Inter-temporal Stability of the European Credit Spread Co-movement Structure

Jan Annaert
Ghent University and University of Antwerp
St. Pietersplein 5
9000 Ghent
Belgium
jan.annaert@rug.ac.be

Anouk G.P. Claes
University of Antwerp
Prinsstraat 13
2000 Antwerp
Belgium
anouk.claes@ufsia.ac.be

Marc J.K. De Ceuster*
University of Antwerp
Prinsstraat 13
2000 Antwerp
Belgium
marc.deceuster@ufsia.ac.be
&
Director Risk Management Deloitte & Touche
Lange Lozanastraat 270
2018 Antwerp
Belgium
mdeceuster@deloitte.com

* Corresponding author

Abstract

Corporate bonds expose the investor to credit risk, which will be reflected in the credit spread. Based on the EMU Broad Market indices, we study the inter-temporal stability of the covariance and correlation matrices of credit spread changes. Within a multivariate framework, the Box and Jennrich tests are most commonly used test statistics in the literature. However, we show that for small samples these tests are not well specified when the normality assumption is relaxed. A bootstrap-based statistical inference provides evidence that correlations between various (investment grade) credit spread changes remain stable over the 1998-2000 period. Covariances on the other hand, turn out to be time-varying over that period.

JEL: Classification: G110, G150

Keywords: credit spreads, diversification, correlations, corporate bonds

1. Introduction

Recently, yields on government bonds have come to historical lows. Many investors try to pick up a higher return by investing in the corporate bond market. In order to be attractive to investors, risky bonds need to offer a higher yield than risk-free bonds with the same maturity and coupon rate, thereby compensating for the credit risk borne. The difference between the yield on the risky bond and the yield of the equivalent risk-free bond is defined as the credit spread. This spread should not just compensate the investor for the expected loss on the risky bond, but it should also incorporate a risk premium in order to recompense the investor for accepting the risk to suffer higher losses.

Credit spread risk, especially for European bonds, is hardly documented. Annaert *et al.* (2000) and Kiesel, *et al.* (2001) are notable exceptions. This paper focuses on the co-movement dynamics of credit risk. Obviously a profound knowledge of the time series characteristics of correlations is important for various financial applications. Portfolio optimisation requires precise estimates of expected returns, variances and correlations/covariances. Risk managers will need to model the correlation dynamics in order to implement Value at Risk or to build sensible stress testing scenarios. The pricing of credit linked copulas also presupposes a profound understanding of the multivariate time series dynamics of credit spreads. In order to contribute to our understanding of credit spread risk, this paper examines whether the co-movement structure is constant over time. Using appropriate statistical tests, we will show that both (unconditional) variance-covariance matrix and the (unconditional) correlation matrix are relatively stable over the time period studied. Obviously, these results are important for e.g. risk management and for choosing appropriate multivariate time series models.

The remainder of this paper is structured as follows. The next section (section 2) will describe the data used and will provide details on the construction of our credit spread data set. Descriptive statistics of the credit spreads will be discussed subsequently (section 3). In section 4 the appropriateness of several approaches to test the inter-temporal stability of covariance and correlation matrices is scrutinised. Given the multivariate nature of the issue, we followed Kaplanis (1988) in using the Box test on the equality of covariance matrices and the Jennrich test on the equality of covariance or correlation matrices. Sub-section 4.1 defines the test statistics and discusses the conjecture of Tang (1995, 1995b) to use the Box test (on covariances) on standardised data in order to implicitly test for stability of correlations, Sub-section 4.2 will document two problems with the proposed tests. First, we show that the suggestion of Tang (1995, 1995b) to perform the Box test on standardised data in order to test equality of correlations leads to misspecifications under the null and should hence not be used. Second, we provide additional insights to the well-known result that departures from the assumption of multivariate normally distributed data lead to extremely misspecified tests. Fat tailedness, which is a stylised fact of all kinds of financial data, causes the standard 5% type I error to boost to over 80% causing unwarranted rejections of the null. This size misspecification is seriously underestimated in the finance literature (e.g. even the monthly data used by Kaplanis (1988) probably are hampered with some leptokurtosis) where these kind of tests lead authors to conclude in favour of unconditional instability. In order to avoid this flaw, we apply (in section 5) the Zhang and Boos (1992) bootstrap procedure to both the Box test (for covariances only) and the Jennrich test (for both covariances and correlations). It turns out that correlations between various (investment grade) credit spread changes remain stable over time. Covariances on the other hand turn out to be time-varying over the period under study. The last section concludes.

2. Data and spread construction

In order to calculate credit spreads, we obtained bond index data constructed by Merrill Lynch. Its EMU Broad Market indices are based on secondary market prices of bonds issued in the Eurobond market or in EMU-zone domestic markets and denominated in euro – or one of the currencies that joined the EMU. Besides direct government bond indices, investment grade corporate bond indices were used as well. The latter are based on publicly traded bonds, issued by companies domiciled in the European Union and exclude convertible securities. For our study, we will use the sub-indices for 5 maturity classes and for 3 rating categories (see Table 1 for the exact indices used). This gives us a set of 15 indices with differing maturity and credit quality.

All these indices are based upon the composite rating of Moody's and Standard & Poor's, if the issue is rated by both. If ratings do not coincide, an average rating is used which is 'rounded' downwards. The composition of each index is determined on the last business day of the previous month. During the month, each bond will stay in the index, regardless whether or not the bond is downgraded or upgraded, or whether the maturity no longer fits the classification. Also when bonds are called during the month, they are not removed from the index until the end of the month. A similar rule holds for changes in the amount outstanding during a month: face values are kept constant and are adjusted at the start of a new month. All issues have a fixed coupon (including step-ups) and a minimum remaining term to maturity of one year.

Table1: Average Number of Issues and Average Market Value Outstanding

	All Ratings	AAA	AA	A
All Maturities	937.14	308.29	362.12	223.35
	(340.89)	(111.61)	(128.39)	(84.84)
1-3 years	244.35	92.62	99.77	45.09
	(65.17)	(24.03)	(26.34)	(13.14)
3-5 years	259.26	80.95	102.22	61.90
	(92.87)	(28.17)	(34.14)	(23.99)
5-7 years	151.57	51.26	44.77	40.15
	(59.57)	(19.78)	(17.45)	(17.24)
7-10 years	224.32	58.03	93.06	66.30
	(98.97)	(27.06)	(41.49)	(27.70)
10-more years	57.64	25.44	22.30	9.91
	(24.31)	(12.57)	(8.96)	(2.77)

Calculations are based upon Merrill Lynch data. Ratings are composite Moody's and Standard and Poor's ratings. Maturity buckets include the lower boundary and exclude the upper boundary. Average number of issues is computed from daily data over the period 31 March 1998 through 29 November 2000 (693 observations). Market values are expressed in billions of euro and are printed in parentheses.

Although the bond index data start on 31 December 1995, daily data are only available from 31 March 1998 onwards. This implies that our daily sample, which runs till 29 November 2000, only contains 693 observations per index. In Table 1 we indicate the average number of issues and the average market value for each maturity bucket and rating class over this sample period. From this Table, it can be seen that both the number of issues and the total market value of issues indicate that bonds with maturities longer than 10 years are relatively scarce. Moreover, the average maturity (shown together with other characteristics in Table 2) of these bonds range between 10.95 years (A-rated bonds) up to 12.41 years (AAA-rated bonds), implying that more than ten years is effectively not really that long a maturity.

From Table 2 no clear pattern can be discerned as to the relation between rating class and average maturity within maturity baskets. However, there is a visible difference between the average coupon rates: as could be expected, coupon rates and yields to maturity are generally larger for lower rated bonds. As the 'coupon effect' may have an impact on the relation between interest rates and yield spreads, this should be taken into account in the empirical analysis. Also, bond indices with longer maturities contain larger issues on average than the shorter dated issues. This may have an impact on the liquidity premia contained in the bond spread.

Table 2: Average Characteristics for the Merrill Lynch Euro Corporate Bond Indices

		Maturity	Modified Duration	Conven-tional Yield to Maturity	Coupon	Average Market Value per Issue
1-3 years	AAA	2.00	1.82	4.26	5.84	259.8
	AA	1.97	1.79	4.35	6.36	263.3
	A	2.06	1.85	4.49	6.45	295.0
3-5 years	AAA	4.02	3.47	4.58	5.45	345.8
	AA	3.95	3.39	4.70	5.80	337.0
	A	4.14	3.55	4.91	5.83	368.4
5-7 years	AAA	5.89	4.81	4.84	5.79	393.7
	AA	5.90	4.80	4.95	5.87	398.1
	A	6.00	4.85	5.23	5.93	437.0
7-10 years	AAA	8.59	6.55	5.17	5.64	463.3
	AA	8.70	6.65	5.32	5.52	435.9
	A	8.85	6.67	5.59	5.74	405.4
10+ years	AAA	12.41	8.52	5.33	5.64	501.4
	AA	11.55	8.64	5.47	5.46	416.0
	A	10.95	8.11	5.84	5.73	278.6

Calculations are based upon Merrill Lynch data. Ratings are composite Moody's and Standard and Poor's ratings. Maturity buckets include the lower boundary and exclude the upper boundary. Averages are computed from daily data over the period 31 March 1998 through 29 November 2000 (693 observations). 'Maturity', 'Modified Duration' and 'Conventional Yield to Maturity' are equally-weighted averages over all issues. 'Coupon' is a market value-weighted average. 'Average Market Value per Issue' is expressed in billions of euro.

Although Merrill Lynch computes option-adjusted spreads, these data are only available since 31 March 1999, which leaves us with far too few observations to do any analysis. We therefore had to compute yield spreads simply as the difference between the yield to maturity as reported by Merrill Lynch and a comparable government bond yield. Because of the differences in coupon rates across rating categories, shown in Table 2, we chose to take the yield of a government bond with similar duration rather than similar maturity.

Table 3 Average Characteristics for the Merrill Lynch AAA-rated Euro-Sovereigns Bond Indices

	Number of Issues	Market Value Outstanding	Maturity	Modified Duration	Conventional Yield to Maturity	Coupon
All issues	158.51	1232.9	7.20	4.97	4.59	6.47
1-3 years	49.63	292.2	1.95	1.77	4.01	6.32
3-5 years	40.71	293.4	3.96	3.39	4.31	6.10
5-7 years	24.72	198.6	5.93	4.77	4.53	6.62
7-10 years	25.36	283.8	8.45	6.52	4.77	5.56
10-more years	18.10	164.8	22.30	11.65	5.29	6.54

Calculations are based upon Merrill Lynch data. Maturity buckets include the lower boundary and exclude the upper boundary. Averages are computed from daily data over the period 31 March 1998 through 29 November 2000 (693 observations). 'Maturity', 'Modified duration', 'Conventional yield to maturity' and 'Coupon' are averages over all issues. 'Market value' is in billions of euro.

The government bond yields used were the yields of indices of AAA-rated Euro-sovereigns, also computed by Merrill Lynch. Average characteristics for these bond indices can be found in Table 3. There are fewer government issues than corporate bond issues, but the former are much larger, as evidenced by the average "Market Value Outstanding" column. Obviously, average yields for government bonds are lower than the corporate bond yields. However, we did not compute spreads by simply subtracting the government bond yield from the bond yields in the corresponding maturity bucket. This would not properly account for potential differences in maturity or coupon rate between the government and the corporate bond index. Therefore, for each corporate bond index and each day in our sample we subtracted the yield of the government bond index with the same average (modified) duration as the corporate bond index. Of course, in most cases the average duration of the corporate bond index did not match the average duration of a government bond index. In these cases we took an interpolated government bond yield. This yield was exponentially interpolated from the two adjacent government bond yields.[1]

3. Descriptive statistics of the credit spreads

Summary statistics of the daily series can be found in Table 4. As could have been expected, average spreads increase monotonically the lower the credit rating. The relation is clearly not linear: the difference between the BBB-rated indices and AA-rated indices is generally much higher than between other adjacent rating classes. These spreads are considerably lower than the spreads reported by Duffee (1998). For AAA-rated bonds, our spreads vary between 24 basispoints (bp) and 43 bp, whereas Duffee finds at least 67 bp. The picture is similar for the other rating categories. Of course, our data period is much shorter than Duffee's and does not overlap with his. Our average spreads do seem consistent with the findings of Pedrosa and Roll (1998) for US investment grade spreads covering the period 1995-1997. Unfortunately, they do not provide an estimate of the average spread on the US market, but the graphs they present, show spread levels similar to our data.

---

[1] The corresponding author will provide further details of the credit spread construction on request.

Table 4 Summary Statistics of Yield Spreads

		Average Spread (in %)	Standard Deviation (in %)	First Order Auto-correlation	Second Order Auto-correlation	Third Order Auto-correlation	Ljung-Box Statistic at 20 lags
1-3 years	AAA	0.235	0.074	0.946	0.935	0.931	10707
	AA	0.328	0.059	0.957	0.938	0.928	10752
	A	0.457	0.099	0.855	0.814	0.792	7255
3-5 years	AAA	0.244	0.081	0.962	0.949	0.938	10723
	AA	0.381	0.086	0.962	0.947	0.938	10835
	A	0.567	0.134	0.978	0.967	0.957	11066
5-7 years	AAA	0.298	0.095	0.972	0.959	0.947	10844
	AA	0.416	0.135	0.982	0.973	0.967	11911
	A	0.688	0.216	0.987	0.979	0.971	12170
7-10 years	AAA	0.431	0.106	0.969	0.950	0.935	10167
	AA	0.608	0.125	0.975	0.961	0.951	10771
	A	0.881	0.203	0.987	0.979	0.971	11747
10+ years	AAA	0.408	0.111	0.973	0.961	0.950	10933
	AA	0.539	0.129	0.961	0.951	0.932	10412
	A	0.967	0.232	0.988	0.981	0.974	12032

Calculations are based upon Merrill Lynch data. Yield spreads are spreads relative to government bond yields with similar duration. Ratings are composite Moody's and Standard and Poor's ratings. Maturity buckets include the lower boundary and exclude the upper boundary. Statistics are computed from daily data over the period 31 March 1998 through 29 November 2000 (693 observations). For the A-rated 1-3 year bonds, one negative spread, possibly due to a data error, was replaced by the average of the adjacent spreads.

As far as the relation between spread and maturity is concerned, we generally find an upward sloping credit curve, as was also the case for Duffee (1998). Nevertheless, in some cases the relation is not monotone: see e.g. '10 and more years' bucket for AAA and AA-rated bonds. This might be due to a liquidity effect, since the longer dated issues are the largest. In addition, it may be the case that relatively less credit risky issuers issue longer dated bonds.

Furthermore, we note that, as in Longstaff and Schwartz (1995), the standard deviation of spreads is generally also increasing when credit rating deteriorates. No clear relation with maturity can be spotted. Finally, we also observe that the first lag auto-correlation coefficient of all spread series is relatively large, and often near 0.95 or higher. We do not present stationarity tests because of the reported low power of unit root tests on observation periods as short as this one. In any case, if we accept on economic grounds that the credit spread series are stationary, it is clear from the high auto-correlation coefficients that they revert only slowly to their long-run average.

In Table 5 summary statistics for *changes* in credit spreads are presented. To some extent these are more important in a risk management context, as it is important to understand how credit spreads behave through time. We present descriptive statistics for both daily and weekly data. The descriptives on daily data are more comparable to the existing literature. Our research on inter-temporal stability of co-movements will be applied to weekly data in order to obtain better estimates of correlations and covariances (see section 5).

It can be noticed that average daily changes in spreads are very small and insignificantly different from zero. This could have been expected as they measure the trend of credit spreads over the time period investigated. When we look at the standard deviation of spread changes, there is not a clear relation to rating class. This contrasts both with Pedrosa and Roll (1998), who find lower standard deviations for lower rating classes, and with Duffee (1998) or Longstaff and Schwartz (1995) who find the opposite effect. Looking at the maturity class, we neither have a clear picture, which is consistent with Duffee (1998), but not with Pedrosa and Roll (1998), who find decreasing volatility for longer maturities. These inconsistent results cannot be explained by the calculation method of the spreads. Whereas Duffee (1998) calculates spreads as simple first differences, Longstaff and Schwartz (1995) and Pedrosa and Roll (1998)) use logarithmic differences instead. Re-computing our standard deviations using logarithmic differences does not alter our results.

Table 5: Summary Statistics of Yield Spread Changes

Panel A: Changes of the daily spreads

		Average Spread Change (in bp)	Standard Deviation (in bp)	Skew-ness	Kurtosis	First Order Auto-correlation	Second Order Auto-correlation	Third Order Auto-correlation
1-3 years	AAA	0.026	0.0240	-0.92*	48.04*	-0.424*	-0.054	0.058
	AA	0.016	0.0171	0.09	7.27*	-0.290*	-0.110*	-0.058
	A	0.014	0.0513	0.14	24.11*	-0.363*	-0.063	0.030
3-5 years	AAA	0.040	0.0209	0.05	6.80*	-0.370*	-0.023	-0.022
	AA	0.039	0.0226	-0.02	14.08*	-0.337*	-0.080*	-0.016
	A	0.081	0.0228	-0.02	7.29*	-0.338*	-0.007	-0.037
5-7 years	AAA	0.045	0.0207	-0.01	5.77*	-0.333*	0.003	-0.046
	AA	0.057	0.2342	0.37*	12.29*	-0.341*	-0.055	0.002
	A	0.104	0.0278	-0.24*	10.05*	-0.228*	-0.019	-0.059
7-10 years	AAA	0.047	0.0244	-0.71*	16.30*	-0.216*	-0.052	-0.004
	AA	0.062	0.0254	0.17	19.62*	-0.258*	-0.089*	0.100*
	A	0.117	0.0233	1.02*	17.22*	-0.286*	-0.037	0.021
10+ years	AAA	0.048	0.0240	-0.20*	11.70*	-0.316*	-0.013	0.001
	AA	0.079	0.0325	0.06	41.00*	-0.456*	0.151*	-0.121*
	A	0.105	0.0295	0.26*	9.42*	-0.298*	0.026	0.016

Panel B Changes of (non-overlapping) weekly spreads

		Average Spread Change (in bp)	Standard Deviation (in bp)	Skew-ness	Kurtosis	First Order Auto-correlation	Second Order Auto-correlation	Third Order Auto-correlation
1-3 years	AAA	0.13	0.0225	-0.02	4.43*	-0.269*	-0.100	0.114
	AA	0.08	0.0223	-0.01	3.11	-0.323*	0.068	-0.109
	A	0.11	0.0665	0.30	9.83*	-0.387*	-0.168	0.172*
3-5 years	AAA	0.19	0.0277	0.43*	4.93*	-0.271*	-0.105	0.021
	AA	0.19	0.0283	0.23	4.26*	-0.227*	-0.020	-0.158
	A	0.37	0.0317	0.06	5.07*	-0.159	0.045	0.008
5-7 years	AAA	0.22	0.0293	0.23	4.14*	-0.224*	-0.008	-0.019
	AA	0.28	0.0300	0.88*	5.23*	-0.167	0.004	-0.066
	A	0.49	0.0461	0.02	8.61*	-0.308*	0.027	-0.028
7-10 years	AAA	0.23	0.0412	-0.57*	7.75*	-0.210*	-0.124	0.023
	AA	0.29	0.0404	0.23	8.94*	-0.172*	0.002	-0.023
	A	0.56	0.0402	0.45*	12.29*	-0.090	-0.096	-0.011
10+ years	AAA	0.24	0.0342	0.63*	4.49*	-0.191*	-0.092	-0.150
	AA	0.39	0.0336	-0.22	6.21*	-0.234*	0.051	-0.056
	A	0.52	0.0479	0.17	5.55*	-0.130	0.100	-0.145

Calculations are based upon Merrill Lynch data. Yield spreads are spreads relative to government bond yields with similar duration. Ratings are composite Moody's and Standard and Poor's ratings. Maturity buckets include the lower boundary and exclude the upper boundary. Statistics in Panel A are computed from daily data over the period 1 April 1998 through 29 November 2000 (692 observations). Panel B provides the results of 138 weekly spreads. Weekly spreads were constructed by summing 5 consecutive daily spreads. Asterisks denote auto-correlation, skewness or excess kurtosis coefficients more than two (asymptotic) standard deviations away from zero.

Table 5 Panel A also shows the first three order auto-correlation coefficients of the spread changes. Without any exception, all first order coefficients are negative and usually significantly different from zero.[2] This is in contrast to Duffee (1998) who uses monthly data and mostly finds positive auto-correlations. The negative auto-correlation may be a reflection of poor liquidity in the European corporate bond market. By frequently bouncing between the bid and the ask quote, negative auto-correlation may be introduced. This explanation is unlikely, however, since aggregation of the daily spreads to weekly spreads does not mitigate this negative correlation (Compare with Panel B). Higher order auto-correlation does hardly seem present in the series.

In any case, from the skewness and kurtosis coefficients in Table 5, it is clear that spread changes are not normally distributed. This is especially due to excess kurtosis. All kurtosis coefficients are significantly higher than three, which implies that the distributions have higher peaks and thicker tails than the normal distribution does. If the latter distribution is used for risk management purposes, this means that more extreme changes would effectively occur than expected. Therefore, 99% Value-at-Risk measures based upon the normal distribution will be underestimating true risk and capital requirements will be insufficient. Table 5 Panel B summarises the same descriptive statistics for weekly data. In general the same stylised facts pop up. Excess kurtosis is much lower but still overwhelmingly indicating non-normality. The first order auto-correlation is somewhat less predominantly present but remains exclusively negative.

---

[2] The asymptotic standard deviation of the auto-correlation coefficients is $\sqrt{1/692} = 0.038$ for the daily data and $\sqrt{1/138} = 0.0851$ for weekly data.

4. Methodological issues in assessing inter-temporal stability of co-movements

Several approaches have been advanced to study the multivariate structure of stock returns. Principal component analysis e.g. reveals the number of (significant) risk factors (Makridakis and Wheelwright (1974), Philippatos and Christofi (1983) and Cheung and Ho (1991)). Obviously, one can perform principal component analysis over several sub-periods and check whether the same risk factors pop up. This, however, is not a test on the stability of the covariance and/or correlation structure. The same critique applies to applications of cluster analysis applied by e.g. Cheung and Ho (1991). Moreover, the choice of the distance metric and the clustering method inevitably brings some arbitrariness into the research design. The Normal Distribution test, as used by Haney and Lloyd (1978), Watson (1980), Maldonado and Sounders (1981) and Cheung and Ho (1991), also has its weaknesses. Using this test, the evaluation of the null hypothesis of equality of a correlation coefficient between two time periods becomes possible. As it is merely a pair-wise method, it only tests the stability of the individual correlation coefficients. Meric and Meric (1989) indicate that this test can lead to misleading results, since in the formation of ex-ante optimal portfolios our major concern is the stability of the entire matrix, not just the stability of the individual coefficients. Moreover, it is well known that independent univariate tests may reject the null hypothesis while the multivariate test does not. However, the reverse is possible as well. Hence, we follow Kryzanowski and To (1987) and Kaplanis (1988) in applying multivariate tests on equality of two covariance/correlation matrices across different time periods. Traditionally, two tests have been advanced for this purpose: one developed by Box (1949), another by

Jennrich (1970). Kaplanis (1988) notes that both tests are asymptotically equivalent, but for small samples, the test results may lead to different conclusions. The Jennrich test is designed for testing both the equality of covariance and correlation matrices. The Box test, however, only applies to covariance matrices. Tang (1995, 1995b) mistakenly - as we will show - conjectured that the Box test could also be used to test for the equality of correlation matrices provided that the data are first standardised.

In this section we will first define the test-statistics used. Next, we will examine their size, which will reveal the shortcomings of the Tang (1995, 1995b) procedure and the extreme sensitivity to departures from multivariate normality.

4.1 Test statistics

4.1.1 The Box test on equality of covariance matrices

The Box (1949) test tests for the equality of the two covariance matrices and consequently avoids the problems of the pair-wise test. Let C_1 and C_2 be the population covariance of the first and second period. The null hypothesis that the covariance matrix is inter-temporally stable is equivalent to the hypothesis that the matrices across the two different time periods are jointly equal to each other. We therefore test the hypothesis:

$$H_0\text{: } C_1 = C_2$$

Following Morrison (1976), the Box test statistic, M, can be computed as follows:

$$M = (T^{(0)}\text{-}2)\ \ln|V| - (T^{(1)}\text{-}1)\ \ln|V^{(1)}| - (T^{(2)}\text{-}1)\ \ln|V^{(2)}|,$$

where: ln(.) denotes the natural logarithm and |.| stands for determinant;

$T^{(i)}$ denotes the number of observations on the N bonds, where i = 0, 1, 2; and $T^{(0)} = T^{(1)} + T^{(2)}$;

V denotes the pooled covariance matrix based on biased sub-period covariance matrices;

$V^{(i)}$ denotes the unbiased sample covariance matrix of asset returns, where i = 1 or 2.

A Chi-squared approximation can be obtained for M by multiplying it by a scalar $c = (1-A_1)$, where

$$A_1 = \frac{2N^2 + 3N - 1}{6(N+1)}\left(\frac{1}{T^{(1)}-1} + \frac{1}{T^{(2)}-1} - \frac{1}{T^{(0)}-2}\right)$$

The Chi-squared approximation has d_1 degrees of freedom, with $d_1 = \frac{N(N+1)}{2}$. However, when N exceeds 4 or 5, this approximation might not be appropriate anymore. In those cases an F-approximation is preferable (Cooley and Lohnes, 1971). If we define

$$A_2 = \frac{(N-1)(N+2)}{6}\left(\frac{1}{(T^{(1)}-1)^2} + \frac{1}{(T^{(2)}-1)^2} - \frac{1}{(T^{(0)}-2)^2}\right)$$

we can discern two cases:

Case 1: $A_2 > A_1^2$

Let $d_2 = \frac{d_1 + 2}{A_2 - A_1^2}$ and $b = \frac{d_1}{1 - A_1 - \frac{d_1}{d_2}}$, then $\frac{M}{b}$ is approximately distributed as an F-distribution with d_1 and d_2 degrees freedom.

Case 2: $A_2 < A_1^2$

In this case, let $d_2 = \frac{d_1 + 2}{A_1^2 - A_2}$ and $b = \frac{d_2}{1 - A_1 - \frac{2}{d_2}}$, such that $\frac{d_2 M}{d_1(b - M)}$ is approximately distributed as an F-distribution with d_1 and d_2 degrees of freedom.

4.1.2 The Adjusted Box test on equality of correlation matrices

Tang (1995, 1995b) re-observed[3] that this Box statistic on covariance matrices can also be used for testing the equality of correlation matrices when performed on standardised data, since for standard scores the two equal. Moreover since the correlation between two random variables equals the correlation between the standardised variables, Tang (1995, 1995b) thought to have found a simple way to test for the equality in correlation structures.

4.1.3 The Jennrich-test on the equality of correlation and covariance matrices

The Jennrich test can be used either to test the equality of covariance matrices or to test the equality of correlation matrices. If we want to evaluate the structure of covariance matrices, we define:

[3] Although Tang (1995, 1995b) does not quote an older working paper of Gibbons (1981), Kryzanowski and To (1987) cites that Gibbons also standardised his data to apply the Box test on correlation matrices.

$$Z = \left(\frac{T^{(1)} T^{(2)}}{T^{(1)} + T^{(2)}} \right) \overline{V^{-1}} \left(\overline{V^{(1)}} - \overline{V^{(2)}} \right).$$

The Jennrich (1970) statistic J_1 to test equality of two covariance matrices is computed as follows:

$$J_1 = \frac{1}{2} tr(Z^2)$$

where tr(.) denotes the trace-function. The J_1 is asymptotically distributed as a chi-squared with $\frac{N(N-1)}{2}$ degrees of freedom.

Testing the equality of correlation matrices can be done by defining Z as follows:

$$Z = \left(\frac{T^{(1)} T^{(2)}}{T^{(1)} + T^{(2)}} \right) C^{-1} \, (C^{(1)} - C^{(2)}) \, .$$

The J_2 statistic, testing equality of the two correlation matrices, is then:

$$J_2 = \frac{1}{2} tr(Z^2) - [dg(Z)]' \, W^{-1} \, [dg(Z)]$$

where $dg(Z)$ stacks the main diagonal of Z into a column vector, and the elements w_{ij} of W are defined to be $w_{ij} = \delta_{ij} + c_{ij} \, c^{ij}$, in which

δ_{ij} denotes Kronecker's delta such that $\delta_{ij} = 1$ if $i = j$ and zero else;

c_{ij} is the ij-th element of C; and

c^{ij} is the ij-th element of C^{-1}.

J_2 is also asymptotically distributed as a chi-square with $\frac{N(N-1)}{2}$ degrees of freedom.

4.2. Size of the Box, the adjusted Box and the Jennrich test

Kryzanowski and To (1987) rightfully observe that the Box test assumes that the data are multivariate normal. They state that “if this assumption is not satisfied exactly, then the power of the test… is less than what would be expected theoretically.” Unfortunately, it is not only the power that is affected but also the size and hence the specification of the test as well!

Before implementing the proposed tests in our empirical framework, we first take a close look at the size of the Box test, the adjusted Box test and the Jennrich test. Keeping in mind that we want to apply the tests to our weekly credit spread data, we have set up simulation experiments with correlation dimensions of 3 (the three rating categories), 5 (the five maturity buckets) and 15 (the fifteen indices) and with short time series of 69 observations. Clearly, we would like to know whether the asymptotic properties of the tests still carry over to small samples. Table 6 documents the size of the tests under multivariate normality.

Table 6: Size of the Box test, the adjusted Box test and the Jennrich test under multivariate normality

		Dimension		
	Type I error	3	5	15
Box test	1%	0.0107	0.0088	0.0118
	5%	0.0554	0.0456	0.0513
	10%	0.1106	0.0967	0.0978
Adjusted Box test	1%	0.0011	0.0009	0.0008
	5%	0.0070	0.0060	0.0065
	10%	0.0173	0.0151	0.0180
Jennrich test on	1%	0.0069	0.0074	0.0051
Covariances	5%	0.0444	0.0413	0.0362
	10%	0.0959	0.0912	0.0819
Jennrich test on	1%	0.0118	0.0087	0.0083
Correlations	5%	0.0559	0.0523	0.0541
	10%	0.1069	0.1044	0.1191

The size of the Box test and the Jennrich test on correlations was determined based on 10,000 simulation runs where the correlation matrices of dimension 3, 5 and 15 were calculated based on 69 observations drawn from a multivariate standard normal distribution. The size of the adjusted Box test and the Jennrich test on covariances was determined based on a covariance matrix equal to 2 times the identity matrix. Each cell entry lists respectively the 1%, 5% and 10% rejection rate (first, second and third row).

Table 6 reveals that (even) for small samples the Box and Jennrich tests are relatively well specified. The adjusted Box test proposed by Tang (1995, 1995b), however, dramatically underestimates the size at all the standard significance levels. It does so, irrespective of the dimension of the correlation matrix. These findings immediately rule out the adjusted Box test for testing the equality of correlations.

Considering the excess kurtosis of the credit spread changes as shown in Table 5 Panel B, a fat tailed distribution would be more apt to model the stochastic behaviour of credit spreads. The literature on the distribution of speculative prices has spawned a plethora of candidate univariate distributions including sum-stable distributions (Mandelbrot, 1963), Mixtures of Normals (Kon, 1984), the Student-*t* distribution (Blattberg and Gonedes, 1975) and Mixed Jump diffusion processes (Ball and Torous, 1983). The main stylised fact all these authors want to capture is the fat-tailed character and the excess kurtosis of the empirical distributions of several kinds of speculative prices. In order to study the size under an alternative distribution, we need a multivariate alternative to the multivariate normal, which at least can capture the behaviour in the tails in an easy way. The multivariate Student-*t* distribution, *mvt*, fits this purpose. It can easily be generated based on a multivariate normally distributed random variable *x,* with a given mean vector and a given variance-covariance structure:

$$mvt = \frac{x}{\sqrt{\frac{y}{v}}}$$

where y denotes a Chi-Square distribution with v degrees of freedom (Tong, 1990). A nice property of the multivariate Student-*t* is that it – as its univariate counterpart – nests the multivariate normal distribution for $v \to \infty$. At the same time v indicates the tail thickness. The lower v, the fatter the tails of the distribution is. In order to study the impact of departures of multivariate normality, we repeat our simulations (10,000 runs, 69 observations, and correlation matrices of dimensions 3, 5 and 15), with the same covariance structures but we let the degrees of freedom vary from 50 to 3. The first case closely approximates multivariate normality, the last one possesses very fat tails. In order to preserve space, Table 7 only reports the results for v = 50, 30, 5 and 3.

Table 7: Size of the Box test, and the Jennrich tests under the multivariate Student-*t* distribution

	Box test			Jennrich test on covariances			Jennrich test on correlations		
	1%	5%	10%	1%	5%	10%	1%	5%	10%
v				Dimension = 3					
50	0.015	0.069	0.129	0.011	0.059	0.117	0.011	0.058	0.114
30	0.018	0.082	0.143	0.012	0.069	0.130	0.015	0.063	0.117
15	0.031	0.113	0.191	0.024	0.097	0.176	0.019	0.079	0.146
5	0.248	0.419	0.524	0.220	0.391	0.505	0.108	0.230	0.324
3	0.628	0.760	0.822	0.593	0.742	0.809	0.312	0.478	0.575
				Dimension = 5					
50	0.014	0.066	0.131	0.011	0.059	0.123	0.011	0.0603	0.120
30	0.018	0.076	0.143	0.013	0.068	0.134	0.012	0.0646	0.130
15	0.037	0.128	0.210	0.026	0.114	0.200	0.021	0.0931	0.170
5	0.390	0.585	0.687	0.349	0.564	0.676	0.205	0.3912	0.511
3	0.831	0.913	0.942	0.806	0.906	0.939	0.598	0.7610	0.830
				Dimension = 15					
50	0.021	0.087	0.160	0.010	0.064	0.133	0.011	0.062	0.132
30	0.034	0.130	0.221	0.015	0.095	0.189	0.017	0.095	0.175
15	0.118	0.290	0.416	0.071	0.235	0.373	0.059	0.207	0.339
5	0.892	0.960	0.978	0.838	0.945	0.973	0.740	0.895	0.942
3	0.999	0.999	1.000	0.997	0.999	0.999	0.991	0.998	0.999

The size of the Box test and the Jennrich test on correlations was determined based on 10,000 simulation runs where the correlation matrices of dimension 3, 5 and 15 were calculated based on 69 observations drawn from a multivariate Student-*t* distribution with 3, 5, 15, and 50 degrees of freedom. The size of the Jennrich test on covariances was determined based on a covariance matrix equal to 2 times the identity matrix. For every test, the cell entries list respectively the 1%, 5% and 10% rejection rate (first, second and third column).

Table 7 clearly shows that the size of both the Box and the Jennrich tests rapidly deteriorates, with diminishing degrees of freedom. Unfortunately, univariate estimates of the degrees of freedom for the Student-*t* distribution often lead to estimates in the neighbourhood of 3 to 5 (e.g. Blattberg and Gonedes (1974) for stocks, Boothe and Glassman (1987) for foreign exchange rates).

Table 7 also displays the fact that tests based on matrices with a larger dimension lead to even more misspecification than those performed on small correlation matrices. It is clear from these simulations that no serious conclusions can be drawn based on a standard application of the proposed tests under the presence of leptokurtosis and fat tails. It is hardly surprising that a lot of authors find rejections of the null of stability. Deviations from the normality assumption caused by fat tails indeed lead to extremely liberal tests and thus are hardly useful.

If we still want to perform these tests on financial data, we need an approximation of the small sample distribution of the test statistic under fat tailed alternatives. In order to do so, Zhang and Boos (1992) develop a bootstrap procedure to assess the significance of the Box test. They show that a bootstrap procedure can be used to estimate critical values for test statistics testing the equality of covariance, when the assumption of multivariate normality does not hold. Consistent with the findings of Zhang and Boos (1992), Table 7 indicates that asymptotically Chi-square distributed tests are rather sensitive to non-normality. We therefore follow their recommendation in using the bootstrap procedure for the data analysis.[4] The proposed bootstrap procedure starts by demeaning the two sub-sets of credit spreads. Subsequently, 500 samples of two sub-sets (each with 69 observations) were drawn with replacement

[4] Boos and Zhang (1992) show that the bootstrap procedure might seem to be a little bit conservative for larger dimensions. Unfortunately, assessing the size and power of bootstrap procedures leads to biased power and size estimates if done with a simple Monte Carlo simulation (Boos and Zhang, 2000). Further statistical research is needed to obtain 'absolute' certainty of the specification of the bootstrap-based statistical inference.

from the pooled demeaned data sets. In order to obtain the bootstrap results for e.g. the Box test, we calculate the Box statistic for each of the 500 samples/couples. Apart from the 500 bootstrapped Box test values, the Box statistic is also calculated for the original data set. The bootstrapped *p*-value reports the percentage of bootstrapped Box statistics that are higher than the Box statistic obtained from the original credit spread changes data. Bootstrapped *p*-values for the Jennrich tests on covariances and correlations were calculated analogously.

5. Inter-temporal stability of credit spread co-movements

A comparison of the stability results using the Box, Jennrich test and the Bootstrap procedure is presented in Table 8. We report the results for the weekly data (5 days) and each correlation or covariance matrix thus is estimated based on 69 observations. It has often been noticed that non-synchronous trading and other microstructure effects can produce artificially low correlations (e.g. Best, 1999). Moreover, since our data are quoted prices, data errors might still occur on a daily basis. It is therefore advisable to study the data on a lower frequency e.g. on a weekly basis. We show the results of the overall data set as well as for 2 different types of sub-samples. First the indices are grouped according to their ratings. These subsets thereby contain five 'asset categories', each originating form a different maturity bucket. Alternatively, the sample is regrouped based on the time to maturity. Within each of these maturity buckets three indices of different investment grade qualities are considered.

Table 8: Equality of Covariance and Correlation Matrices

	Covariance		Correlation
	Box	**Jennrich**	**Jennrich**
Overall	485.06	297.6	191.63
	(0.0000)*	(0.0000)*	(0.0000)*
	[0.0000]*	[0.0000]*	[0.0120]*
AAA	39.07	36.23	22.09
	(0.0011)*	(0.0016)*	(0.0147)*
	[0.0700]	[0.0680]	[0.0860]
AA	62.84	49.51	20.04
	(0.0000)*	(0.0000)*	(0.0289)*
	[0.0000]*	[0.0000]*	[0.0580]
A	166.89	100.08	22.94
	(0.0000)*	(0.0000)*	(0.0110)*
	[0.0000]*	[0.0000]*	[0.1580]
1-3	149.87	77.61	10.44
	(0.0000)*	(0.0000)*	(0.0152)*
	[0.0000]*	[0.0000]*	[0.0140]*
3-5	8.13	8.06	5.66
	(0.2288)	(0.2336)	(0.1296)
	[0.5840]	[0.5620]	[0.3460]
5-7	39.38	31.93	4.56
	(0.0000)*	(0.0000)*	(0.2074)
	[0.0380]*	[0.0440]*	[0.3560]
7-10	70.85	52.3	24.93
	(0.0000)*	(0.0000)*	(0.0000)*
	[0.0060]*	[0.0040]*	[0.0020]*
10+	15.94	15.34	2.62
	(0.0141)*	(0.0178)*	(0.4532)
	[0.0980]	[0.0940]	[0.4520]

Calculations of the Box and Jennrich statistics are based on a total sample of 138 weekly observations. Each sub-period (2 populations) consist of 69 observations. Weekly spread changes are constructed by adding 5 daily spread changes. An asterisk denotes significantly different covariance or correlation matrices at the 5% significance level. *p*-values of the standard statistics are given between parentheses, the bootstrapped *p*-values are placed between squared brackets.

Table 8 reports the test results on the stability of the covariances and correlations of the credit spread changes based on the standard (i.e. asymptotically) and bootstrapped Box and Jennrich tests. We observe that the results on covariance stability are the same irrespective of the test used (compare the Box (column2) and Jennrich (column 3) *p*-values).

The stability of the covariance structure computed from the 15 rating and maturity based indices, is firmly rejected. If we study the indices conditioned on rating/maturity (leaving us within the rating/maturity bucket with 5/3 sub-indices differing with respect to maturity/rating) this result is generally confirmed. At a 5% significance level, the bootstrapped *p*-values and their asymptotic counterparts, lead to the same conclusions. Only two exceptions can be spotted: the AAA bucket and the 10+ maturity bucket. For these buckets, the bootstrap procedure does not reject the null of stability.

Since the covariances are based on both variances and correlations, the instability found can be due to changing correlations. We now turn to the analysis of the correlation matrices to find out whether the correlations are more stable than the covariances. The standard and the bootstrapped Jennrich test performed on the correlation matrices based on the 15 indices both reject the null hypothesis at a 5% significance level. However, at the 1% level, the bootstrapped *p*-value no longer rejects the null. This evidence shows that correlations in general are more stable than covariances. Turning to the sub-samples, one can observe that the bootstrapped *p*-values convincingly confirm this. For only two of the sub-samples, the bootstrapped tests reject the equality of correlation matrices at the 5% significance level. Finally, it is clear from the rating buckets in Table 8 that the asymptotic statistical inference can lead to false conclusions. Taking these results into account, we can induce that the covariances are in general less stable than correlations.

6. Conclusion

Corporate bonds expose the investor to credit risk, which will be reflected in the credit spread. Based on the EMU Broad Market indices, we study the inter-temporal stability of the covariance and correlation matrices of credit spread changes. Before implementing stability tests on weekly data, we give full details on the construction of the credit spreads used. Descriptive statistics of the credit spreads are provided as well. Given the multivariate nature of the data set, we use the Box test (on the equality of covariance matrices) and the Jennrich test (on the equality of covariance or correlation matrices) to test the stability of the co-movement structure of credit spreads. First, we show that procedure proposed by Tang (1995a, 1995b) to test equality of correlations (using a Box-test based on standardised data), leads to a misspecified test under the null. Second, the appropriateness of the standard Box and Jennrich tests is checked while relaxing the assumption of multivariate normality. This analysis leads us to conclude that these departures also lead to extremely misspecified tests. We therefore are forced to apply a bootstrap procedure to obtain correct statistical inference. Using the standard (asymptotical) Box and Jennrich test statistics, neither the covariance nor the correlation matrix is stable over time. The bootstrap procedure, however, shows that, for the period under study, this conclusion is falsified. Correlations turn out to be less instable as often claimed.

References

Annaert, J., M. J. K. De Ceuster, F. De Jonghe (2000) "Modelling European credit spreads", *Research report*, Antwerp, Deloitte & Touche.

Ball, C. A. and W. N. Torous (1983) "A simplified jump process for common stock returns", *Journal of Financial and Quantitative Analysis*, **18** (1), 53-65.

Best, P. W. (1998) *Implementing Value at Risk*, New York, Wiley Series in Financial Engineering.

Blattberg, R. C. and N. J. Gonedes (1974) "A comparison of the stable and student distributions as statistical models for stock prices", *Journal of Business*, **47** (2), 244-280.

Boos, D. D. and J. Zhang, 2000, "Monte Carlo evaluation of resampling-based hypothesis tests", *Journal of the American Statistical Association*, **June**, **90** (450), 486-492.

Boothe, P. and D. Glassman, (1987) "The statistical distribution of exchange rates: Empirical evidence and economic implications", *Journal of International Economics*, 297-319.

Box, G. E. P. (1949) "A general distribution theory for a class of likelihood criteria", *Biometrika*, **36**, 317-346.

Cheung, Y. L. and Y. K. Ho (1991) "The inter-temporal stability of the relationships between the Asian emerging equity markets and the developed equity markets", *Journal of Business Finance and Accounting*, **18** (2), 235-254.

Cooley, W. W. and P. R. Lohnes (1971) *Multivariate data analysis*, New York, Wiley.

Duffee, G. R. (1998) “The relation between Treasury yields and corporate bond yield spreads,” *Journal of Finance,* **53** (6), 2225-2242.

Fridson, M. S., and J. G. Jónsson (1995) “Spread versus Treasuries and the riskiness of high-yield bonds,” *Journal of Fixed Income*, 79-88.

Gibbons, M. R. (1981) “Empirical examination of the return generating process of the arbitrage pricing theory”, *Working paper*, Stanford University, Stanford, CA.

Haney, Jr. R. L. and W. P. Lloyd (1978) “An examinations of the stability of the inter-temporal relationships among national stock market indices”, *Nebraska Journal of Economics and Business*, **17** (2), 55-65.

Jennrich, R. I. (1970) “An asymptotic χ^2 test for the equality of two correlation matrices”, *Journal of the American Statistical Association*, **65**, 904-912.

Kiesel, R., W. Perraudin, and A. Taylor (2001) “The structure of credit risk: spread volatility and ratings transitions”, *Working Paper*, Bank of England, **131**.

Kaplanis, E. C. (1988) “Stability and forecasting of the co-movement measures of international stock market returns”, *Journal of International Money and Finance*, **7**, 63-75.

Kon, S. J. (1984) “Models of stock returns: a comparison”, *Journal of Finance*, **39** (1), March, 147-165.

Kryzanowski, L. and M. C. To (1987) “The E-V stationarity of secure returns. Some empirical evidence”, *Journal of Banking and Finance,* **11**, 117-135.

Longstaff, F. A., and E. S. Schwartz (1995) “A simple approach to valuing risky fixed and floating rate debt,” *Journal of Finance,* **50** (3), 789-819.

Makridakis, S. G. and S. C. Weelwright (1974) "An analysis of interrelationships among the major world stock exchanges", *Journal of Business Finance and Accounting*, **1** (2), 195-215.

Maldonado, R. and A. Saunders (1981) "International portfolio diversification and the inter-temporal stability of international stock market relationships: 1957-1978", *Financial Management*, 54-63.

Mandelbrot, B. B. (1963) "The variation of certain speculative prices", *Journal of Business*, 394-419.

Meric, I. and G. Meric (1989) "Potential gains from international portfolio diversification and inter-temporal stability and seasonality in international stock market relationships", *Journal of Banking and Finance*, **13** (4-5), 627-640.

Morrison, D. F. (1976) *Multivariate statistical methods*, second ed., Toronto, McGraw-Hill.

Pedrosa, M. and R. Roll (1998) "Systematic risk in corporate bond credit spreads," *Journal of Fixed Income*, 7-26.

Philippatos, G. C., A. Christofi and P. Christofi (1983) "The inter-temporal stability of international stock market relationships: another view", *Financial Management*, **12** (4), 63-69.

Tang, G. (1995) "Inter-temporal stability in international stock market relationships: a revisit", *Quarterly Review of Economics and Finance*, **35**, 579-593.

Tang, G. (1995b) "Stability of international stock market relationships across month of year and holding intervals", *European Journal of Finance*, **1** (3), 207-218.

Tong, Y. L. (1990) *The Multivariate Normal Distribution*, Berlin: Springer-Verlag.

Watson, J. (1980) "The stationarity of inter-country correlation coefficients: a note", *Journal of Business Finance and Accounting*, **7** (2), 297-303.

Zhang, J. and D. D. Boos (1992) "Bootstrap Critical Values for Testing Homogeneity of Covariance Matrices", *Journal of the American Statistical Association*, June, **87** (418), 425-429.