

Handling risk-on/risk-off dynamics with correlation regimes and correlation networks

Jochen Papenbrock · Peter Schwendner

© Swiss Society for Financial Market Research 2015

Abstract In this paper, we present a framework for detecting distinct correlation regimes and analyzing the emerging state dependences for a multi-asset futures portfolio from 1998 to 2013. These correlation regimes have been significantly different since the financial crisis of 2008 than they were previously; cluster tracking shows that asset classes are now less separated. We identify distinct “risk-on” and “risk-off” assets with the help of correlation networks. In addition to visualizing, we quantify these observations using suitable metrics for the clusters and correlation networks. The framework will be useful for financial risk management, portfolio construction, and asset allocation.

Keywords Regime switching · Correlation regimes · Clustering · Correlation networks · Risk management · Portfolio construction · Asset allocation

JEL Classification C14 · G11 · G01 · D85

J. Papenbrock
PPI AG, Frankfurt, Germany

J. Papenbrock (✉)
Firamis UG, Oberursel, Germany
e-mail: jp@firamis.de

P. Schwendner
Center for Alternative Investments and Risk Management,
Zurich University of Applied Sciences, Winterthur, Switzerland
e-mail: peter.schwendner@zhaw.ch

1 Introduction

Since the global financial crisis, the most frequently used phrase for describing asset price behavior has been “correlations jump to one”. This is due to a new market environment that oscillates between two situations—“risk on” and “risk off” (RoRo)—from one trading day to another. When risk appetite is high, commodities, equities, and high-yielding currencies tend to rally together. When risk appetite fades, those assets fall and there is a flight to safe-haven assets. This is in contrast to pre-crisis dynamics when high correlations were only associated with market contractions (see [Fengler and Schwendner 2004](#)) and high volatilities.

As a result, between mid-2007 and the beginning of 2013, most markets and asset classes were highly synchronized, despite the stock market recovery and bond rally that occurred between March 2009 and March 2013, which was accompanied by decreasing volatilities.

There was a long-term positive drift of equities and bond prices despite their negative short-term return correlation.

This bipolar attitude toward risk makes conventional investment processes difficult:

1. Due to the extreme market coordination, there is less potential for diversification so diversification fails where it is most needed. For example, after the Lehman default in September 2008, not only equities, but also commodities, credit, and most hedge funds substantially declined in value or even became illiquid.
2. Since asset classes “have lost their identity” ([Flood 2012](#)), fundamental asset class evaluations are dominated by correlated movements triggered by macro events ([Lee 2012](#)). Also, it is hard to add value since there is little differentiation within asset classes.
3. Parameter estimation of risk and expected return is an even greater challenge.
4. Former relative price relations become ineffective. Factor sensitivities of individual stocks relevant for quantitative investment strategies become more and more similar, therefore decreasing the value of these models ([Lee 2012](#)).
5. Strategies that involve relative changes between assets are less effective due to the new market synchronization and the stress in the money market ([Khandani and Lo 2007](#)). On the other hand, increased correlations offer new opportunities for pair trading ([Stopford 2012](#)).
6. Fundamental forecasting is nearly impossible as macro events occur suddenly and have different regional input.

It is even more difficult to persistently succeed with investment strategies, and there will be a wider dispersion of potential portfolio outcomes and a higher degree of portfolio concentration.

2 How to deal with RoRo?

Using a simple correlation matrix approach to identify RoRo dynamics has drawbacks. First, a correlation heat map’s visual outcome can vary based on the ordering of markets. Second, the short time window used to compute the correlation matrix introduces

statistical noise in the results (Laloux et al. 1999; Plerou et al. 2002; Fengler et al. 2007).

An indirect method of dealing with RoRo is to estimate its driving factors. According to Lee (2012), the most recent risk-based portfolio construction approaches estimate the factor structure, time the factor returns, and dynamically change portfolio exposures. Taking this approach to its extreme, the only skill that should be rewarded when correlations go to one is factor timing. However, there are some difficulties with this method.

- The breadth of decisions is relatively limited and entirely relies on the quality of factor information.
- It is not clear what the right number of factors is or what to do in time of crisis when the number of factors may be even fewer. Also, if the risk factor approach is fully executed it converges to the asset class approach. However, asset class-specific behavior was observed less since the Lehman default.
- The right factor timing is an open question.

Finding the risk factors statistically has the advantage that it is not necessary to make any assumptions about which variables correspond to which factors and factors are directly defined from the return time series. In another statistical approach, HSBC (2010) assigns the first principal component of the correlation matrix to the RoRo factor. The authors use the correlation between an individual asset and the first principal component to classify this asset:

1. If the correlation between the asset returns and the RoRo factor is high and positive, we have a typical risk-on asset.
2. If the correlation between the asset returns and the RoRo factor is high and negative, we have a typical risk-off asset.
3. If there is only low correlation, it is a RoRo-neutral asset.

This analysis can be dynamic due to rolling correlation windows. Also, it uses the relevant parts of a noisy correlation matrix. Finally, a breakdown of the usual relations can be identified, indicating the end of the RoRo paradigm.

In the next section, we introduce cluster-based identification of different correlation regimes. Conditional on each regime, there is a typical conditional correlation structure and we also identify which time period was characterized by which regime. From the regimes and their timing we can deduce early warning signals, design a correlation stress test, conduct distinct correlation scenario analysis, properly calibrate portfolio optimization programs and risk models, and obtain an overview of correlation structures. Also, we can identify a typical RoRo regime and use network technologies to find the most typical risk-on or risk-off assets. This information, together with the correlation regimes, can be the basis for trading strategies, asset allocation, portfolio construction, and risk management strategies.

The network-based representation also provides insight into “the most similar market/asset” to a specific one, which is interesting for proxy hedge purposes. Finally,

we analyze the cluster structure of each regime and track the merging, splitting, birth, and death of the clusters in time. This could be the basis for pairs-like reversal trading strategies but with pairs corresponding to clusters.

3 Correlation clusters and correlation networks

We use techniques that can detect changes in the correlation structure of several markets in very short periods, such as a few trading days. The resulting correlation structure switches are the basis for finding several distinct correlation regimes. Similar to [HSBC \(2010\)](#), we analyze 25 different markets covering four different asset classes: government bond futures, equity index futures, commodity futures, and foreign exchange futures (see [Table 1](#)) over the period 01/07/1998 to 31/01/2013, which includes several financial crises and recoveries. In contrast to HSBC, we rely purely on time series of rolled futures contracts instead of a mix of futures and cash instruments. This makes the return computations across asset classes more comparable and the correlation computations more reliable. The influence of different currency denominations on futures returns is only of second order, as these instruments do not involve a funding component.

We identify correlation regimes with the following six steps:

1. Split the data set of daily returns into 175 non-overlapping subsets each containing the daily returns for each of the 175 corresponding months. Each of the 175 matrices defines a fundamental asset return correlation time slice of the markets.
2. Compute the 25×25 asset return correlation matrices for each of the 175 months.
3. Transform the 175 correlation matrices into 175 correlation-based distance matrices.
4. Filter each correlation distance matrix with a hierarchical clustering technique because of the matrices' high noise, which is due to the relatively short estimation windows ([Laloux et al. 1999](#)).
5. Compute the pair-wise similarity between the filtered correlation distance matrices (similarity is measured with respect to their correlation structure). The result is a 175×175 matrix with pair-wise similarities of all filtered correlation distance matrices.
6. Find discrete state clusters ("regimes"), where each cluster contains very similar monthly time slices, in terms of their correlation structure. The state clusters are expected to exhibit high discriminatory power. We apply the well-known flat k -means clustering with $k = 5$ clusters but also evaluate the quality of an alternative clustering based on a hierarchy that is cut at a certain point to obtain the $k = 5$ flat clusters. The choice of five clusters is mainly driven by the desire to have a low number of clusters for the sake of simplicity, on the one hand, and, on the other, to have economically meaningful clusters for such a long time period that also discriminate between different market phases in the time period before the financial crisis. For an analysis that simply describes just the most prominent aspects of the data, two clusters would be sufficient.

Table 1 Futures markets used to compute the correlation matrices

| Bloomberg ticker | Long name | Asset class | Color |
|------------------|---------------|---------------------|--------|
| TY1 Comdty | 10Y US T-Note | Gov bond future | Orange |
| G 1 Comdty | 10Y Gilt | Gov bond future | Orange |
| JB1 Comdty | 10Y JGB | Gov bond future | Orange |
| RX1 Comdty | 10Y Bund | Gov bond future | Orange |
| XM1 Comdty | 10Y Australia | Gov bond future | Orange |
| CN1 Comdty | 10Y Canada | Gov bond future | Orange |
| SP1 Index | S&P500 | Equity index future | Green |
| Z 1 Index | FTSE | Equity index future | Green |
| NK1 Index | Nikkei | Equity index future | Green |
| VG1 Index | EuroSTOXX50 | Equity index future | Green |
| GX1 Index | DAX | Equity index future | Green |
| HI1 Index | Hang Seng | Equity index future | Green |
| BZ1 Index | Bovespa | Equity index future | Green |
| HG1 Comdty | Copper | Commodity future | Yellow |
| GC1 Comdty | Gold | Commodity future | Yellow |
| CL1 Comdty | WTI Oil | Commodity future | Yellow |
| NG1 Comdty | Natural Gas | Commodity future | Yellow |
| S1 Comdty | Soybeans | Commodity future | Yellow |
| W1 Comdty | Wheat | Commodity future | Yellow |
| BP1 Curncy | GBP | FX future | Blue |
| JY1 Curncy | JPY | FX future | Blue |
| EC1 Curncy | EUR | FX future | Blue |
| AD1 Curncy | AUD | FX future | Blue |
| CD1 Curncy | CAD | FX future | Blue |
| RA1 Curncy | ZAR | FX future | Blue |

We realize that it is possible to estimate the proper number of clusters by some cluster quality criterion. However, there is a whole range of these criteria and each focuses on a different aspect of the clustering. These could be evaluated by some sort of aggregate ranking of the clustering procedures based on a number of cluster quality criteria. In this analysis, however, we determine the number of clusters based on financial market experience, which resembles domain knowledge in the machine-learning context.

Results for the alternative clustering based on a hierarchy are given in Appendix 2. In the following, we detect our states based on the k -means analysis. We use this method because both ways of clustering deliver relatively similar results with respect to a structural break related to the beginning of the financial crisis in 2007. The hierarchical clustering, however, is less noise sensitive and focuses more on outliers whereas the k -means flat clustering suggests clusters with evenly distributed sizes, which is closer to the idea of finding distinct regimes/states of the market.

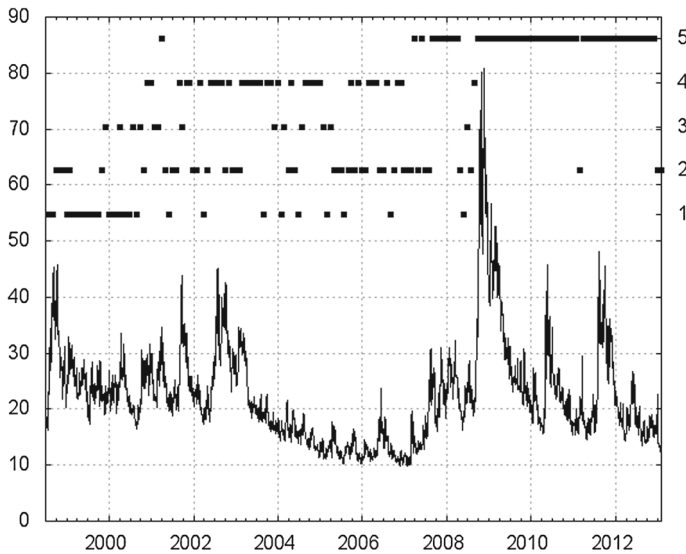


Fig. 1 Correlation regimes and VIX in time

4 Correlation regimes in time

Figure 1 shows the occurrence of the $k = 5$ clusters in time, laid over the VIX index as 175 black squares, one for each month. The VIX index measures the implied volatility of S&P 500 index options and is often viewed as a “fear index”.

Our procedure for detecting states of financial markets is similar to that of [Münix et al. \(2012\)](#), who use the original noisy state correlations and then use a state distance/dissimilarity measure that strongly suppresses noise. However, under our approach, we first filter the monthly asset return correlations by extracting a hierarchical asset cluster structure and then use a state distance/dissimilarity measure that matches the hierarchical data structure.

For Step 3, we choose a widespread correlation-based distance. This is the necessary input for the filtering based on hierarchical asset clustering. The correlation-based metric is widely used in financial applications by practitioners of clustering (see, e.g., [Lisi and Corazza 2008](#); [Mantegna 1999](#); [Tola et al. 2008](#); [Tumminello et al. 2010](#); [Dose and Cincotti 2005](#)). The Pearson correlation coefficient often is used as a measure of strength of linear dependence between two variables:

$$\rho_{ij} = \frac{\langle r_i(t, \Delta t) r_j(t, \Delta t) \rangle_T - \langle r_i(t, \Delta t) \rangle_T \langle r_j(t, \Delta t) \rangle_T}{\sqrt{\langle r_i^2(t, \Delta t) \rangle_T - \langle r_i(t, \Delta t) \rangle_T^2} \sqrt{\langle r_j^2(t, \Delta t) \rangle_T - \langle r_j(t, \Delta t) \rangle_T^2}},$$

where $i, j = 1, \dots, n$ denote the markets, T is the time horizon, and $r_i(t, \Delta t)$ is the return of market i in the subperiod $[t, t + \Delta t]$ of T .

However, the correlation coefficient of a pair of asset returns cannot be used as a distance because it does not fulfill the axioms that form a metric. A real metric can be

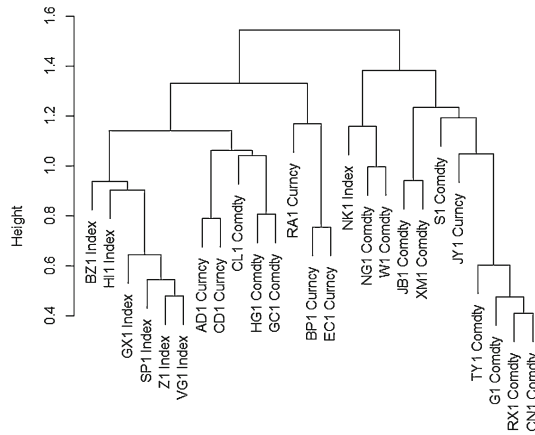


Fig. 2 Dendrogram for 1 month

designed using a function of the correlation coefficient ρ . It can be rigorously determined by a transformation of the correlation coefficient so that the distance between variables decreases if correlation between them increases (Gower 1966):

$$d(i, j) = \sqrt{2(1 - \rho_{ij})}$$

This distance fulfills the usual metric properties, including the triangle relation (see, e.g., Mantegna 1999).

In Step 4, the correlation distance matrices are input for a hierarchical clustering called average linkage (see Tola et al. 2008). Tola et al. (2008) compare the following four procedures in terms of portfolio risk and weight concentration:

1. Markowitz basic estimation,
2. Random matrix theory,
3. Single linkage clustering, and
4. Average linkage clustering.

For their specific data set, Tola et al. (2008) find that average linkage clustering is a successful filtering method with respect to level of portfolio risk and the concentration of weights across a portfolio. Thus, we also use the average linkage clustering procedure to filter the monthly return correlation distance matrices. The average linkage clustering for the correlation distance matrix in one of the monthly time slices has the dendrogrammatic representation shown in Fig. 2.

Two hierarchical asset clusterings of two different time slices can be compared to each other in terms of their dendrogrammatic similarity. One such measure is the cophenetic distance. The cophenetic distance between two markets that have been clustered is defined to be the intergroup dissimilarity at which the two markets are first combined into a single cluster (height on the y-axis of the dendrogram). In comparison to the correlation distance matrix, which has $n(n-1)/2$ distinct entries, the cophenetic distance matrix has only $n-1$ distinct entries.

A dendrogram is an appropriate summary of some data if the correlation between the original distances and the cophenetic distances is high. The cophenetic correlation coefficient is a measure of how faithfully a dendrogram preserves the original correlation distance matrix. It is defined by:

$$c_{st} = \frac{\sum_{i < j} (s_{ij} - \bar{s})(t_{ij} - \bar{t})}{\sqrt{\left[\sum_{i < j} (s_{ij} - \bar{s})^2 \right] \left[\sum_{i < j} (t_{ij} - \bar{t})^2 \right]}}$$

where s_{ij} is a matrix entry of the original correlation distance matrix, t_{ij} is an entry of the cophenetic distance matrix, \bar{s} is the average of the s_{ij} , and \bar{t} is the average of the t_{ij} . Averaged over the 175 filtered monthly correlation distance matrices, the cophenetic correlation between the original correlation distances and the hierarchical cluster representation is 0.84, meaning that the hierarchical cluster structure is a good approximation of the original data.

In Step 5, we need to compute a similarity matrix between the 175 dendrograms. As a similarity measure, we also use the cophenetic correlation coefficient between all months (s, t), each from 1 to 175.

The representation of correlation regimes shown in Fig. 1 has many advantages.

1. If new data arrives, it can be immediately determined whether their correlation structure is similar to any of existent regimes or whether they constitute a completely new regime.
2. Regimes can be found that occur only in stressful periods, where, in our case, “stress” is defined as a high VIX index. These regimes could be called “correlation at risk” or “correlation at stress” and they could be suitable for stress testing and for a parameterization of diversification models. For example, following [Fabozzi and Focardi \(2010\)](#), these regimes could be used to diversify trends/drifts using those correlation structures typical for crisis scenarios in well-known models like Markowitz portfolio construction. The latter approaches are known for their vanishing diversification properties in times of crises. So a parameterization of Markowitz approaches with conditional correlations like “correlation at stress” could improve modeling accuracy. Also, the use of conditional correlations could re-justify the normal distribution assumption that has to hold when employing Markowitz.
3. Any correlation-based model that has to be parameterized with historical data can benefit from the regimes as the model input would not be most recent data (e.g., data from the last 250 trading days) but the most similar data. In simple terms, model input would be the returns data from a past regime that is the same as the present regime.
4. Another application is of a forecasting nature with respect to future dependence structures. For a certain correlation regime in one step, there are probabilities for transition to another (or the same) regime in the next time step.

Looking at correlation regime evolution in time, it can be observed that from July 1998 until the run on Northern Rock in September 2007, Regimes 1–4 dominate. In mid-2007, as liquidity problems in the money markets arose, Regime 5 occurred

for the first time. After the Fed-organized takeover of Bear Stearns by JP Morgan in March 2008, markets traded in an inflationary scenario and Regime 5 disappeared for a few months. The Lehman fall marks a major reconstruction of correlations: after that event, Regimes 1, 3, and 4 almost disappear, replaced by Regimes 2 and 5. Regime 5 seems to occur at almost all crisis events after Lehman. This is the “risk-on, risk-off” regime. In 2010 (Eurozone crisis) and 2011 (downgrade of the USA by S&P), the VIX experienced two strong spikes, but this did not lead to the emergence of another correlation regime. It would be interesting to discover the typical correlation structure of the five regimes. Therefore, we average the monthly correlation matrices belonging to each regime and plot the five typical regime correlation matrices as heat maps¹ (Fig. 3).

White cells in the heat map stand for high correlation and black cells are anti-correlation. The individual assets are sorted into asset classes (from left to right or bottom to top: government bond futures, equity index futures, commodity futures, and FX futures). For Regimes 1–4, most asset classes build their own cluster of high intra-correlation. The range of gray in the heat map implies that there are many separate market forces driving different assets in a nontrivial way. This range of forces leads to many different behaviors and, consequently, to a large number of uncorrelated assets before the crisis. Regimes 2 and 4 show especially negative correlations between equities and bonds. The heat map of Regime 5 is very extreme in terms of RoRo coupling, as it merely shows two blocks: bonds and JPY as “risk-off” assets against all other assets which form a “risk-on” block. However, looking at several statistics and, later, at dynamic cluster analysis, it is obvious that there are also huge differences between Regimes 1–4.

5 Statistical characterization

Table 2 compares the five regimes in terms of mean correlations, mean absolute correlations, mean daily returns, mean absolute daily returns, mean realized volatilities, and mean absolute Sharpe ratios.

The “mean” averages shown in Table 2 are always across the 25 futures markets. The most striking result is the strong separation of mean absolute correlations due to the regimes. The strong coupling in Regime 5 leads to an average absolute correlation $>30\%$. The same value is $<20\%$ in all other regimes. For the mean absolute daily returns, the separation due to the regimes is much stronger than it is for the mean daily returns: the mean absolute return in Regime 5 is clearly higher than that of all others. The full ANOVA tables in Appendix 1 emphasize this finding with a failed F test for the mean daily return time series and a passed F test for the mean absolute daily return time series. The reason could be the negative correlations between assets that blur the effect of the regimes on the mean returns across the assets. Also, volatilities are higher in Regime 5 than in the other regimes. To account for the effect that the volatility of commodity and equity futures is much larger than that of FX futures and

¹ “‘Risk On–Risk Off’—How a Paradigm is born”, *Currency Weekly*, HSBC Global Research, August 2, 2010.

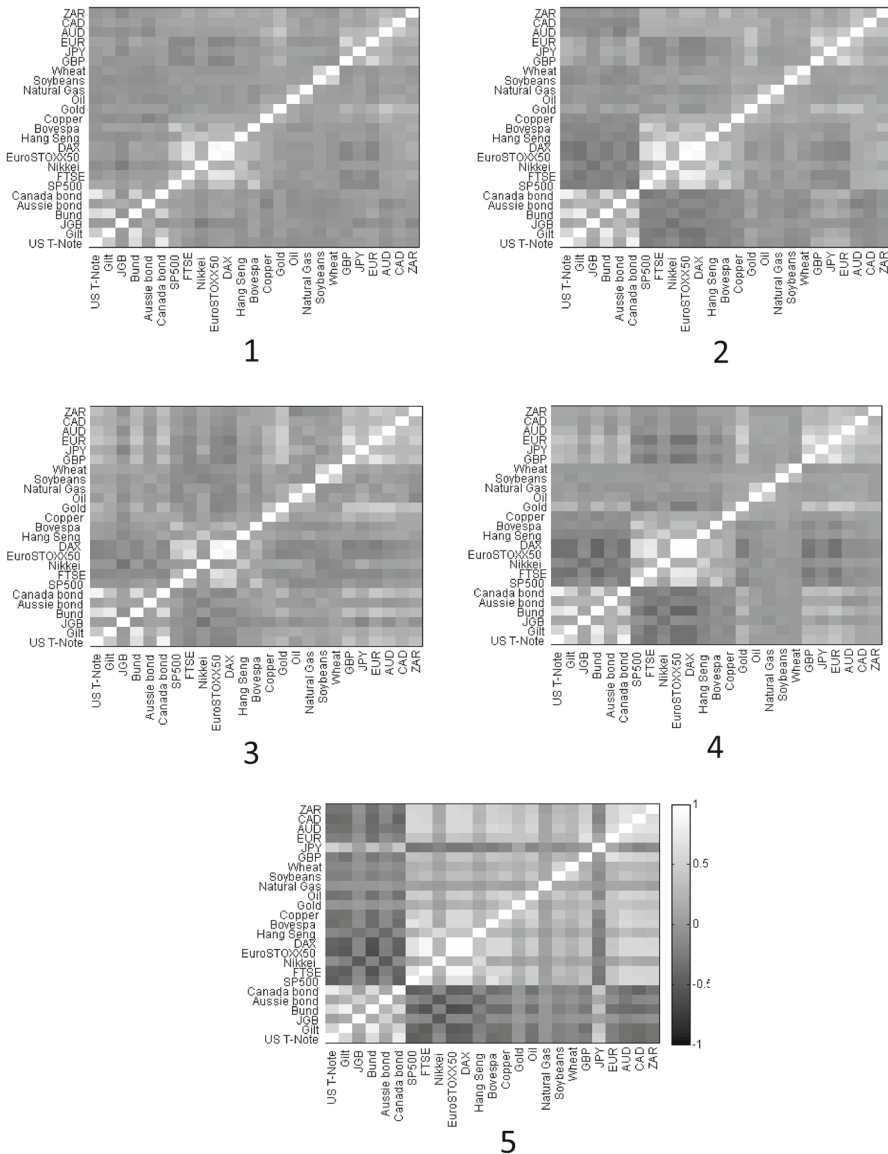
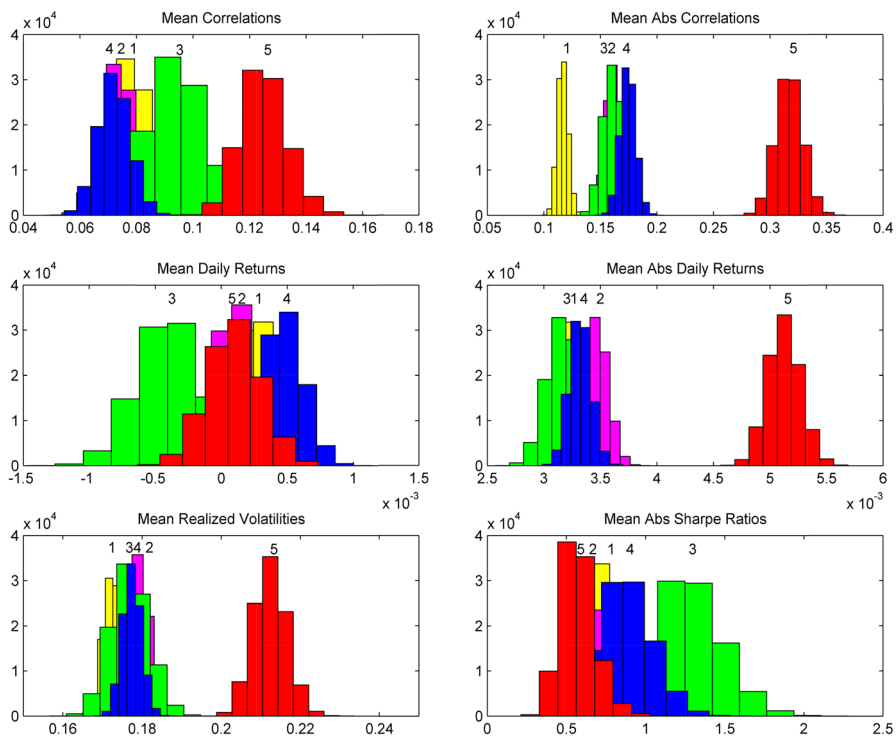


Fig. 3 Heat maps of the five regimes; each heat map represents the average correlation matrix of all unfiltered time slices belonging to this regime

government bond futures, we also compute the Sharpe ratios for each market and average across markets. The mean absolute Sharpe ratio of Regime 3 is the highest, as the negative mean returns of Regime 3 were accompanied by low realized volatility, whereas the returns of Regime 5 showed high absolute levels and high absolute levels of correlations, but less mean correlations at high mean volatilities, so the assets hedged against each other in Regime 5 and did not have a steady trend, as reflected by the low

Table 2 Average returns and correlations within the five regime correlation matrices

| | Regime 1 | Regime 2 | Regime 3 | Regime 4 | Regime 5 | F | F crit | P value |
|----------------------------|-----------|-----------|------------|-----------|-----------|-------|----------|-----------|
| Mean correlations | 0.0779 | 0.0739 | 0.0931 | 0.072 | 0.1251 | | | |
| Mean Abs correlations | 0.1092 | 0.1577 | 0.149 | 0.1706 | 0.3165 | | | |
| Mean daily returns | 0.0002600 | 0.0001190 | −0.0004040 | 0.0004700 | 0.0000945 | 1.36 | 2.37 | 0.2445 |
| Mean Abs daily returns | 0.0032 | 0.0035 | 0.0032 | 0.0033 | 0.0051 | 25.94 | 2.37 | 3.11E−21 |
| Mean realized volatilities | 0.1728 | 0.1792 | 0.1777 | 0.1776 | 0.2125 | | | |
| Mean Abs Sharpe ratio | 0.5314 | 0.4631 | 1.0828 | 0.783 | 0.4268 | | | |


Fig. 4 Bootstrap densities in the five regimes

Sharpe ratio. This is consistent with the interpretation of Regime 5 as the RoRo regime. As we do not have different time series of volatilities, correlations, and Sharpe ratios in each regime, but only one realization, it is not possible to present an ANOVA table for these observables. Therefore, in Fig. 4 we illustrate our interpretation with results of a bootstrap procedure for all observables of Table 2. We pull 100,000 draws from the historical return time series of the respective regimes with replacement and show the histograms. The bootstrap approach was pioneered by Efron (1979). A financial

application of mapping bootstrapped correlation matrices to model price densities is presented by [Fengler and Schwendner \(2004\)](#). The idea of using bootstrapping to validate results of correlation networks is presented by [Tumminello et al. \(2007\)](#). [Dütsch \(2014\)](#) applies this method to time series of statistical moments so as to validate regime separation. Our bootstrap results confirm the failed F test for the mean daily returns and the successful F test for the mean absolute daily returns. Comparing the regime separation between the mean correlation densities and the mean absolute correlation densities gives a result similar to comparing the regime separation between the mean daily return densities and the mean absolute daily return densities: Regime 5 is completely separated from the rest when absolute values of returns or absolute values of mean correlations are considered, but less so when the signed observables are examined. Comparing the densities of mean correlation with those of mean daily returns and mean realized volatilities in Regimes 1–4 confirms that “high correlation means high volatilities and negative returns”, which was a common pattern before the crisis. The strong correlations in Regime 5 obviously were formed in a period of broad return distributions and high volatilities, but low trends, as confirmed by the low mean and narrowness of the Sharpe ratio density in this regime. A market practitioner might characterize Regimes 1 and 4 as “bullish”, Regime 3 as “bearish”, Regime 2 as an “uncorrelated sideways market”, and Regime 5 as “correlated risk on, risk off”.

To this point, we have described the behavior of our multi-asset futures universe in the five regimes from a market perspective via their mean returns, mean volatilities, and mean correlations. Note that it is also possible to interpret the regimes from an economic perspective based on mean returns, mean volatilities, and mean correlations, even though the regimes were determined by comparing full correlation matrices.

In the next sections, we employ methods from cluster and network analysis.

6 Correlations between regimes

We use the same method to compare dendrograms with cophenetic correlations as before but this time just for the five averaged regime correlation matrices. This allows us to discover how different the five regimes are in terms of their cophenetic correlation structure (Fig. 5).

We observe two blocks, the first between Regimes 1 and 2 and the second between Regimes 3 and 4, which supports the visual observation that Regimes 1 and 2 are similar and Regimes 3 and 4 are similar, but that Regime 5 is distinct from each of the other four. However, the degree of difference between the regimes is now quantified in a differentiated way. This method emphasizes the difference between Regime 5 and the block comprised of Regimes 3 and 4.

7 Cluster tracking

We now look at how the markets group in each regime, how the grouping structure changes, whether there are groups that always exist throughout the different regimes, and whether certain markets/assets play functional roles across the whole system.

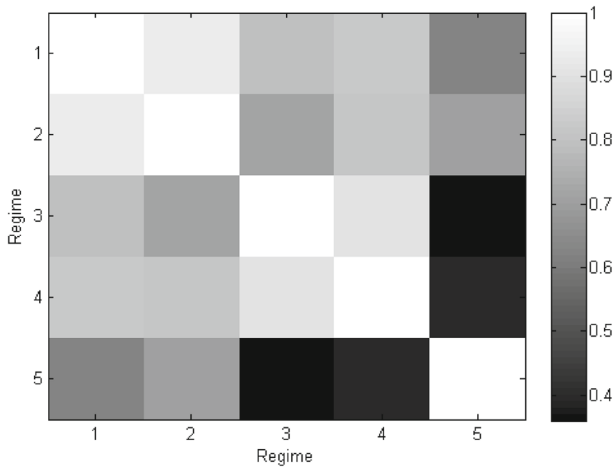


Fig. 5 Heat map of cophenetic correlations between regimes

We mine for distinct clusters in each regime based on the k -means clustering. Also, we use an algorithm for tracking the clusters in each regime to discover the following evolving cluster structures:

1. Birth of a cluster never before observed,
2. Death of an existing cluster,
3. Merging of one or more clusters,
4. Splitting of a cluster,
5. Contraction of a cluster, and
6. Growth of a cluster.

We use a procedure similar to that taken by [Fenn et al. \(2012\)](#) and focus on a market-centric cluster analysis. [Fenn et al. \(2012\)](#) mention that many approaches require defining core markets of clusters, which seems overly restrictive. Thus, instead of tracking whole clusters, we identify clusters from the perspective of individual nodes, which allows us to study the temporal dynamics of functional market roles.

A market's identity is known in all five regimes; moreover, its cluster membership in each regime is known. Thus, we can track the cluster evolution from the perspective of individual markets. We investigate the persistence through time of a market's cluster by defining a cluster autocorrelation. For market i with cluster $c_i(s)$ in regimes, and cluster $c_i(t)$ in regime t the autocorrelation is defined by:

$$a_i^{st} = \frac{c_i(s) \cap c_i(t)}{c_i(s) \cup c_i(t)}$$

Using the k -means clustering algorithm (2:15 clusters, where the clustering quality measure is average silhouette width), we split each regime correlation matrix into discrete clusters. We use the market-centric cluster tracking version, which does not require us to determine which cluster in each regime represents the descendant of a cluster in another regime. Figure 5 tracks the clusters and labels each cluster with the

Table 3 Adjusted Rand index for the five regimes

| Regime | 1 | 2 | 3 | 4 | 5 |
|--|------|------|------|------|------|
| Adjusted Rand index (reference partition is asset class) | 0.37 | 0.35 | 0.46 | 0.32 | 0.28 |

number of market constituents. The five regimes are on the x -axis; the cluster number is on the y -axis. A square signifies the occurrence of a specific cluster in a certain regime. The numbers in the squares are the number of markets contained in a specific cluster in a specific regime. No cluster is contained in all regimes. Clusters 1–5 and 7–9 only occur in Regimes 1–3. Clusters 10 and 11 only occur in Regime 4 and Clusters 12 and 13 occur only in Regime 5. The clusters in Regimes 4 and 5 are much bigger than the clusters in Regimes 1–3.

It was expected that after Lehman, there would be strong RoRo dominance and, indeed, this is seen in Regime 5. Remarkably, even before the crisis, there is a bipolar Regime 4. However, the two clusters in Regimes 4 and 5 are very different from each other, which is illustrated by the relatively low cophenetic correlation between these two regimes (see Fig. 5) and by the cluster tracking, which indicates that all clusters are different from each other.

As will be shown by the network analysis, natural gas plays a special role and populates a network community of its own, which indicates that it has nothing to do with RoRo dynamics. Table 3 shows the degree of accordance in asset classes and intrinsic clusters.

The adjusted Rand index for Regime 3 is 0.46, which means that its discrete clustering is quite similar to the partition of asset classes. As expected, bipolar Regime 5 has relatively low accordance with the asset classes.

This analysis makes it clear how to get rid of the RoRo influence. One way is to engage in a diversification program similar to Markowitz, parameterized with the correlation matrix of a crisis regime. This should increase diversification if the market switches to this regime in the future. Another way to avoid RoRo is to conduct a finer-grained clustering of time slices in even more clusters and find the regime in which RoRo occurs the least. The correlation structure of this regime will reveal how assets classes are truly correlated without the RoRo influence.

Implementing the RoRo dynamics in active management, however, requires an additional analysis that needs to find those specific markets during a period of RoRo. Using both the correlation regimes as well as correlation networks, we could set up a process to detect the timing and the markets relevant for actively trading in the RoRo regime. In the following section we provide some background on correlation networks.

8 Network analysis

There are many types of networks in finance: threshold networks, causal networks, influence networks, partial networks, bipartite networks, and many others. A wide range of financial assets and markets have been investigated using networks, including equities, currencies, commodities, bonds, and interest rates.

The class of topological filter networks dates back to Mantegna (1999). In a system with n markets/assets, the approach selects $n - 1$ correlations from the $n(n - 1)/2$ distinct entries of a correlation matrix to construct a spanning tree, a network with no cycles. When constructing the spanning tree, only the highest (and therefore highly significant) possible correlations are chosen. The result is called a “minimal spanning tree” (MST) as the highest correlations translate into minimal distances (a correlation of 1 translates to a distance of 0) from which the tree is constructed. Such network approaches consider the interaction of several markets/assets as a complex system. They’re composed of many interacting elements and can exhibit numerous forms of “emergent” collective dynamics without the need for any external organizing principle. Such dynamics typically cannot be explained by studying the constituent parts in isolation, so a complex system must be analyzed as a whole. Networks reflected by graphs provide a tractable framework and controlled environment for the quantitative analysis of many complex systems by distilling them down to their key dependence structure. These networks focus on the most important “backbone” dependence structure and reduce complexity, smoothing the way for more sophisticated analyses. The elements of the system are represented as the network’s nodes and the important interactions between them are links. Indeed, networks naturally reflect the variety of elements in the system by their set of vertices, and the plurality of the interrelations between elements by their edges. Networks can be analyzed by network statistics on a global, local, or dynamic scale. These relations and changes can be visualized so that subsequent analyses can focus on the relevant structures. For example, central or decoupled assets are easily identified.

We generate² MSTs for the five correlation regimes (Fig. 6) and color them based on the asset class codes defined in Table 1.

To compare the MSTs of the five regimes, we use the edge survival measure between regime s and t :

$$ES(s, t) = \frac{1}{N - 1} \left| E^s \cap E^t \right|,$$

where E^s refers to the set of edges of the MST of regime s , E^t refers to the set of edges of the MST of regime t , and the \cap operation intersects the two MSTs. So $ES(s, t)$ is simply the fraction of surviving edges to all edges. Figure 7 shows all combinations of survival ratios for the five regimes.

Like in Fig. 5, in Fig. 7, Regimes 1 and 2 are similar. But, in contrast to the cophenetic correlation, the tree survival rates emphasize a larger similarity between Regimes 4 and 5 than between Regimes 3 and 4. Note, however, that Fig. 7 is based on a different analysis than Fig. 5. The latter was produced looking at the asset hierarchies within the five regimes. Figure 7 was produced by looking at the asset networks of each regime. The obvious thing the figures have in common is their database: they use the same asset correlations per regime as input. The similarity between Regimes 4 and 5 displayed in Fig. 7 reminds us of the similarity in the number of clusters per regime

² We used the software FNA (<http://www.fna.fi>) for the network figures.

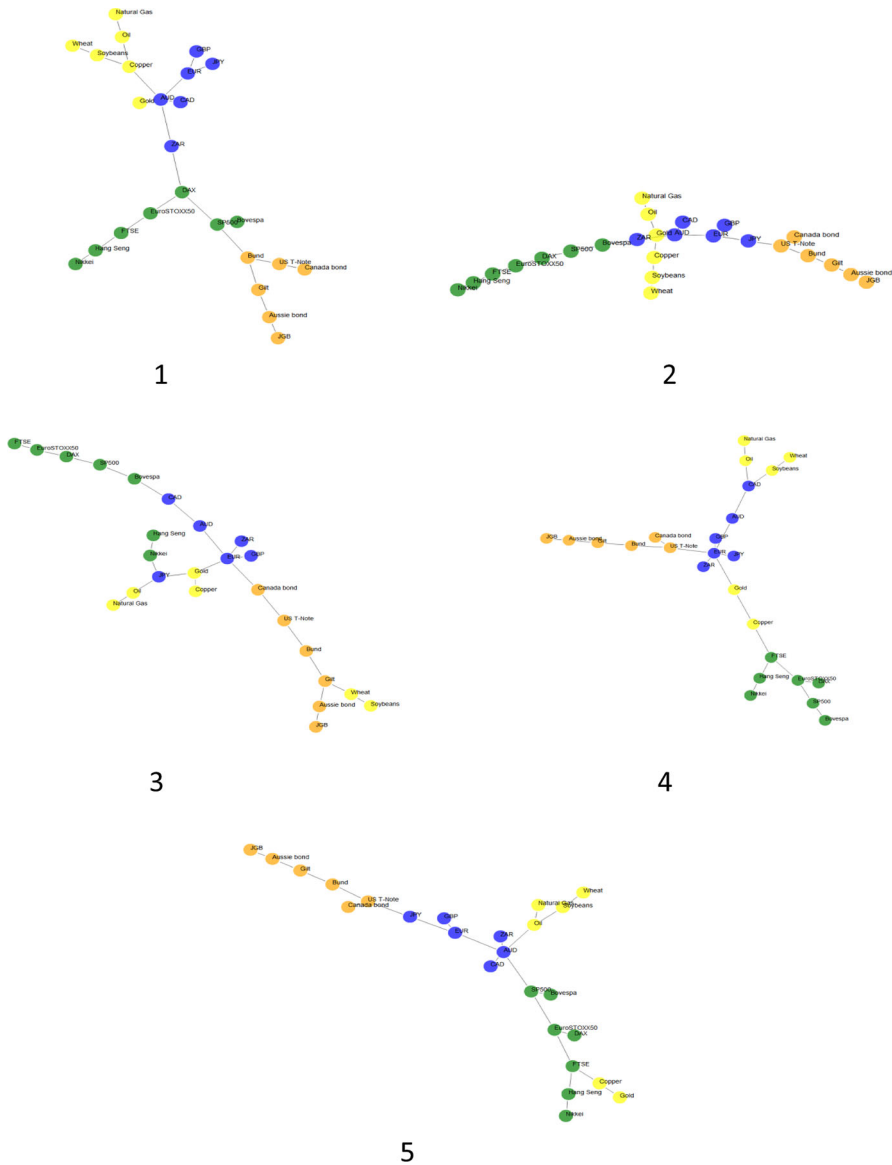


Fig. 6 Minimum spanning trees (MST) generated from the five regimes

discussed in relation to Fig. 8. From an inspection of the MSTs in Fig. 6, we discover that the gold and copper markets are sorted between the FX futures and equity index futures for Regime 4, but are behind the equity index markets for Regime 5. Apart from this shift in hierarchies, the two networks appear very similar. This fits with the higher similarity between Regimes 4 and 5 found by the tree survival rate measure than that found from the cophenetic correlation that compares the hierarchies.

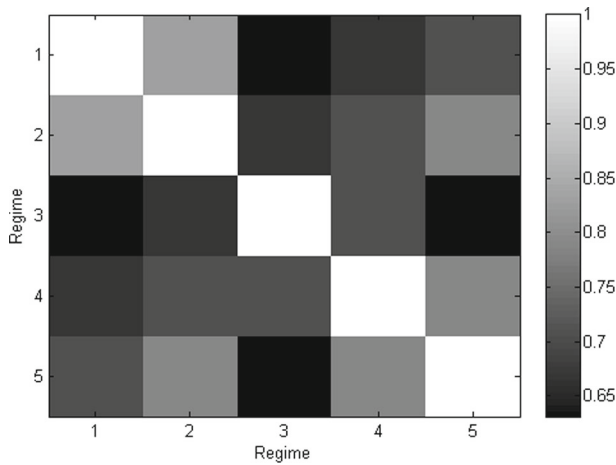


Fig. 7 Tree survival rates between the regimes

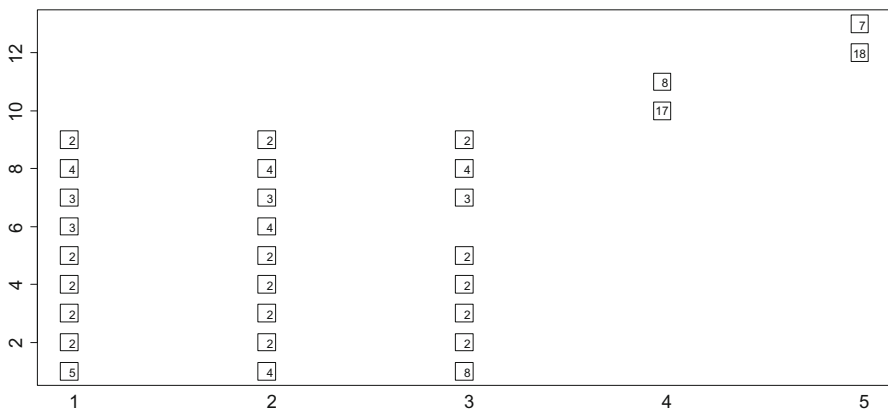


Fig. 8 Cluster tracking through regimes

In the next step, we focus solely on crisis Regime 5. We intend to discover the weakest links in the MST. To do so, we introduce an increasing threshold for the absolute correlations and remove the links in the MST below this threshold. At a threshold level for the absolute correlations of 0.24, the MST starts to break into two trees. This is a different analysis than simply a flat clustering of Regime 5 as done before.

The results of the cutoff (shown in Fig. 9) are two branches with numerous nodes and one single node: natural gas. One of the branches is mostly populated with assets commonly known as risk on and the other branch is characterized by typical risk-off assets. The centroid of the risk-on assets is AUD; the centroid of the risk-off assets is the US T-Note future.

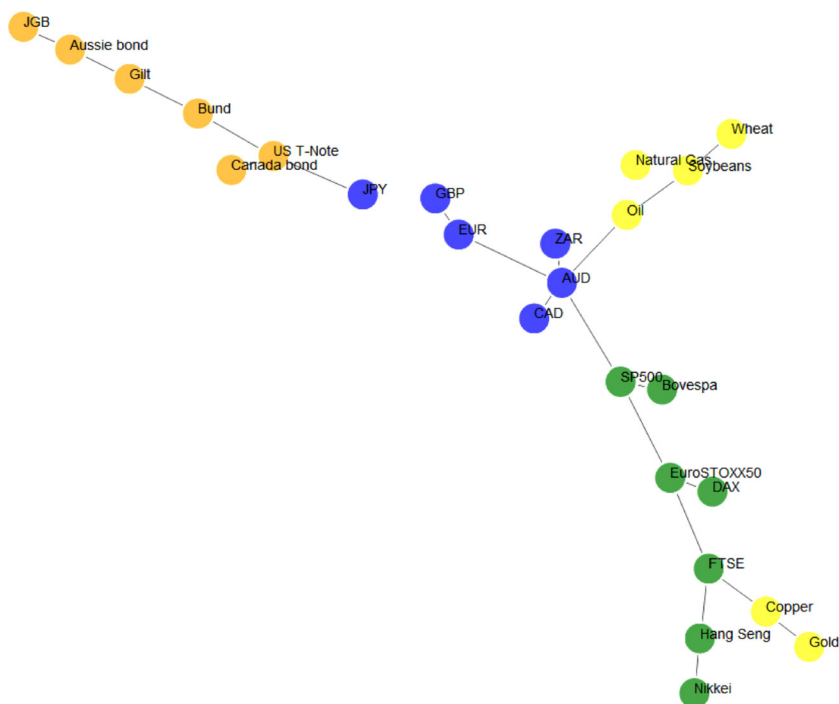


Fig. 9 Breakup for the MST of Regime 5

9 Conclusion

We took the following steps to analyze the “risk-on, risk-off” dynamics of a multi-asset futures portfolio from July 1998 to January 2013.

1. We partitioned the time axis into equidistant time slices. For each time slice, we calculated the correlation matrix and filtered it using average linkage clustering so as to reduce noise.
2. We computed similarity matrices from the filtered correlation matrices and performed k -means clustering to detect correlation regimes. With $k = 5$, we found four distinct regimes (1–4) that tended to appear before the Lehman default, and one regime (5) dominating afterward. Regime 5 is a specific crisis regime that occurs only “post-subprime” (mid-2007). An alternative hierarchical clustering technique also revealed this crisis regime.
3. For each of the five regimes, we calculated the average correlation matrix. The “post-subprime” regime correlation matrix for Regime 5 shows markedly higher absolute correlations than the others.
4. We derived clusters from each of the five regime correlation matrices using k -means and quantified the accordance of the intrinsic clusters with asset classes. We found far less asset class separation in the post-subprime period.

5. We computed the return time series for the five regimes and characterized the univariate and multivariate behavior of the 25 markets in these five regimes using ANOVA tables and a bootstrap approach.
6. A network analysis using minimum spanning trees (MST) for the five regime correlation matrices was quantified using edge survival ratios. We found similarities between Regimes 1 and 2.
7. For Regime 5, we introduced a cutoff level for the absolute correlations and increased it until the MST broke into two trees. One of the two trees was distinctly characterized by risk-on assets, the other by risk-off assets.

The proposed steps not only confirm the observation of higher correlations in the post-subprime period, but also allow assigning distinct correlation regimes to specific time slices and following the time evolution of the regime classification. The correlation network of the mean correlation matrix within a specific regime shows the tightest couplings between markets. The clustering and network techniques for time slices and assets thus do not only result in an appealing visualization of correlation matrices, but also offer a quantitative framework for analyzing the asset dynamics that can be used in risk management, trading, and portfolio management. The presented framework is a macro-to-micro connection from the regime classification at the top (macro view) to the characterization of individual markets in terms of location within a network or a cluster (micro view). The dynamics of networks from one time slice or regime to the next is an interesting field for further studies.

Acknowledgments We thank Frank Fabozzi and an anonymous referee for valuable comments that significantly improved the quality of this article.

Appendix 1

See Appendix Table 4 and Fig. 10.

Table 4 Analysis-of-variance tables for the mean daily returns and mean absolute daily returns in the k -means regimes with $k = 5$

| SUMMARY for mean daily returns | | | | | | |
|---|--------|-----------|-----------|----------|----------------|-------------------|
| Group | Number | Sum | Mean | Variance | | |
| Regime 1 | 571 | 0.1487 | 2.60E−04 | 1.72E−05 | | |
| Regime 2 | 809 | 0.0963 | 1.19E−04 | 2.09E−05 | | |
| Regime 3 | 263 | −0.1063 | −4.04E−04 | 1.52E−05 | | |
| Regime 3 | 735 | 0.3454 | 4.70E−04 | 1.76E−05 | | |
| Regime 4 | 1,287 | 0.1216 | 9.45E−05 | 5.33E−05 | | |
| ANOVA for mean daily returns | | | | | | |
| Sources of variance | SS | <i>df</i> | MS | <i>F</i> | <i>P</i> value | <i>F</i> critical |
| Between the regimes | 0.0002 | 4 | 4.17E−05 | 1.362 | 0.2445 | 2.374 |
| Within the regimes | 0.1121 | 3,660 | 3.06E−05 | | | |
| Sum | 0.1122 | 3,664 | | | | |
| SUMMARY for mean absolute daily returns | | | | | | |
| Group | Number | Sum | Mean | Variance | | |
| Regime 1 | 571 | 4.5909 | 8.04E−03 | 5.86E−06 | | |
| Regime 2 | 809 | 6.5925 | 8.15E−03 | 1.00E−05 | | |
| Regime 3 | 263 | 2.1289 | 8.09E−03 | 7.60E−06 | | |
| Regime 3 | 735 | 6.0889 | 8.28E−03 | 7.70E−06 | | |
| Regime 4 | 1,287 | 12.2443 | 9.51E−03 | 2.78E−05 | | |
| ANOVA for mean absolute daily returns | | | | | | |
| Sources of variance | SS | <i>df</i> | MS | <i>F</i> | <i>P</i> value | <i>F</i> critical |
| Between the regimes | 0.0016 | 4 | 3.89E−04 | 25.94 | 3.11E−21 | 2.37 |
| Within the regimes | 0.0548 | 3,660 | 1.50E−05 | | | |
| Sum | 0.0564 | 3,664 | | | | |

For the mean daily returns, the F test fails at a 0.05 significance level. For the mean absolute daily returns, the F test is met

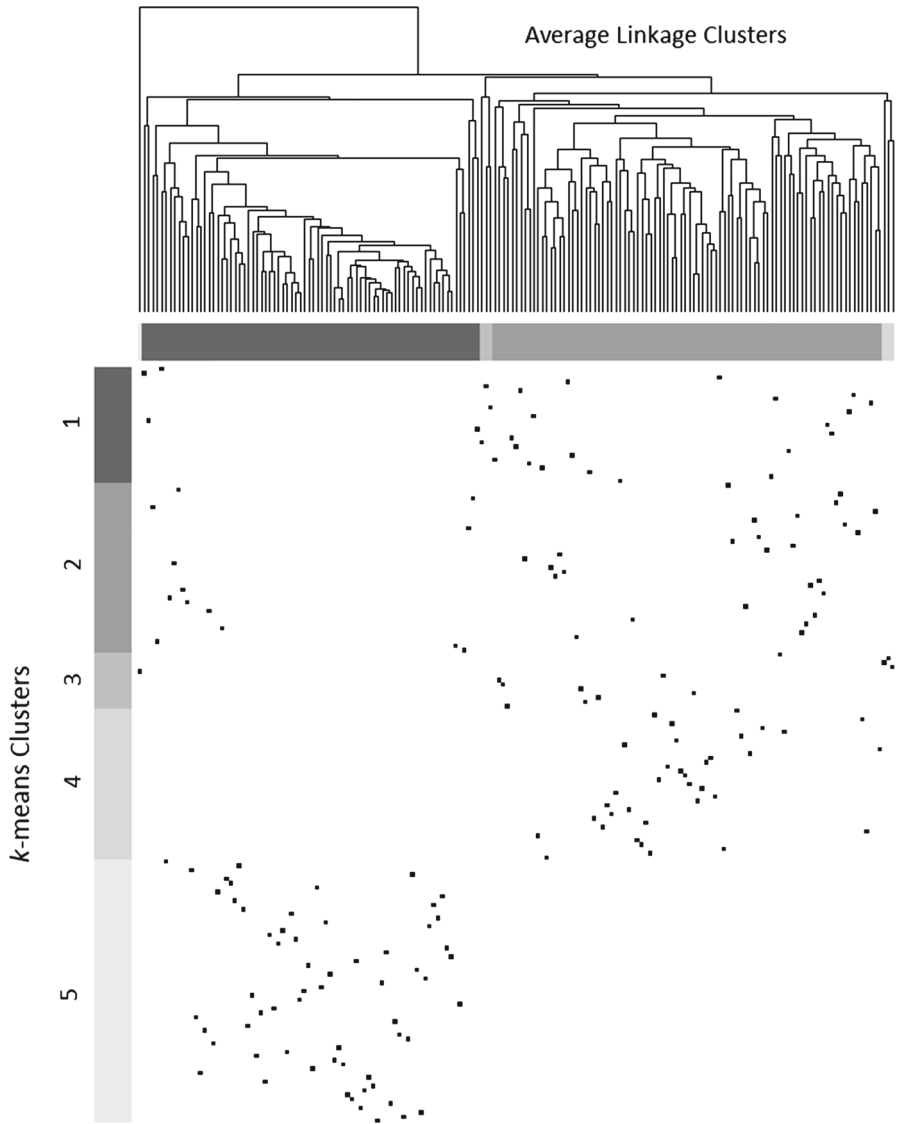


Fig. 10 Assignment of the 175 monthly correlation matrices to the clusters found by k -means and to the clusters found by the average linkage procedure. Both clusterings use the parameter $k = 5$

Appendix 2

Detecting the market states is a prerequisite for all subsequent analyses, which is why we not only employed the k -means as a clustering method to find the market states, but also used a representative of hierarchical clustering: the average linkage procedure. Flat clusters can be deduced, for example, by cutting the dendrogram at a certain level so as to obtain the desired number of clusters. The hierarchical superstructure between

the clusters and the hierarchical substructure within the clusters is ignored when flat clusters are defined by the dendrogram cut. In the following figure, we plotted the amount of overlap between the two clusterings— k -means and average linkage—with $k = 5$ flat clusters in both procedures. This was done by starting with an ordered heat map of the 1:175 time slices (months) where only the diagonal has black dots and all off-diagonal elements are white. We then reordered the rows according to the k -means clustering and the columns according to the average linkage procedure. The vertical gray-scale bars on the left-hand side of the figure correspond to the $k = 5$ clusters from the k -means clustering. At the top of the figure, the dendrogram shows the hierarchical clustering, also leading to five flat clusters denoted by the horizontal gray-scale bars. We expect to see rectangular clouds of dots in the rearranged heat map where clusters from both clustering procedures share the same time slice constituents.

The two flat clusterings correspond in finding the “risk on, risk off” regime (Regime 5 from k -means) as can be observed at the point cloud in the bottom-left corner of the heat map. The overlap of the other clusters is less structured.

References

- Dose, C., Cincotti, S.: Application to index and enhanced-index tracking portfolio. *Physica A* **355**(1), 145–151 (2005)
- Dütsch, P.: Anwendung von Machine Learning Techniken im Quantitativen Trading. Unpublished Master's thesis, Zurich University of Applied Research (2014)
- Efron, B.: Bootstrap methods: another look at the Jackknife. *Ann. Stat.* **7**, 1–26 (1979)
- Fabozzi, F.J., Focardi, S.M.: Diversification: should we be diversifying trends? *J. Portf. Manag.* **36**(4), 1–4 (2010)
- Fengler, M., Schwendner, P., Pilz, K.F.: Basket volatility and correlation. In: *Volatility as an Asset Class*, pp. 95–131. Incisive Media, London (2007)
- Fengler, M.R., Schwendner, P.: Quoting multiasset equity options in the presence of errors from estimating correlations. *J. Deriv.* **11**(4), 43–54 (2004)
- Fenn, D., Porter, M.A., Mucha, P.J., McDonald, M., Williams, S., Johnson, N.F., Jones, N.S.: Dynamical clustering of exchange rates. *Quant. Finance* **12**(10), 1493–1520 (2012)
- Flood, C.: Investors ignore flak over volatility ETPs. *Financial Times* April 29 (2012)
- Gower, J.D.: Some distance properties of latent root and vector methods used in multivariate analysis. *Biometrika* **53**, 325–338 (1966)
- HSBC: Risk on–risk off: the full story. *Currency Quant Special*, HSBC Global Research, November (2010)
- Khandani, A.E., Lo, A.W.: What happened to the quants in August 2007? *J. Invest. Manag.* **5**(4), 5–54 (2007)
- Laloux, L., Cizeau, P., Potters, M., Bouchaud, J.-P.: Random matrix theory. *Risk* **12**(3), 69 (1999)
- Lee, W.: Risk on/risk off. *J. Portf. Manag.* **38**(3), 28–39 (2012)
- Lisi, F., Corazza, M.: Clustering financial data for mutual fund management. In: Perna, C., Sibillo, M. (eds.) *Mathematical and statistical methods in insurance and finance*. Springer, Milan (2008)
- Mantegna, R.N.: Hierarchical structure in financial markets. *Euro. Phys. J. B* **11**, 193–197 (1999)
- Münix, M.C., Shimada, T., Schäfer, R., Leyvraz, F., Seligman, T.H., Guhr, T., Stanley, H.E.: Identifying states of a financial market. *Nat. Sci. Rep.* **2**, 644 (2012)
- Plerou, V., Gopikrishnan, P., Rosenow, B., Amaral, N., Guhr, T., Stanley, H.E.: Random matrix approach to cross correlations in financial data. *Phys. Rev. E* **65**, 1–18 (2002)
- Stopford, J.: Diversification in a correlated world. *Investec* (2012)
- Tola, V., Lillo, F., Gallegati, M., Mantegna, R.N.: Cluster analysis for portfolio optimization. *J. Econ. Dyn. Control* **32**(1), 235–258 (2008)
- Tumminello, M., Coronello, C., Lillo, F., Micciche, S., Mantegna, R.N.: Spanning trees and bootstrap reliability estimation in correlation based networks. *Int. J. Bifurc. Chaos* **17**(7), 2319–2329 (2007)
- Tumminello, M., Lillo, F.: Correlation, hierarchies, and networks in financial markets. *J. Econ. Behav. Organ.* **102**(30), 10421–10426 (2010)

Jochen Papenbrock is a managing consultant at PPI AG and managing director of Firamis UG. He earned a Ph.D. in Business Engineering at Karlsruhe Institute of Technology (KIT). His research interests and areas of financial product development are financial networks, tail risk modeling, and investment management.

Peter Schwendner is a senior lecturer at Zurich University of Applied Research (ZHAW), Center for Alternative Investments and Risk Management. He earned a Ph.D. in Physics at the University of Göttingen. His research interests are financial markets, investment products, and quantitative trading. Before joining ZHAW, Peter worked in derivatives research and systematic asset management.