Main Comments of Referee 1

1. On page 4 it is said that the data period is 2007-2012, but Figure 2 graphs data starting 2004. If one is to test non-crisis dependence and crisis dependence, I guess an even much longer period prior to 2007 is needed. Furthermore, if one wants to test the difference between the euro crisis and credit crisis episodes, this is another objective, as both were crisis period.

Indeed, that part of the sample was excluded from the analysis. We did it for two reasons:

- at the inception of the index, it lacked liquidity. Before the end-2007, the bid-ask spread associated to the CDS index was very high as compared with the rest of the sample. Between mid-2004 and 03/12/2007, the average bid-ask spread was near 12.0%, whereas in the remainder of the span it was below 1.75%. The percentage of trading days with no CDS spread changes was also very high during the first span, which is consistent with the presence of large transaction costs. Taken together, these results suggest that short-term frictions in the CDS markets were very high prior to 2008, which could lead to biased results
- further, as Figure 1 suggests, there is a clear structural break in the conditional mean and variance of CDS index returns in the beginning of the financial turmoil.

For all the above mentioned reasons, we excluded the span prior to 03/12/2007 from the subsequent analysis.

2. On page 6 residuals are estimated assuming a GARCH process. Subsequently, the dependency is investigated on these residuals. This poses the question whether those data are really proper (since these are estimated). In fact it has been shown that if a GARCH process with normal innovations is estimated, the residuals are still fat tailed due to estimation errors. For the same reason, these errors may still carry dependency that was to be removed by the GARCH filter. This possible bias has to be addressed.

It is clear that the use of a two-stage estimation procedure may bring about estimation error. First, the correct specification for the conditional mean and variance of the returns may not be that specific ARMA-GARCH process employed in the analysis. Concerning that issue, we tested several alternative specifications in terms of the number of lags used. We also tested alternative GARCH models as the asymmetric GARCH model. Based on the AIC/BIC, we chose the ARMA-GARCH filtering process presented in the paper. Second, even when the specification is correct, the parameters are estimated with error due to the finite-sample bias. As far as we know, there is no simple way to deal with that problem.

We also performed two additional tests to ascertain whether the filtered returns/innovations still carried dependency (BDS and runs test). The hypothesis of randomness was not rejected.

3. In middle of page 6 it is reported that the univariate distributions are 'estimated' nonparametrically. Then the question is why one would estimate the (tail) copula parametrically, see the bottom page 6? A parametric copula is biased by what happens in the center of distribution (due to the wealth of observations in the center) and imposes a mold that may not fit. So if one is interested in tail dependency, why not proceed nonparametrically? This is potentially even easier, see the book by De Haan and Ferreira (2006) or Embrechts, Frey and

With the aim of checking the robustness of the results, we will provide the results obtained when using parametric marginal distributions, namely the Gaussian and student's t univariate distributions (in a revised version). K-S tests reject the hypothesis that the filtered returns follow a normal distribution, but they do not reject the hypothesis that they follow a student's t distribution.

Overall, the results do not change significantly when one of the alternative parametric distributions is used to model the marginal distribution of the copula instead of the Empirical Distribution Function.

4. Section 3 starts with measuring the quantile dependence. But nowhere is this concept defined, while this is crucial for the entire paper. Figures 2 and 3 are hard to interpret for this reason.

In the new version of the paper, we will introduce a formal definition for quantile dependence.

5. It is quite suspect that left and right tail confidence bands in Figure 2 are so widely different. Without knowing how these are constructed, we cannot understand why this is the case?

Idem.

6. Page 9: The time variation in the rank correlation is certainly of interest. But is the i.i.d. bootstrap to test the symmetry of both sides of the sample on bottom of page 9 correct under this null?

The idea is that the filtered returns/innovations are random. Notice that we employ several metrics to gauge the presence of autocorrelation and ARCH effects in the residuals, including the Ljung-Box-Pierce test, the Breusch-Godfrey test, the runs test and the BDS test.

There were no signs of the presence of dependence in the residuals. This test was also implemented in other papers as in Patton (2012)[1].

7. Page 2, Table 5: The time varying normal copula somewhat automatically adapts to fat tails, so this is why the unconditional Student-t distribution does best. Why not let the Student-t be time varying to test for time variation. Since time variation may erroneously confuse for fat tails; see also page 12 bottom.

In the new version of the paper we will add the time-varying Student's t copula to the analysis. Moreover, we change the frequency of the data from daily to weekly data (and extend the span to May 2014) to remove spurious time-varying dependence.

---

[1] Patton, A. J. (2012). Copula Methods for Forecasting Multivariate Time Series. Handbook of Economic Forecasting (2).

Extracted from the new version of the paper:

« Although daily information is available for the two indexes, we conduct the analysis using weekly data to make our results immune to microstructure noise that stems from the bid-ask bounce and non-synchronicity between the two indexes. In addition, information is assimilated at different paces by stock prices and spreads within each index, causing autocorrelation of the indexes returns and affecting spuriously their conditional means and variances. »

Under this new data frequency, the AIC/BIC measures "confirm" that Student's t copula outperforms other copula functions (static and dynamic) in describing and fitting the data.

8. Bottom page 11: It is totally unclear what is meant by "the dependency structure is not concentrated in one of the tails but rather in the centre"?

Modified in the new version.

Small remarks

1. *Page 4: Entries (firms) in both indices did change during the credit and euro crises. How important is this effect? Was it possible to replicate these investments without making considerable extra losses due to entry and exit?*

Indeed, the rebalancing of the indices poses several questions to empirical researchers. However, both indices are well diversified and capture the systematic trend of the two markets. In that sense, we consider that the rebalancing effect would have little impact in the results.