not go back far enough. For any long data series (such as the Moody's US dataset) we are probably forced to analyse spreads relative to government bonds.

7.4 Our analysis has looked at corporate bond data; other considerations will apply for more complex credit instruments such as callable bonds, CDOs, asset backed securities etc.

7.5 Stresses could vary by sector (financials, industrials etc). However there are question marks regarding whether there is sufficient data to support this level of granularity without introducing spurious accuracy.

7.6 In addition the analysis we have done so far does not necessarily pick up all aspects of credit risk, and as explained in (2.7) the analysis of spreads using the Moody's US dataset does not pick up the impact of defaults and downgrades, except to the extent that bond spreads widen in anticipation of a downgrade/default event occurring.

7.7 Many credit indices are re-balanced only periodically (e.g. once a month), which could introduce excess volatility in the index. Let us illustrate this with a simple example:

- Suppose a bond index of AAA bonds is rebalanced on the first trading day of each month and that the yield on AAA bonds does not change over time.

- Suppose then that a significant proportion of the index (e.g. one or two large constituent issuers) is downgraded to AA sometime during that month.

- This will cause the average yield or spread on the index as the yield on the downgraded bonds increases and they are still included in the index

- On the first trading day of the following month, the index is rebalanced and these bonds are no longer included in the AAA index but are instead included in the AA index.

- This will then cause the yield or spread on the index to fall to its original level before the bonds were downgraded.

- In reality therefore, although spreads on AAA bonds have actually remained static, there has been upwards and downwards movement in the index. This has then caused us to overestimate the index volatility as positive and estimating a stress to AAA spreads whereas in fact volatility of AAA spreads is zero and so is the stress to AAA spreads.

- The same effect would happen, in reverse, in the case of upgrades e.g. if AA bonds were upgraded to AAA, causing yields on the AA index to fall and then to increase back to their original level after the index was re-balanced.

These simple examples illustrate how relatively infrequent re-balancing of an index can induce excess volatility.

7.8 Excess volatility in an index may also be induced where the index is if the index consists of a small number of constituents (e.g. the iBoxx AAA index for UK corporate bonds has typically consisted of one or two constituents over the last two years) or when an index is disproportionately weighted towards a small of constituents (e.g. using market values as weights).

7.9 For any business where the liability discount rate includes an allowance for a liquidity premium, any discussion of credit risk is incomplete without consideration of how the liquidity premium changes under the credit stresses applied. There is considerable research, controversy and debate on this topic elsewhere!

Appendix A – Further analysis on GARCH models

A.1 Fitting the GARCH model to a shorter term dataset resulted in unintuitive results: This is partly to the paucity of the dataset, in that the model fitting is ‘blind’ to the context, and assumed that the extreme event seen in the short dataset is representative of the whole

iBoxx Data	**Gilts**	**Corporates**	**Corporates XS Gilts**
μ	1.838×10^{-4}	1.798×10^{-4}	5.081×10^{-5}
α_1	0.03073	0.02506	0.1843
β_1	0.9662	0.9714	0.8401
ω	3.789×10^{-8}	3.158×10^{-8}	4.934×10^{-9}
VaR	25.97%	17.18%	114.03%

A.2 We should note that the iBoxx data sets contain both a benign period, from 1998 to 2007, and the much more volatile period of 2008 and early 2009. In the Moody’s data sets, both the period of the Great Depression and the 2008-early 2009 period show much greater volatility than other periods. One modelling approach that might be considered in response to this is what is known as a *threshold GARCH* model, under which the parameter α_1 takes a different value according to whether y_{t-1} is above or below a threshold y_0. Typically we would expect α_1 to be lower above than below y_0, with the interpretation that adverse events tend to be followed by periods of high volatility. The main drawback of this approach is that little research has been done into systematic procedures for estimating y_0. The most commonly used procedures are graphical, and this introduces an element of subjectivity to the estimation process. We shall not pursue this approach further here.

A.3 Another possible modelling approach is a *Markov switching GARCH* model, under which the parameters of the GARCH process at a given time depend upon the state of an unobservable Markov chain representing the ‘regime’ the market is in. Dueker (1997) fits four different formulations of this model to US equity return data, using a two-state Markov chain. It is found that for this data set, one of these formulations, the model ‘GARCH-DF’, achieves a particularly significant improvement in fit over the basic GARCH(1,1) model.

A.4 The paper by Dueker does not describe the iterative algorithm used to estimate the parameters. To date the Working Party has been unable to estimate the parameters of any of the models in the paper for either the iBoxx or the Moody’s credit data. It is not known whether this is because the algorithms the Working Party has attempted to use do not converge or whether they converge extremely slowly. These difficulties in fitting must be regarded as a drawback of using such more complex models.