Analysis of Correlation Based Networks Representing DAX 30 Stock Price Returns

Jenna Birch[1], Athanasios A Pantelous[2] and Kimmo Soramäki[3]

Abstract

In this paper, we consider three methods for filtering pertinent information from a series of complex networks modelling the correlations between stock price returns of the DAX 30 stocks for the time period 2001-2012 using the Thomson Reuters Datastream database and also the FNA platform to create the visualizations of the correlation-based networks. These methods reduce the complete 30x30 correlation coefficient matrix to a simpler network structure consisting only of the most relevant edges. The chosen network structures include the Minimum Spanning Tree (MST), Asset Graph (AG) and the Planar Maximally Filtered Graph (PMFG). The resulting networks and the extracted information are analysed and compared, looking at the clusters, cliques and connectivity. Finally, we consider some specific time periods (a) a period of crisis (Oct 2008 – Dec 2008) and (b) a period of recovery (May 2010 – Aug 2010) where we discuss the possible underlying economic reasoning for some aspects of the network structures produced. Overall, we find that network based representations of correlations within a broad market index are useful in providing insights about the growth dynamics of an economy.

JEL Classification: D85; G01; C82; C88
Keywords: MST; PMFG; AG; DAX 30; correlation networks; European sovereign-debt crisis.

---

[1]Institute for Financial and Actuarial Mathematics, Department of Mathematical Sciences, University of Liverpool, United Kingdom.
[2]*Corresponding Author*. Institute for Financial and Actuarial Mathematics, Department of Mathematical Sciences, and Institute for Risk and Uncertainty, University of Liverpool, United Kingdom. *Tel*: +44 (0)151 794 5079, *Email*: app@liv.ac.uk.
[3]Financial Network Analytics (www.fna.fi), London UK

1. Introduction

Over the last two decades there has been much focus on how network theory can be used to explain and better understand financial markets. Networks can be used to model the interactions between banks and other financial institutions. Interbank markets have been covered extensively in the literature, for example Boss *et al.* (2004) constructed networks to model the Austrian banking system which consists of many sectors and tiers. Soramäki *et al*. (2007) described the topology of the interbank payment system in the USA. Iori *et al*. (2008) used network topology to analyse the Italian overnight money market and the lending/borrowing that occurred between foreign banks and Italian banks of various sizes. Li (2010) used data from Japan to construct a directed network model and also provided a summary of the banking systems in several other countries.

As well as looking at the structure of financial systems the literature has covered robustness and contagion in financial networks. Allen and Gale (2000) and Leitner (2005) both modelled contagion in the banking networks. Becher *et al.* (2008) used data from CHAPS Traffic Survey 2003 to illustrate the broad network topology of the interbank payments in the UK. Galbiati and Soramäki (2012) modelled clearing systems as networks whose function is to transform exposures and studied how their topology affects the resulting exposures and margin requirements.

One area with significant recent developments is that of *correlation* based networks. These networks can be used to reduce complexity of financial dependencies and to understand and forecast the dynamics in financial markets. Mantegna (1999) introduced a method for finding a hierarchical arrangement for a portfolio of stocks by extracting the *minimum spanning tree* (MST) from the complete network of correlations of daily closing price returns for US stocks. This method has been expanded using coordination numbers by Vandewalle *et al.* (2001) and also applied to other markets such as global stock exchange indices by Bonanno *et al*. (2000) and currency markets by Mizuno *et al*. (2006). More recently Brookfield *et al*. (2012) examined the properties of the MST as applied to the book-to-market ratio and market returns. This technique was studied further by Onnela *et al*. (2002) who considered the effect of a stock market crash on the minimum spanning tree, or asset tree, using the 1987 stock market crash as evidence. They concluded that there was strong shrinkage of the asset tree during the crash, with the normalised tree lengths decreasing and remaining low for the duration of the crash. Onnela *et al*. (2003) extended their study with the introduction of the Asset Graph (AG) - a network structure similar to the minimum spanning tree where a network with n vertices has $n-1$ edges; however the algorithm for the asset graph selects the largest $n-1$ correlations regardless of the resulting structure. The MST and AG are methods for reducing the complete network to a basic minimum structure that contains only the most relevant information and, in the case of the MST, the general hierarchical structure. To build from these Tumminello *et al*. (2005) proposed an algorithm where the complete network can be filtered at a chosen level, by varying the genus of the resulting filtered graph. So if a graph is embedded on a surface with genus = g, as g increases the resulting graph becomes more complex and so reveals more information about the clusters formed, while keeping the same hierarchical tree as the corresponding MST. The simplest form of this graph is the PMFG, on surface $g = 0$.

The aim of this paper is to consider these advantages of the three methods for filtering pertinent information from a series of complex networks modelling the correlations between stock price returns of the DAX 30 blue chip[1] stocks for the time period 2001-2012 (see Appendix A). The dataset, created from Thomson Reuters Datastream[2], consists of the closing prices, adjusted for dividends and splits, of the 30 stocks that form the DAX 30 for the time period between 2001 and 2012. This is a significant time period for the German economy as the euro area was established 1st January 1999 and Germany officially accepted the euro as its legal tender on 31st December 2001. Since its establishment in 1999 the euro area has had several periods of financial crisis; however these have not always been reflected by the German economy - the largest economy within the euro area (see Figure 1).

[*Insert FIGURE1 about here*]

[1]**Appendix A** contains a table of all stock symbols used throughout the paper as well as the sector and subsector of each company
[2]Thomson Reuters Datastream 5.0 (www.thomsonreuters.com)

There have been several periods of recession for the EU and the euro area. After the introduction of the euro certain countries within Europe suffered a decline in their GDP between 2001 and 2004[3]. After a period of recovery and economic growth, Europe was affected by the 2007-2009 financial crisis led by the U.S. subprime mortgage crisis. Finally on the 15th November 2012 the euro area officially entered recession for the second time in four years, despite continuing growth in the largest economies of the area – Germany and France. The latter period is often referred to as the European sovereign-debt crisis (or the euro zone crisis). It is important to note that a financial crisis in the euro zone will affect countries at different times and the rate of recovery will vary depending on the state of the country's economy prior to the crisis.[3]

The German economy is predominantly based on exports; with exports accounting for almost 38% of its GDP[4]. This means that the status of the exports market can be a significant factor for growth within the German economy. As the value of the euro increased through 2002 the German economy once again fell into recession (see Figure 1) with a possible factor being the undesirable exchange rate between the euro and major currencies affecting the export markets with the increased price of goods produced in Germany. The financial crisis 2007-2009 also had an effect on the export markets when a lack of orders and sales resulted in a severe fall in German exports from 2008 Q4[5] (in 2008 Germany was the 3rd largest exporter in the world). However, a weak euro can have a positive effect on the export market and thus on the German economy - a record high of 2.2% GDP growth was reported for the 2nd quarter of 2010 (see Figure 1). As we can see from Figure 1 the quarter-on-quarter volume growth of GDP for 2012 Q1, Q2 and Q3 were 0.5%, 0.3% and 0.2% respectively, meaning that Germany has avoided a further recession, unlike many countries within the euro area (e.g. Greece, Spain etc).

The structure of the remainder of the paper is as follows: the formation of the correlation matrix from the daily adjusted prices is discussed in Section 2. We discuss each network structure in Sections 3, with Sections 3.1-3.3 covering the Minimum Spanning Tree, Asset

---

[3]For more details please refer to 'The euro and economic growth' (April 2005) speech by Lucas Papademos and ECB statistics for Quarter-on-Quarter growth of GDP and expenditure components.

[4]For further details refer to Germany's Economy: Domestic Laggard and Export Miracle. Michael Dauderstädt. Friedrich Ebert Stiftung. November 2012.

[5]ECB statistics. Year-on-Year volume growth of GDP and expenditure components: 2.4 Exports (Q-on-Q).

Graph and Planar Maximally Filtered Graph respectively. Section 4 is an analysis of the three network structures for two specific time periods Oct 2008 – Dec 2008 and May 2010 – Aug 2010. In this section we also discuss the possible underlying economic reasoning for some aspects of the network structures produced. This analysis includes 3-clique and 4-clique analysis from the PMFG. Finally, an overall conclusion is offered in Section 5.

2. Data

We begin by taking the daily closing prices, adjusted for dividends and splits, of the 30 stocks that form the DAX 30 for the time period between 2001 and 2012. The members of the DAX 30 can change as it is reviewed quarterly and so we take the current 30 members for each time period considered.

Denoting the adjusted closing price of stock *i* on day *t* as $P_i(t)$ we calculate the daily logarithmic returns of the stock prices Y_i, as:

$$Y_i = \ln P_i(t) - \ln P_i(t - \Delta t). \qquad (1)$$

Bonanno *et al.* (2004) considered the affect that varying the time horizon, Δt, has on the hierarchical organisation of stocks. For our work we use one trading day, setting $\Delta t = 1$. To look at the affiliation between the price returns of stocks *i* and *j* we calculate the pair-wise correlation coefficient using Pearson product-moment correlation for all trading days in the time period:

$$\rho_{ij} = \frac{\langle \mathbf{Y}_i \mathbf{Y}_j \rangle - \langle \mathbf{Y}_i \rangle \langle \mathbf{Y}_j \rangle}{\sqrt{\left(\langle \mathbf{Y}_i^2 \rangle - \langle \mathbf{Y}_i \rangle^2\right)\left(\langle \mathbf{Y}_j^2 \rangle - \langle \mathbf{Y}_j \rangle^2\right)}}, \qquad (2)$$

where $\langle \cdot \rangle$ is an average over the time period. We use a moving window technique when calculating the correlation coefficient matrices – so the data is separated into annual sets and then we consider a time period of 23 observations (based on an average number of trading days per month) with an interval of 10 days chosen for simplicity. Note that the final window for each annual set will not necessarily contain 23 observations but will end with the last observation for that specific year. For *n* stocks, this results in an $n \times n$- matrix with all entries within the interval $[-1,1]$. These end values correspond to total anti-correlation between stocks i and j and complete linear correlation between stocks i and j. $\rho_{ij} = 0$ represents no correlation between stocks i and j.

As discussed, the DAX 30 is reviewed quarterly so members can be removed or added to the DAX 30 during certain time windows we consider. For consistency we remove the stocks that are not present throughout the entire time window resulting in some having 28 or 29 stocks rather than 30. For example, 22nd September 2003 saw the regular exit of MLP and the entry of Continental (CON). When modelling the 2003 data we have a time window ranging from 10th September – 10th October 2003 which had 29 stocks as MLP and CON were both omitted. This is done automatically with FNA[6].

In the following sections we consider various correlation based networks that have been presented in the literature as a way of filtering the most relevant data from the complete networks.

3. Network Structures

3.1 Minimum Spanning Tree (MST)

The first structure that we consider is the Minimum Spanning Tree (MST). As discussed above, the MST was used by Mantegna (1999) to show the hierarchical arrangement of a portfolio of stocks. The MST extracts the most relevant connections from the correlation matrix and directly gives the subdominant ultrametric hierarchical arrangement of stocks. The stocks are clustered in a way that is entirely based on their correlations and Mantegna (1999) noted how this seems to be related to their economic sector.

Let $G(V,E)$ be a connected, undirected graph, where V is the set of vertices and E is the set of edges. A spanning tree $S(V,E)$ of the graph G is a sub-graph that is a tree connecting all vertices of G, so if the number of vertices $|V| = n$ then the number of edges $|E| = n-1$. For a graph $G(V,E)$ with positively weighted edges we can select the MST - a spanning tree where the sum of the edge weights is less than or equal to that of all other spanning trees. The MST is unique if all of the edge weights are distinct. Various algorithms have been proposed to construct a MST such as Kruskal (1956) and Prim (1957). In our paper, the Kruskal's algorithm has been applied as a method most commonly used in the literature. To be able to con-

[6]See FNA.fi for further details.

struct a MST we need to define the distance between the vertices and the main method used in the literature is to construct the network using the Euclidean metric.

The distance between the stocks is defined so that the three axioms[7] of a metric space are satisfied.

We cannot construct a MST directly from the correlation coefficient matrix as using the correlations as distances would not satisfy these metric axioms – in particular, they do not satisfy the positive definiteness axiom as the correlations range from –1 to 1. Also a stock correlated with itself would give a correlation of 1 and not 0 as required by the first axiom. Furthermore, it is possible to have a high correlation between two stocks but for each of these stocks to have a low correlation with a third stock, which would thus not satisfy the third axiom. To transform the correlation matrix into a distance matrix, a metric function that incorporates the correlation coefficient and satisfies all axioms is needed. We have used a distance function used by Mantegna (1999) based on work by Gower (1966):

$$d(i,j)=\sqrt{2\left(1-\rho_{ij}\right)}. \tag{3}$$

where $d(i,j)$ is the distance between stock i and stock j and ρ_{ij} is the Pearson product-moment correlation coefficient between stock i and stock j. With this distance function we create networks where the shorter the edge length between the vertices (i.e. stocks) the higher the correlation between the stocks.

The 30×30 correlation coefficient matrix, C, is converted to a 30×30 distance matrix, D, using the distance function shown in (3). The $n\left(n-1\right)/2=435$ distances from the upper triangular section of D are then placed in ascending order, so that we can apply Kruskal's algorithm.

The advantage of constructing this network compared to other methods (AG, PMFG) is that, when calculated in this way, the MST directly determines the subdominant ultrametric

[7] 1) For all $p,q,r\in S$ we have $d\left(p,q\right)\geq 0$ and $\left(p,q\right)=0\Leftrightarrow p=q$ (Positive Definiteness), 2) we have $d\left(p,q\right)=d\left(q,p\right)$ (Symmetry) and 3) we have $d\left(p,r\right)\leq d\left(p,q\right)+d\left(q,r\right)$ (Triangular Inequality), where S is a set and d is a metric on S.

distance matrix. The axioms[8] for an ultrametric space are similar to that of a metric space. The subdominant ultrametric is a *unique* ultrametric space that satisfies these axioms and also $u(p,q) \le d(p,q)$. The subdominant ultrametric distance matrix, $D^<$, can be calculated where the entry $d^<(i,j)$ shows the maximum value of any Euclidean distance from all edges in the shortest path connecting *i* and *j* in the MST. This means that a stock *i* with two different Euclidean distances between itself and two other stocks, say *j* and *k*, can have the same ultrametric distance between itself and stocks *j* and *k*. These stocks with the same ultrametric distance can then be clustered together, leading to another method for data reduction: hierarchical clustering. This can be shown using a hierarchical clustering structure (known as a hierarchical tree or dendrogram) which can also be obtained using methods such as Single Linkage Cluster Analysis (SLCA) and Average Linkage Cluster Analysis (ALCA). The SLCA converts the original correlation matrix C into the subultrametric distance matrix $D^<$ by reducing C using an algorithm that selects the maximum correlations. The ALCA reduces the correlation matrix in a similar way; however the resulting matrix and dendrogram vary slightly to that produced by the SLCA as the algorithm uses average subultrametric distances between vertices. Although we will not discuss these algorithms further here, we note that Tumminello *et al*. (2008) provides a detailed explanation.

Schaeffer (2007) defined graph clustering as the task of grouping vertices of a graph into clusters taking into consideration the edge structure of the graph so that there are many edges within each cluster but relatively few between the clusters. As the MST does not contain cycles we consider clusters as the groups of vertices with high weighted edges between them. Possible reasons for the formation of these clusters are discussed in Section 3.2. The MST is probably the most severe form of data reduction. To satisfy the construction algorithm for the MST we may have to omit higher correlations in place of lower correlations so as to keep the resulting graph acyclic. This can be misleading – implying relationships exist between some stocks when they do not.

[8] 1) For all $p,q \in S$, $u(p,q)=0 \Leftrightarrow p=q$, 2) we have $u(p,q)=u(q,p)$ and 3) we have $u(p,r) \le \max[u(p,q),u(q,r)]$, where *S* is an ultrametric space and *u* is an associated distance function.

3.2. Asset Graph

The Asset Graph (AG) was introduced by Onnela *et al.* (2003) as a network similar to the MST but as one that includes all the strongest correlations. The time dependent graph $G^t\left(V^t, E^t\right)$is constructed from either the $n(n-1)/2$ entries of the upper or lower triangular section of the distance matrix, D^t. Note that the distance matrix is calculated using the distance function in equation (3). The $n(n-1)/2$ distances are placed in ascending order. As with the MST, the asset graph has $n-1$ edges however now the set of edges chosen are the $n-1$ smallest distances from the ordered list. With this selection the set of edges E^t are the $n-1$ strongest correlations (as shorter distances correspond to stronger positive correlations) and are chosen regardless of whether they form cycles within the network. A similar approach to the AG is to create threshold networks. Tse *et al*. (2010) constructed a threshold network from closing price data on US stocks, using a winner-take-all approach. This method reduces the complete network to a less complex one by including an edge between two stock prices if their cross correlation is larger than a set threshold value. The complexity of the resulting network can be determined by varying this threshold value.

The AG is useful as it again gives us an idea of the clusters formed by the stocks. The graphs created tend to consist of some clique[9] components with the remaining vertices forming either 1 or 2 edges with other vertices or being completely unconnected. As both the AG and the MST contain $n-1$ edges we can make comparisons between the two networks, with the AG being useful in identifying any misleading selections made by the MST construction algorithm.

From the MST and AG we get a clear indication of the clusters that form between the stocks. These can be stocks from within the same economic sector, for example if we take the set of stocks Bayer (BAYN), BASF (BAS) and Linde (LIN), all within the chemical sector, and look at the 25 MSTs for 2007 we see that at least 2 of these stocks are connected in 80% of the networks and actually all 3 are connected in 32% of the networks. In addition we notice that ThyssenKrupp (TKA) is often connected to this set of stocks, sometimes forming the link between 2 of the connected stocks from the set and the third stock. TKA is in the indus-

[9]For a graph $G(V,E)$, a subset of vertices $C \subseteq V$ is called a clique if $G(C)$ is a complete graph.

trial sector; LIN produces industrial gases and so is classed as being in the industrial gases subsector. Thus the 4 stocks would form a cluster based on their sectors and subsectors. With the AGs from the same time period we see that it is BAYN that is the central vertex in this group – with a connection between BAYN and BAS in 48% of the networks (an average correlation of 0.7107). A similar example can be seen with the set of stocks BMW, Daimler (DAI) and Volkswagen (VOW3), all within the automobile sector, and the networks for 2004 data. The MSTs show that least 2 of these stocks are connected in 80% of the networks and all 3 are connected in 40% of the networks. The AGs for 2004 also show the strong correlations between these stocks, but they also identify that there are strong correlations between two insurance companies, Allianz (ALV) and Munich Re (MUV2), and BMW and DAI. This was something not shown with the MSTs.

Another example is between the two stocks that belong to the utilities sector, EOAN and RWE. If we look at the 25 MSTs and AGs for 2009 we see that the two stocks are connected in 72% of the MSTs and in 64% of the AGs. We can see many clusters of this form are present within the networks and we can identify them using the MST and AG. There are, however, other reasons that these clusters may form that may not be immediately clear. Companies from different sectors can form partnerships or be involved in mergers and acquisitions. For example in January 2003 Siemens (SIE) acquired majority control in Sinius GmbH, a technology service set up by Deutsche Bank (DBK)[10]. In the 25 MSTs and AGs for 2003 SIE and DBK are connected in 56% of the MSTs and 72% of the AGs (an increase from the previous 2 years). Although we have not considered any social influences, e.g. companies having the same board members, the impact this can have on the networks has been discussed in Halinen and Tornroos (1998).

The disadvantage to this method is that we do not get a complete image due to the disconnected vertices. Also, as with the MST, it favours strong, positive correlations. To show this disadvantage we highlight from our data the correlations for VOW3 from 26th August 2008 – 18th December 2008. After several years of acquiring VOW3 shares, Porsche owned 42.6% of VOW3 shares outright and had derivative contracts for a further 31.5% by October 2008 (with 20% of VOW3 shares being Government owned) when they revealed plans to increase

[10] See http://www.highbeam.com/doc/1P1-70855773.html

this stake to 75% during 2009[11]. There was a rapid increase in the price of VOW3 shares, which was encouraged by Porsche buying options to purchase more shares. On 29th October Porsche announced they would settle up to 5% of VOW3 options, resulting in a fall in the price of VOW3 shares. During this time period the returns of VOW3 showed some unusual patterns and as a result the correlation matrices showed a negative correlation between the returns for VOW3 and most other stocks (in some cases with all other stocks e.g. the correlation matrix for 23rd September – 23rd October 2008). Due to the nature of the construction algorithms these returns were not shown by the MST or the AG.

3.3. Planar Maximally Filtered Graph

The final network that we discuss is the filtered graph proposed by Tumminello *et al*. (2005) with particular focus on the planar[12] filtered graph (PMFG - created when the graph is embedded into a surface with genus set equal to 0). The networks discussed so far are a severe form of data reduction, containing the minimum number of edges. The proposed filtered graphs allow us to choose how much information we filter from the complete network, so by increasing the genus of the surface we are able to construct a more complex network containing more edges. The PMFG is constructed in a similar way to the MST. For a graph $G(V,E)$ with $|V| = n$ and $|E| = m$, all edges, e_1, e_2, …, e_m, from the upper triangular section of C are placed in descending order $e_{(1)}, e_{(2)}, ..., e_{(m)}$. Select the first edge $e_{(1)}$ and construct a graph with $e_{(1)}$ and the two vertices that it connects. Continue selecting the ordered edges and add them to the network structure only if the resulting network can be drawn on a planar surface without edges crossing. (There are some tests for planarity based on Kuratowski's theorem. For more detail on these and others see Hopcroft and Tarjan (1974)). The algorithm ends when all vertices $v_1, v_2, ..., v_n$ are connected, using $3(n-2)$ edges.

The advantage of the PMFG is that it will always contain the corresponding MST and so shows some of the clusters between stocks, but also provides additional information. Unlike the MST, the PMFG does not have a unique path between each of the vertices. This means

[11] See http://www.economist.com/node/12523898

[12] A planar graph is a graph that can be drawn in a plane without edges crossing (i.e. the only intersection between edges occurs at vertices). A plane graph is a graph that is drawn on a plane in the planar graph structure i.e. drawn without the edges crossing.

that we cannot identify the hierarchical clustering between stocks using the subdominant ultrametric distances in the direct way that we can with the MST. However, as the construction algorithm allows the inclusion of loops the PMFG contains cliques, as with the AG, so we can extract further information from the network by analysing these cliques. Looking specifically at the PMFG, we consider 3- and 4-cliques as the maximum number of elements that can form a clique is 4[13].

We can say that the PMFG is the triangulation of a sphere as the network consists entirely of 3-cliques (triangulation of a surface is a partition of that surface by triangles into facets). With our dataset of 30 stocks, we have a total of $\binom{n}{3} = 4060$ possible combinations of 3-cliques from each complete graph. By constructing the maximally filtered graph we considerably reduce the connectivity of the network leaving the most relevant cliques. The following Lemma is useful for our analysis.

Lemma 1 - *Let* $G(V,E)$ *be a PMFG with n vertices. Then, the number of facets is 2n – 4.*

Proof. Let *n* be the number of vertices, *e* be the number of edges and *f* the number of facets belonging to G. For a PMFG we have $3(n-2)$ edges and can obtain the following from Euler's formula:

$n-e+f=2$ (when we include the unbounded area)

$n-3(n-2)+f=2$, and

$$f = 2n-4. \tag{4}$$

□

As mentioned above the PMFG is the triangulation of a sphere so the facets will all be triangles and from the lemma we see that the maximum number of these surface triangles is $2n-4$. Tumminello *et al*. (2005) states that the maximum number of 3-cliques formed by a PMFG is $3n-8$, which has been proven by Birch *et al.* (2014). This means that the maximum number of triangles on the plane can be significantly less than the number of triangles in the PMFG (the number of 3-cliques) and so indicates that several of the 3-cliques do not appear on the surface. Aste *et al*. (2005) explain these as collar rings – triangular rings that

[13]For more details please refer to G. Ringel in Map Color Theorem. Springer, Chapter 4 (1974).

belong to the tetrahedrons that form when four 3-cliques share common edges (the possible structures of 3-cliques are discussed further in Birch *et al.* (2014)).

We analyse the 4-cliques by showing the sectors that the four stocks belong to as well as the average correlation coefficient inside the clique, the range between the highest and lowest correlation coefficient in the clique and the standard deviation. It is worth noting that Tumminello *et al*. (2005) states the maximum number of 4-cliques formed by a PMFG is $n-3$ and this is also proven by Birch *et al*. (2014).

Let us consider some of the examples highlighted in the previous sections. We have noted from the 2007 MSTs and AGs that BAS, BAYN and LIN often formed a cluster and they all belong to the chemical sector. For the PMFGs for 2007 the three stocks are connected in 60% of the networks and actually form a 3-clique in 44% of the networks. We also considered the cluster of stocks in the automobile sector for the 2004 data. These clusters are also shown in the PMFGs – with the three stocks being connected in 72% of the networks and a 3-clique forming in 44% of the networks. Finally the two stocks in the utilities sector, RWE and EOAN, were connected in a high proportion of the MSTs and AGs for 2009 and this was also the case with the PMFGs with a connection in 84% of the PMFGs for 2009.

Cliques also allow us to identify the most connected stocks so that they can not only be clustered but also separated into two sets: *core* and *periphery*. This can be done using the AG as, due to the construction algorithm, we often have clique components and unconnected vertices. However, the benefit of the PMFG is that, as it is a connected network, we have a better understanding of the relationships between the stocks that are not identified as being within the core.

Unlike the MST and AG the PMFG does not necessarily favour the strong, positive stocks. We highlighted VOW3 as an example of a stock that was not fairly represented in the 2008 networks due to its negative correlation. For the PMFG in 2008 we see that VOW3 is mainly connected to 3 other stocks (84% of the networks) and these are mostly other stocks from the automobile sector (BMW, CON and DAI). At most it is connected to 6 other stocks (this included BMW and DAI). It forms 3-cliques and in some networks a 4-clique, although this 4-clique has a lower average correlation compared to the others from the same PMFG due to the negatively correlated VOW3.

Table 1 provides a summary of all filtering methods covered in this section and their advantages and limitations.

[*Insert TABLE1 about here*]

4. Analysis of DAX 30

So far we have made comparisons between each of the network structures and discussed their construction and the possible information we can extract. The filtered networks extract clusters of stocks from the complete networks which have high correlations between their return prices. These clusters often form between stocks that belong to the same economic sector and subsector with cross-sector clusters appearing less frequently. There may be some economic reasons for these cross-sector clusters; however they could also be due to errors with the multiple simultaneous estimates made when creating the correlation matrices, such as type I errors (i.e. false positives - identifying a correlation when one does not exist). To this end, we have included the Bonferroni correction parameter when constructing the networks with FNA. For the Bonferroni correction the familywise error rate (the probability of making one or more type I error among all hypotheses when performing multiple tests) is set to the chosen level of α (here α = 0.5) and each individual test is performed at significance level $a^* = a / m$ where m is the total number of tests performed. The edges in the network structures can then be classified as being significant or not significant.

We now discuss two specific time periods in more detail, discussing possible economic reasons for some of their features.

4.1 Period of Crisis

The first time period assessed is 7th Oct 2008 – 31st Dec 2008 and includes 2 important dates; in October 2008 the German government, market regulators and other financial institutions agreed a €50 billion rescue plan (originally €35 billion, a later deal with an additional €15 billion was agreed on 5th October 2008) to prevent the collapse of Hypo Real Estate, the second largest commercial property lender. This was a sign of the economic problems in Germany – the GDP had declined 0.4% in the 2nd quarter of 2008 and a further 0.4% in the 3rd quarter of 2008 meaning as of 13th November 2008 the German economy was officially in a state of recession (see Figure 1).

[*Insert FIGURE2 about here*]

The diameter of the MST increases as we move through the time period – this implies that the distances between the vertices is increasing and so the correlations are decreasing. There are some clusters that form – the two stocks from the utilities sector (RWE and EOAN) are strongly connected in 4 of the 5 MSTs. Stocks in the FIRE (Finance, Insurance and Real Estate) sector (particularly the three banks Commerzbank (CBK), DBK and Deutsche Börse (DPB)) are also strongly connected, across the first 4 MSTs. However, in the final MST many of the edges that connect stocks from the FIRE sector to the tree are classified as insignificant – including CBK, DBK, DPB and MUV2. For the remaining MSTs the edges shown to be insignificant were rather predictable – mainly the edges connecting VOW3, HRE and IFX to the networks for the period of crisis. The correlations that the test has found to be insignificant in the MST are the lower correlations that may have only been chosen to satisfy the construction algorithm.

Some of the clusters identified in the complete data set are not present in the MSTs – such as the automobile and the chemical cluster.

[*Insert TABLE 2 about here*]

We have seven stocks that are not included in any of the AGs for this time period. VOW3 has been previously discussed. CON and (Hypo Real Estate) HRE were both excluded from the DAX 30 on the basis of the fast-exit rule in December 2008 and similarly (Deutsche Postbank) DPB and (Infineon Technologies) IFX were excluded in Q1 of 2009. The final 2 stocks that were not included are (Fresenius Medical Care) FME and (Metro) MEO. We can see from the complete dataset that for some years FME does not cluster with any other stock and is included in very few AGs between 2002 and 2004 (actually it is not included in any AG for 2003). This could be because the company is fairly unique, being the only healthcare company included in DAX30 at this point. Let us consider the correlations of the stocks that were included in the AGs. Across the series there is a decrease in correlations – the highest correlated pair falling from 0.9607 for the first AG to 0.8409. Although this is not a significantly large decrease if we consider that for the first AG the lowest correlated pair (the 29th and therefore last edge to be included) was 0.8631 we can see that there has been a decrease in the correlation throughout the complete graph. This supports what is shown by the increase in the diameter of the MSTs.

[*Insert TABLE3 about here*]

[*Insert FIGURE3 about here*]

From the PMFG we can consider the changes observed in the 4-clique analysis (average correlation within the clique, the range and the standard deviations). We also take into account the number of 4-cliques that were observed compared to the maximum total number of 4-cliques that were possible within the graph and the economic sectors that the stocks of each clique belong to. We can see from Table 3 that each PMFG had the maximum number of 4-

cliques possible for the number of vertices included. Interestingly at least one 4-clique formed in each PMFG containing VOW3. The average correlation within this clique was lower than the other averages (due to VOW3 being negatively correlated to all other stocks during this time period). For now we will omit the clique containing VOW3 from the following discussion as this was identified as a special case and explained above (Section 3.2).

Overall we can see a decrease in the average correlation within the 4-cliques - the highest and lowest averages for 7th October - 6th November were 0.9044 and 0.7821 respectively whereas for 2nd December - 31st December the highest was 0.7967 and the lowest 0.4168. When considering the economic sectors that the stocks of the 4-cliques belong to we can see from Table 3 that there are many cliques where all 4 stocks are from a different sector.

To further analyse the 4-cliques we compute a quantity $\langle y \rangle$, as shown by Tumminello *et al.* (2005), to calculate the spread of the correlation among the stocks belonging to each clique (where $\rho_{ij} \geq 0$). $\langle y \rangle$ is the mean value of the disparity measure[14] over the clique and for a clique with all correlations shared evenly between the stocks within the clique $\langle y \rangle = {}^{1}/_{3}$.

For the cliques contained in the PMFGs for this first time period, most have the expected value $\langle y \rangle \approx 0.333$. Within each of the PMFGs for 21st October – 20th November, 4th November – 4th December and 18th November – 18th December there are 3 cliques that have $\langle y \rangle$ slightly higher than a 0.34 (ranging from 0.341 to 0.365). Each of these cliques continued one of the seven stocks mentioned above that were omitted from all AGs for this overall time period. For the final PMFG for this time series (2nd December- 31st December) the value for $\langle y \rangle$ was greater than 0.34 for 11 cliques. The highest value was 0.474 for a 4-clique formed with Deutsche Telekom (DTE), FME, MEO and VOW3. This PMFG is the only one for this time period where VOW3 has non-negative correlations; however they are very small in comparison to the others which would explain the higher $\langle y \rangle$ value.

If we consider the edges that have been classified as insignificant within the PMFG we can see that, as with the MST, it is mainly the edges connecting the vertices VOW3, HRE and IFX for the first networks in the series. However as each vertex in the PMFG has a degree of

---

[14] $y(i) = \sum_{j \neq i, j \in clique} \left[\frac{\rho_{ij}}{s_i}\right]^2$ where s_i is the strength of element i.

at least 3 there were more edges that were classified as insignificant compared with the MST. The final PMFG in the series, representing data from 2nd December to 31st December 2008, actually has a larger number of edges classified as insignificant – with vertices from a range of sectors having all the edges connecting it to the remaining network being insignificant.

4.2 Period of Recovery

The second time period between 7thMay and 3rdAugust 2010 is considered a time of economic success for Germany. With the country officially out of recession in August 2009, there was a significant growth in the country's exports and with that a 3.6% growth to their economy in 2010. The 2nd quarter of 2010 showed a record high in the GDP growth rate (2.2%) (see Figure 1).

For the AGs created for this time period the only stock that is not included is Merck (MRK). FME and (Fresenius) FRE are only included in one AG where they form a separate component with only 1 edge between each vertex. These are the only 3 stocks in the pharmaceutical and healthcare supersector and, although we cannot comment on the performance of the companies based solely on the networks, we can say that their price returns do not appear to follow the same patterns as the other stocks. MEO is also only included in one AG – for the two time periods this is the only stock in the multiline retail subsector. The range from the largest to smallest correlation in the AGs increases across the time period and they are generally not as high as in the previous time period.

We can see again from the MSTs that the edges classified as being insignificant are fairly predictable – mainly the edges connecting FRE, FME and MRK to the networks for the period of recovery. The MSTs show some clear clusters based on the economic sectors that the stocks belong to – particularly the automobile, chemical and FIRE sector.

[*Insert FIGURE4 about here*]

[*Insert TABLE4 about here*]

[*Insert TABLE5 about here*]

[*Insert FIGURE5 about here*]

We can see from Table 5 that, unlike the first time period, the maximum possible number of 4-cliques did not form in the PMFG. The most significant example of this is during the time between 4th June – 6th July 2010 when only 16 from the possible 26 cliques formed. This is interesting as the AG actually included more of the 30 stocks compared to the AGs constructed for the first time period. A possible reason for this could be that only stocks from certain sectors were performing well – stocks in the FIRE sector and companies that produce goods for exports. This is something that we would need to consider in further detail. Overall the average correlations for the 4-cliques were generally lower for the second time period when compared to those of the first. If we compare the values calculated for $\langle y \rangle$ to the values calculated in the first time period we see that there are even fewer cliques that have $\langle y \rangle$ greater than 0.333, with the highest value being 0.355. There are 13 cliques for the whole time series that have $\langle y \rangle$ higher than 0.34 and of these, all but five contain one or more of MEO, FME, FRE or MRK which have been omitted from or shown in only one AG for this time period.

Tumminello *et al*. (2005) used a similar 4-clique analysis to investigate 100 US stocks from January 1995 to December 1998. The total number of 4-cliques formed was 97 and of these, 31 had all four stocks in the same economic sector and 22 had three in the same economic sector. Our 4-clique analysis actually showed that it was more likely for a 4-clique to form with each stock in a different sector or at most two stocks to be in the same sector. Possible reasons for this difference could be that the time periods considered were not ‘average days’ as they were a period of crisis and recovery. To test this further an analysis of the 4-cliques that were observed over the whole time period would need to be considered. The German DAX 30 is also considerably smaller than the 100 US stocks considered by Tumminello *et al.* (2005).

Looking at the edges that have been classified as insignificant within the PMFG we see a similar pattern to the PMFGs for the period of crisis. The first networks in the series show that the edges that connect the vertices FME, FRE and MRK to the remaining network are insignificant (the same vertices identified within the MST). For the remaining PMFG a larger number of edges are shown to be insignificant, although MEO and MRK are the only vertices to have all their edges classified as insignificant. This could show that they are not highly correlated with other stocks with the network and have only been included to satisfy the construction algorithm. This supports what was shown with the AGs for the period of recovery.

Some of the correlations may be driven by a third factor, such as markets moving up or down in general. To control for this we apply Principal Component Analysis (PCA). PCA identifies patterns in data and expresses data in a way to highlight these similarities so we can control the effect of common factors such as the market return. As PCA needs a complete dataset some vertices were omitted if they were not present throughout the whole time period i.e. for the period of crisis BEI, CON, HRE, SZG and TUI were omitted from the networks and for the period of recovery HEI and SZG were omitted. When performing PCA with all components we found that there was very little difference between the resulting networks and the original networks. However, following Laloux *et al.* (1999), for a second analysis we removed the first and largest component as this most likely represented the variance due to the market return and also removed components greater than component 6 as less than 1.5% of the variance was explained by these components. These networks were slightly different to our original networks but this could be due to the missing vertices. They still supported the findings from our analysis.

5. Conclusions

In summary, we have shown three possible methods for filtering information from a complete network of the correlations of the daily adjusted closing prices for DAX 30 stocks. The minimum spanning tree reduces the complete network to the minimum connected structure and can be used to show the hierarchical clustering of the stocks. The clusters that form are likely to be between stocks in the same economic sector. The asset graph separates the complete network into components – generally complete cliques and unconnected vertices. The planar maximally filtered graph combines these two methods by showing some hierarchical clustering, as it will contain the corresponding MST, and also highlighting the most connected stocks, as with the AG.

We have considered two time periods in detail - a period of crisis and of recovery. Overall we can see that during the period of crisis the correlations decreased throughout the time period and they were generally lower than during the time of recovery. The AGs for the period of recovery had less unconnected stocks than the period of crisis, although the stocks not included in the AGs for the first time period seemed to show some companies that would be omitted from the DAX 30 during, or soon after, the crisis time period. There were fewer clusters for the first time period compared to the second time period – which contained clusters of stocks from the same economic sectors. We note from the 4-clique analysis that the cliques that formed in both time periods contained stocks from three or four different sectors, rather than from one sector as in the literature.

As this is a fairly new area of research there is the possibility to develop the methods further. We shall consider other distance metrics that could be used instead of Euclidean distance, as the length of the edges in the current network visualisations are not proportional to the correlations. We would also like to extend the 4-clique analysis to include more networks from across the whole time period to include more ‘average’ trading days. These can then be compared to the cliques formed during the specific time periods that we have considered.

References

Allen, F., & Gale, D. (2000). Financial Contagion. *Journal of Political Economy, 108*: 1–33.

Aste, T., Di Matteo, T., Tumminello, M., & Mantegna, R.N. (2005). Correlation Filtering in Financial Time Series. *Noise and Fluctuations in Econophysics and Finance. Proc. of SPIE, 5848:* 100-109.

Becher, C., Millard, S., & Soramäki, K. (2008). The Network Topology of CHAPS Sterling. *Working Paper* No. 355, *Bank of England.*

Birch, J., Pantelous, A.A., & Zuev, K.. (2014). The Maximum Number of 3- and 4-Cliques Within a Planar Maximally Filtered Graph. *Physica A* (to appear).

Bonanno, G., Vandewalle, N., & Mantegna, R.N. (2000). Taxonomy of Stock Market Indices. *Physical Review E, 62*: R7615-R7618.

Bonanno, G., Caldarelli, G., Lillo, F., Micciche, S., Vandewalle, N., & Mantegna, R.N. (2004). Networks of Equities in Financial Markets. *European Physical Journal B, 38*: 363-371.

Boss, M., Elsinger, H., Summer, M., & Thurner, S. (2004). Network Topology of the Interbank Market. *Quantitative Finance, 4*(6): 677-684.

Brookfield, D., Boussabaine, V., & Su, C. (2012). Identifying Reference Companies using the Book-to-Market Ratio: a Minimum Spanning Tree Approach. *The European Journal of Finance.* 1-25.

Galbiati, M., & Soramäki, K. (2012). Clearing Networks. *Journal of Economic Behavior and Organization, 83*: 609-626.

Gower, J.C. (1966). Some Distance Properties of Latent Root and Vector Methods used in Multivariate Analysis. *Biometrika, 53*: 325-338.

Halinen, A., & Tornroos, J. (1998). The Role of Embeddedness in the Evolution of Business. *Scandinavian Journal of Management, 14*(3): 187-205.

Hopcroft, J., & Tarjan, R. (1974). Efficient Planarity Testing. *Journal of the Association for Computing Machinery, 2*(4): 549-568.

Iori, G., De Masi, G., Precup, O.V., Gabbi, G., & Caldarelli, G. (2008). A Network Analysis of the Italian Overnight Money Market. *Journal of Economic Dynamics and Control, 32*: 259-278.

Kruskal, J.B. (1956). On the Shortest Spanning Subtree of a Graph and the Travelling Salesman Problem. *Proceedings of the American Mathematical Society, 7*: 48-50.

Laloux, L., Cizeau, P., Bouchaud, J.-P., & Potters, M. (1999). Noise Dressing of Financial Correlation Matrices. *Physical Review Letters, 83*.

Leither, Y. (2005). Financial Networks: Contagion, Commitment, and Private Sector Bailouts. *Journal of Finance, 6*(6): 2925-2953.

Li, S., He, J., & Zhuang, Y. (2010). A Network Model of the Interbank Market. *Physica A, 389*: 5587-5593.

Mantegna, R.N. (1999). Hierarchical Structure in Financial Markets. European *Physical Journal B, 11*: 193-197.

Mizuno, T., Takayasu, H., & Takayasu, M. (2006). Correlation Networks Among Currencies. *Physica A, 364*: 336-342.

Onnela, J.-P., Chakraborti, A., Kaski, K., & Kertesz, J. (2002). Dynamic Asset Trees and Portfolio Analysis. *European Physical Journal B, 30*: 285-288.

Onnela, J.-P., Chakraborti, A., Kaski, K., Kertesz, K., & Kanto, A. (2003). Asset Trees and Asset Graphs in Financial Markets. *Physic Scripta*, *T106*: 48-54.

Prim, R.C. (1957). Shortest Connection Networks and Some Generalizations. *Bell System Technical Journal, 36*: 1389-1401.

Schaeffer, S.E. (2007). Graph Clustering. *Computer Science Review, 1*: 27-64.

Soramäki, K., Bech, M., Arnold, J., Glass, R.J., & Beyeler, R.J. (2007). The topology of interbank payment flows. *Physica A, 379*: 317-333.

Tse, C.K., Liu, J., & Lau, F.C.M. (2010). A Network Perspective of Stock Market. *Journal of Empirical Finance, 17*: 659-667.

Tumminello, M., Aste, T., Di Matteo, T., & Mantegna, R.N. (2005). A Tool for Filtering Information in Complex Systems. *Proceedings of the National Academy of Sciences, 102*(30): 10421-10426.

Tumminello, M., Lillo, F., & Mantegna, R.N. (2010). Correlation, Hierarchies and Networks in Financial Markets. *Journal of Economic Behavior and Organization, 75*: 40-58.

Vandewalle, N., Brisbois, F., & Tordoir, X. (2001). Non-random Topology of Stock Markets. *Quantitative Finance, 1*(3): 372-374.

Appendix A

List of all stock symbols and the supersector, sector and subsector that the company belongs to. The details of the various sectors can be found in Guide to the Equity Indices of Deutsche Börse. Version 6.6, Nov 2008. (Deutsche-boerse.com)

Symbol	Company	Supersector	Sector	Subsector
AAA	Altana	Basic Materials	Chemicals	Chemicals, Specialty
ADS	Adidas	Consumer Goods	Consumer	Clothing & Footwear
ALV	Allianz	FIRE	Insurance	Insurance
BAS	BASF	Basic Materials	Chemicals	Chemicals, Specialty
BAYN	Bayer	Basic Materials	Chemicals	Chemicals, Specialty
BEI	Beiersdorf	Consumer Goods	Consumer	Personal Products
BMW	BMW	Consumer Goods	Automobile	Automobile Manufacturers
CBK	Commerzbank	FIRE	Banks	Credit Banks
CON	Continental	Consumer Goods	Automobile	Auto Parts & Equipment
DAI	Daimler	Consumer Goods	Automobile	Automobile Manufacturers
DB1	Deutsche Börse	FIRE	Financial Services	Securities Brokers
DBK	Deutsche Bank	FIRE	Banks	Credit Banks
DGS	Degussa Huls	Basic Materials	Chemicals	Chemicals, Specialty
DPB	Deutsche Postbank	FIRE	Banks	Credit Banks
DPW	Deutsche Post	Industrials	Transportation & Logistics	Logistics
DRB	Dresdner Bank	FIRE	Banks	Credit Banks
DTE	Deutsche Telekom	Telecommunications	Telecommunications	Fixed-Line Telecommunica-tions
EOAN	E. On	Utilities	Utilities	Multi-Utilities
EPC	Epcos	Information Technology	Technology	Electronic Components & Hardware
FME	Fresenius Medical Care	Pharma & Healthcare	Pharma & Healthcare	Healthcare
FRE	Fresenius	Pharma & Healthcare	Pharma & Healthcare	Healthcare
HEI	Heidelberg Cement	Industrials	Construction	Building Materials
HEN3	Henkel	Consumer Goods	Consumer	Personal Products
HNR1	Hannover Re	FIRE	Insurance	Re-Insurance
HRE	Hypo Real Estate	FIRE	Financial Services	Real Estate
HVB	HypoVereinsbank	FIRE	Banks	Credit Banks
IFX	Infineon Technologies	Information Technology	Technology	Semiconductors
KAR	KarstadtQuelle	Consumer Services	Retail	Retail, Multiline
LHA	Deutsche Lufthansa	Industrials	Transportation & Logistics	Airlines
LIN	Linde	Basic Materials	Chemicals	Industrial Gases
LXS	Lanxess	Basic Materials	Chemicals	Chemicals, Commodity
MAN	MAN	Industrials	Industrial	Industrial, Diversified
MEO	Metro	Consumer Services	Retail	Retail, Multiline
MLP	MLP	FIRE	Financial Services	Diversified Financial
MRK	Merck	Pharma & Healthcare	Pharma & Healthcare	Pharmaceuticals
MUV2	Munich Re	FIRE	Insurance	Re-Insurance
RWE	RWE	Utilities	Utilities	Multi-Utilities
SAP	SAP	Information Technology	Software	Software
SCG	Schering	Pharma & Healthcare	Pharma & Healthcare	Pharmaceuticals
SDF	K + S	Basic Materials	Chemicals	Chemicals, Commodity
SIE	Siemens	Industrials	Industrial	Industrial, Diversified
SZG	Salzgitter AG	Basic Materials	Basic Resources	Steel & Other Metals
TKA	ThyssenKrupp	Industrials	Industrial	Industrial, Diversified
TUI	TUI*	Industrials	Transportation & Logistics	Transportation Services
VOW3	Volkswagen Group	Consumer Goods	Automobile	Automobile Manufacturers

Table 1

This table provides a summary of the advantages and limitations of each of the filtering methods discussed in Section 3. For further details on each filtering method and how they are constructed please refer to Section 3.

Filtering Method	Advantages	Limitations
Minimum Spanning Tree	Directly determines the subdominant ultrametric distance matrix. So stocks with the same ultrametric distance can be clustered together. The stocks are clustered in a way that is entirely based on their correlation.	A severe form of data reduction with some data lost. To satisfy construction algorithm we omit higher correlations in place of lower correlations to the graph acyclic. Favours strong, positive correlations.
Asset Graph	Gives a clear indication of the clusters formed between the stocks using clique components. Can identify any misleading selections made by the MST construction algorithm.	Do not get a complete image due to disconnected vertices. Little information known about the disconnected vertices. Favours strong, positive correlations.
Planar Maximally Filtered Graph	Can choose the level of filtering by changing the genus of the surface that the graph is embedded to. Always contains the corresponding MST. Gives a clear indication of the clusters formed between the stocks using clique components. Includes negatively correlated stocks.	Cannot identify the hierarchical clustering between stocks using the subdominant ultrametric distance in the direct way that we can with the MST.

Table 2

The edges that form the AG for 2.a) 7th Oct-6th Nov 2008, 2.b) 4th Nov – 4th Dec 2008 and 2.c) 2nd Dec-31st Dec 2008 listed by the order of their addition and the vertices that the edge connects. We have also shown the correlations corresponding to each edge. Note that the AGs here show the correlations and not the distances so that they can be compared with the correlations in the 4-clique analysis.

	(2.a)			(2.b)			(2.c)		
Edge	Correlation	Vertex	Vertex	Correlation	Vertex	Vertex	Correlation	Vertex	Vertex
1	0.9607	RWE	EOAN	0.9458	DBK	CBK	0.8409	TKA	DPW
2	0.9426	SIE	BAS	0.9326	RWE	EOAN	0.8229	TKA	HEN3
3	0.9314	SIE	DAI	0.9230	DBK	ALV	0.8207	TKA	BMW
4	0.9167	SAP	DAI	0.9194	MAN	DAI	0.8053	DPW	DAI
5	0.9129	DAI	ALV	0.9154	SIE	DAI	0.8039	DPW	BAYN
6	0.9023	RWE	DTE	0.9077	TKA	MAN	0.8030	DAI	BMW
7	0.9023	BMW	ALV	0.9046	MAN	DBK	0.8015	TKA	LHA
8	0.8960	DBK	BMW	0.9026	SIE	MAN	0.7991	TKA	LIN
9	0.8952	SIE	ALV	0.8981	BAYN	ADS	0.7981	SIE	DPW
10	0.8949	RWE	DAI	0.8977	TKA	DAI	0.7933	TKA	MAN
11	0.8926	BAYN	BAS	0.8972	MAN	BAS	0.7909	TKA	BAS
12	0.8893	EOAN	DAI	0.8961	EOAN	DBK	0.7890	SIE	DAI
13	0.8861	SIE	SAP	0.8949	DBK	DAI	0.7868	TKA	DAI
14	0.8851	DAI	BMW	0.8939	CBK	ALV	0.7801	DAI	BAYN
15	0.8848	SIE	MAN	0.8903	TKA	SIE	0.7654	BAS	ALV
16	0.8848	RWE	BAS	0.8903	SIE	BAYN	0.7596	TKA	SIE
17	0.8847	SIE	RWE	0.8894	EOAN	ALV	0.7596	BMW	BAS
18	0.8838	LHA	DBK	0.8892	SDF	MAN	0.7576	SIE	HEN3
19	0.8777	LIN	BAS	0.8861	MAN	DPW	0.7555	LIN	DPW
20	0.8742	EOAN	DTE	0.8847	MAN	CBK	0.7507	LHA	DPW
21	0.8733	SAP	RWE	0.8847	DAI	BAYN	0.7351	HEN3	DAI
22	0.8729	SAP	MUV2	0.8843	HEN3	DAI	0.7339	MAN	DAI
23	0.8689	RWE	MRK	0.8816	DAI	ADS	0.7311	SAP	DBK
24	0.8653	SIE	EOAN	0.8766	TKA	CBK	0.7275	SDF	RWE
25	0.8653	TKA	BMW	0.8766	SAP	RWE	0.7259	DAI	BAS
26	0.8644	TKA	DBK	0.8758	LHA	ADS	0.7258	TKA	BAYN
27	0.8644	TKA	MAN	0.8744	DAI	BAS	0.7183	LHA	ALV
28	0.8637	SIE	LIN	0.8741	EOAN	DAI			
29	0.8631	DAI	BAS	0.8734	SIE	ADS			

Table 3

Results from PMFG Analysis for the first time period (7th October 2008 - 31st December 2008). The column '4-Cliques' shows the observed number of 4-cliques in each PMFG and 'Max 4-Cliques' shows the total number of 4-Cliques possible for each PMFG. The last 4 columns show the number of 4-cliques that formed with stocks in 4 different sectors, in 3 different sectors etc.

	PMFG Analysis						
Dates	Stocks	4-Cliques	Max 4-Cliques	4 Sectors	3 Sectors	2 Sectors	1 Sector
7 Oct - 6 Nov 2008	30	27	27	10	14	3	0
21 Oct - 20 Nov 2008	30	27	27	10	15	2	0
4 Nov - 4 Dec 2008	30	27	27	11	13	3	0
18 Nov - 18 Dec 2008	30	27	27	12	12	3	0
2 Dec - 31 Dec 2008	28	25	25	9	14	2	0

Table 4

The edges that form the AG for **4.a)** 7th May-8th Jun 2010, **4.b)** 4th Jun – 6th Jul 2010 and **4.c)** 2nd Jul-3rd Aug 2010 listed by the order of their addition and the vertices that the edge connects. We have also shown the correlations corresponding to each edge. Note that the AGs here show the correlations and not the distances so that they can be compared with the correlations in the 4-clique analysis.

	(4.a)			(4.b)			(4.c)		
Edge	Correlation	Vertex	Vertex	Correlation	Vertex	Vertex	Correlation	Vertex	Vertex
1	0.9280	VOW3	DPW	0.9099	RWE	EOAN	0.9396	RWE	EOAN
2	0.9119	MUV2	ALV	0.8878	MUV2	ALV	0.8462	FRE	FME
3	0.9119	LIN	ADS	0.8537	DBK	ALV	0.8417	MUV2	ALV
4	0.9092	SIE	DAI	0.8371	TKA	MEO	0.8134	DAI	BMW
5	0.9055	SIE	MUV2	0.8258	TKA	DPW	0.7954	TKA	HEI
6	0.8967	RWE	EOAN	0.8176	MEO	IFX	0.7736	BAS	ALV
7	0.8929	SIE	BAYN	0.8107	MUV2	DPW	0.7591	DBK	CBK
8	0.8910	RWE	DTE	0.8084	DTE	BEI	0.7568	VOW3	MAN
9	0.8905	LIN	DAI	0.7988	VOW3	BMW	0.7553	RWE	DBK
10	0.8875	DAI	BAS	0.7968	LIN	BAS	0.7504	CBK	ALV
11	0.8860	VOW3	DAI	0.7940	LHA	IFX	0.7478	RWE	ALV
12	0.8841	LIN	BAS	0.7897	BAYN	BAS	0.7467	SIE	BAS
13	0.8801	MAN	BAYN	0.7890	DPW	ALV	0.7406	DPW	CBK
14	0.8783	MUV2	BAYN	0.7883	IFX	DBK	0.7308	LIN	BAS
15	0.8707	SIE	ALV	0.7843	HEN3	DAI	0.7308	MUV2	IFX
16	0.8702	VOW3	SIE	0.7807	SAP	IFX	0.7225	EOAN	DBK
17	0.8701	SIE	MAN	0.7799	LHA	DBK	0.7156	BAS	ADS
18	0.8697	DBK	ALV	0.7771	TKA	DB1	0.7154	BEI	BAS
19	0.8681	RWE	MUV2	0.7768	SDF	IFX	0.7103	RWE	IFX
20	0.8681	LIN	DPW	0.7766	SIE	BAS	0.7099	DTE	BAYN
21	0.8678	IFX	DPW	0.7757	VOW3	IFX	0.7098	RWE	MUV2
22	0.8674	VOW3	IFX	0.7756	IFX	DPW	0.7087	MUV2	BAS
23	0.8671	DPW	ADS	0.7704	IFX	BAS	0.7087	EOAN	ALV
24	0.8657	BAYN	ADS	0.7703	SAP	BEI	0.7074	SDF	BAS
25	0.8651	VOW3	MAN	0.7697	DBK	BAYN	0.7057	BAYN	BAS
26	0.8638	SDF	BAYN	0.7684	TKA	MAN	0.7039	IFX	DB1
27	0.8637	TKA	SZG	0.7668	EOAN	DPW	0.7034	IFX	ALV
28	0.8631	SIE	DPW	0.7649	DB1	CBK	0.6978	SIE	ALV
29	0.8622	SIE	ADS				0.6967	EOAN	CBK

Table 5

Results from PMFG Analysis for the second time period (7th May 2010 – 3rd August 2010). The column '4-Cliques' shows the observed number of 4-cliques in each PMFG and 'Max 4-Cliques' shows the total number of 4-Cliques possible for each PMFG. The last 4 columns show the number of 4-cliques that formed with stocks in 4 different sectors, in 3 different sectors etc.

	PMFG Analysis						
Dates	Stocks	4-Cliques	Max 4-Cliques	4 Sectors	3 Sectors	2 Sectors	1 Sector
7 May - 8 Jun 2010	30	24	27	7	15	2	0
21 May - 22 Jun 2010	29	23	26	7	12	4	0
4 Jun - 6 Jul 2010	29	16	26	7	7	2	0
18 Jun - 20 Jul 2010	29	26	26	12	13	1	0
2 Jul - 3 Aug 2010	30	27	27	7	17	3	0

Figure 1

The time series shows the quarter-on-quarter volume growth of GDP and expenditure components for Germany (shown in red) and the Euro Area (shown in black). [Data taken from ECB statistics[3]].

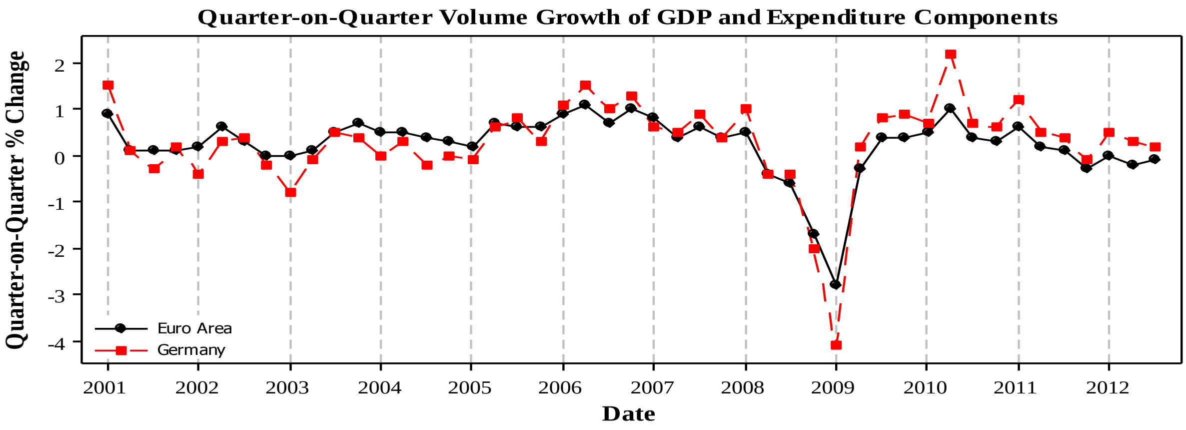
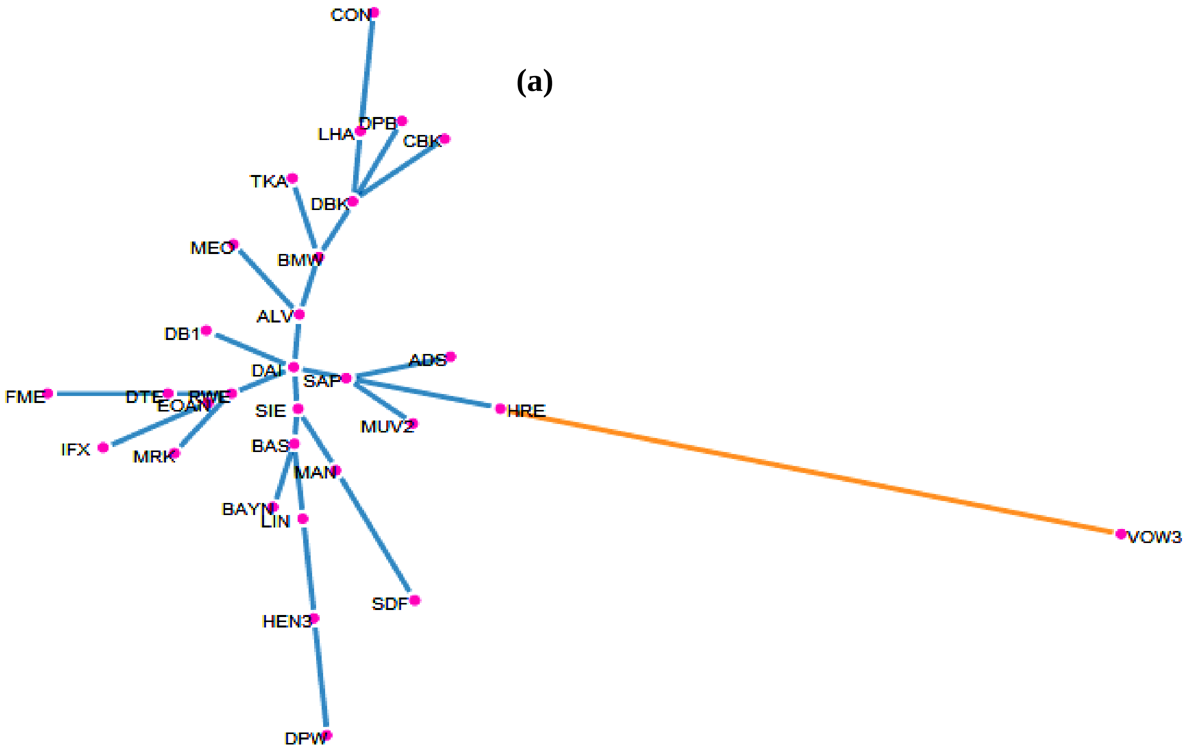


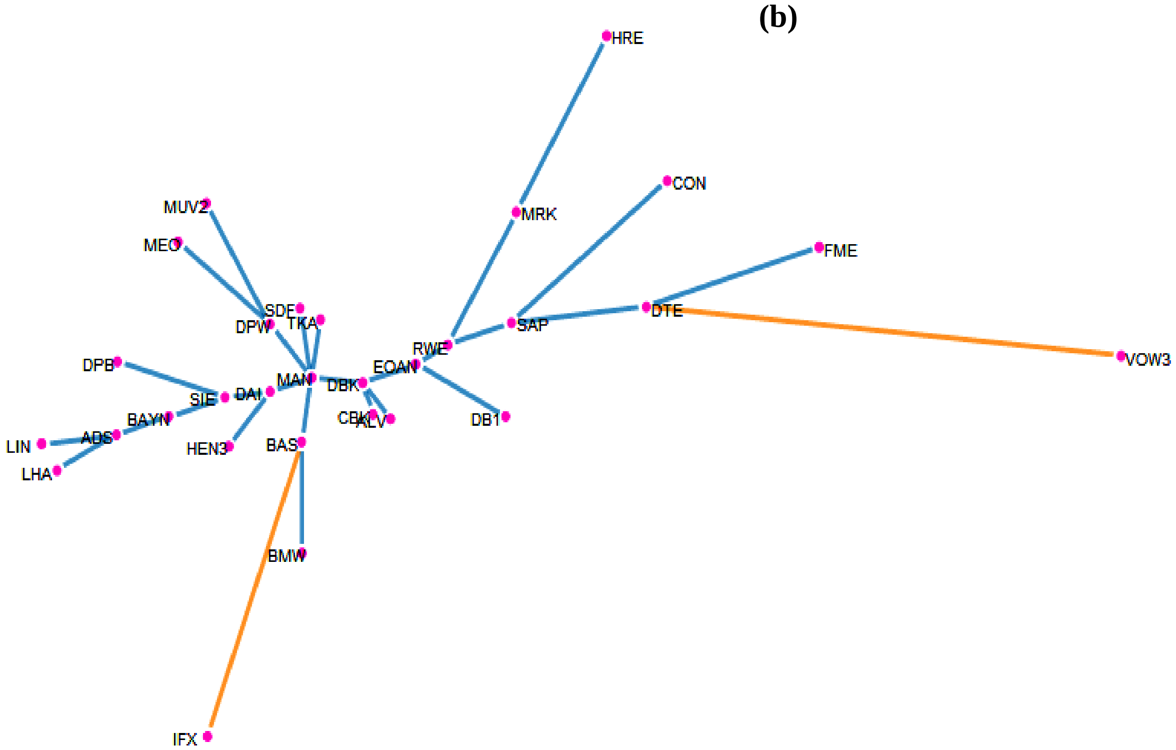
Figure 2

The MST for **2.a)** 7th Oct-6th Nov 2008; **2.b)** 4th Nov – 4th Dec 2008 and **2.c)** 2nd Dec-31st Dec 2008 where the vertices represent the various DAX30 companies, labelled using their stock symbols (please see Appendix A). The edge length is determined by the corr-distance so that shorter edges correspond to higher positive correlations. The edges identified as insignificant by the Bonferroni correction are the orange edges.



(a)

(b)


HRE
CON
MRK
FME
MUV2
MEO
SDF
TKA
DPW
SAP
DTE
VOW3
RWE
EOAN
DPB
MAN
DBK
SIE
DAI
BAYN
CBK
ALV
DB1
LIN
ADS
LHA
HEN3
BAS
BMW
IFX

(c)

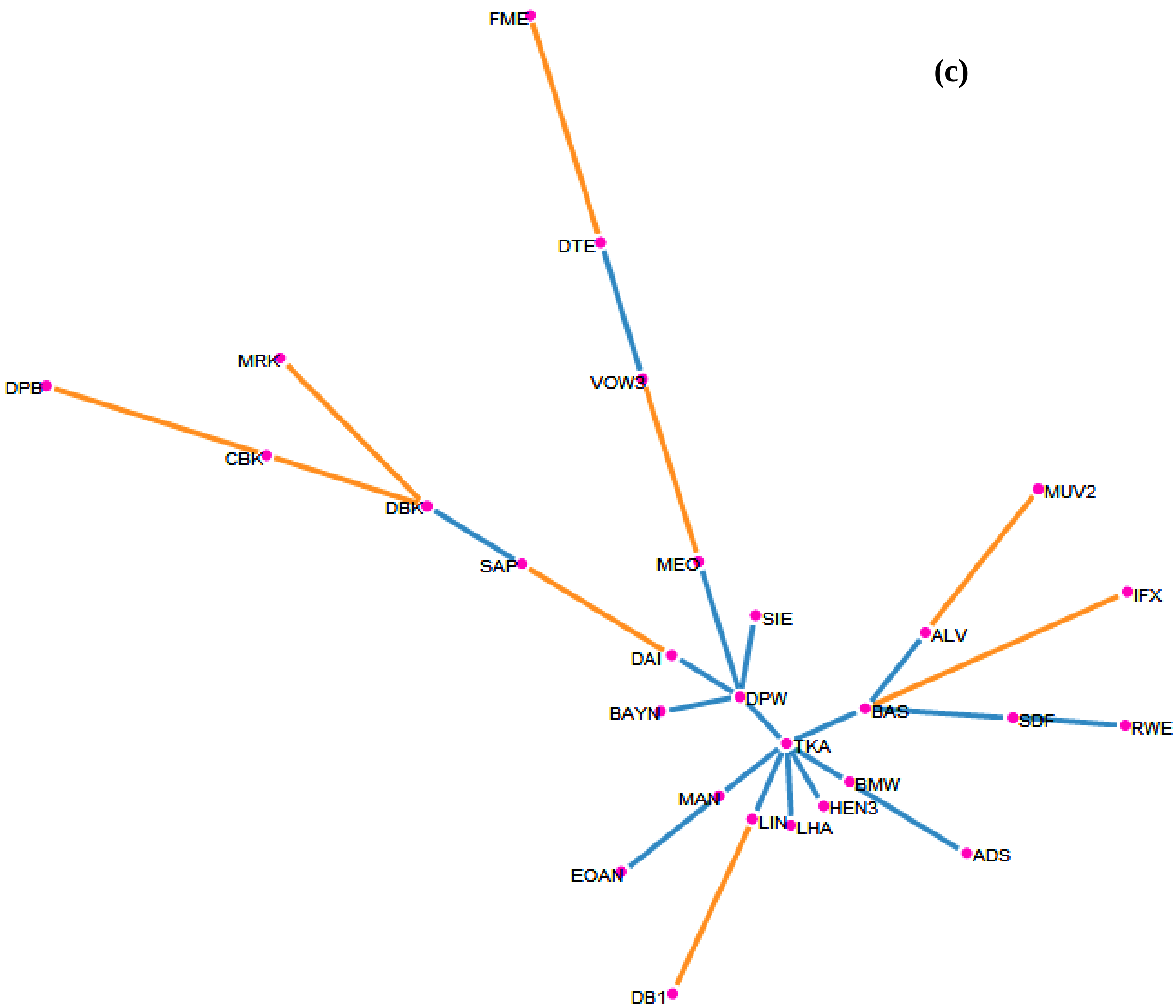
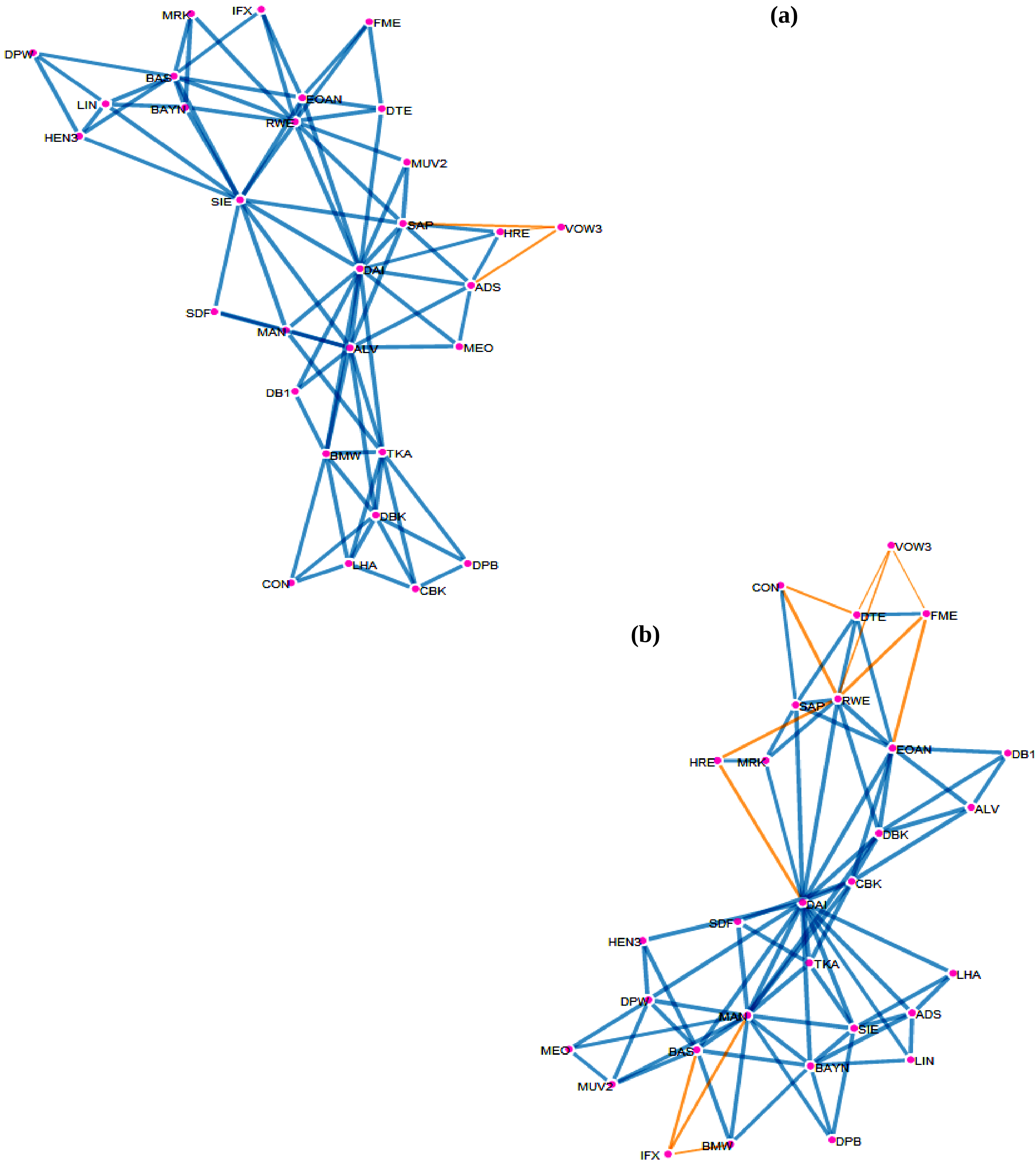
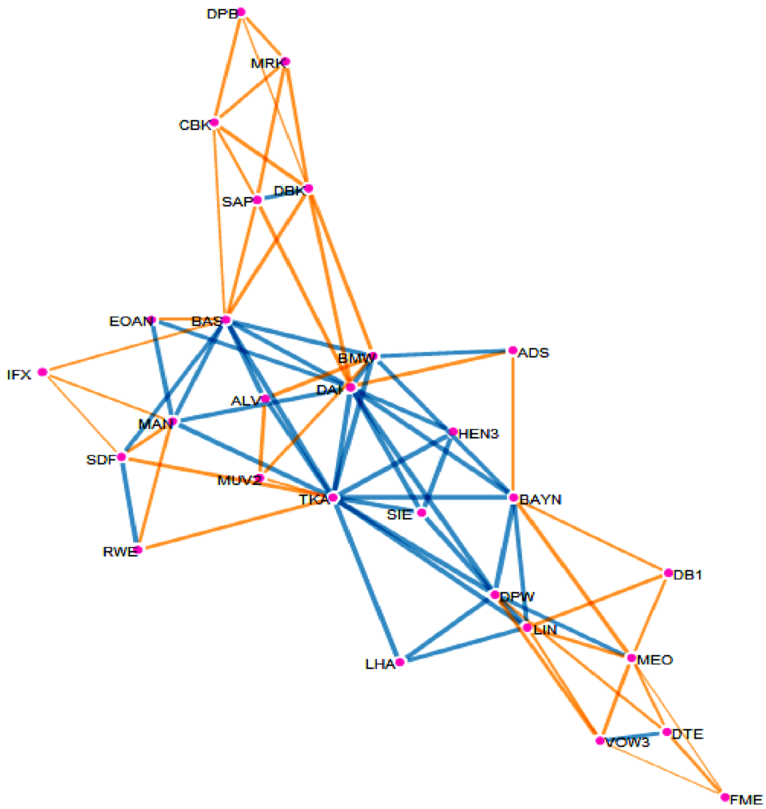

FME
DTE
MRK
DPB
VOW3
CBK
DBK
SAP
MEO
MUV2
SIE
IFX
ALV
DAI
DPW
BAYN
BAS
SDF
RWE
TKA
BMW
MAN
HEN3
LIN
LHA
ADS
EOAN
DB1

Figure 3

The PMFG for **3.a)** 7th Oct-6th Nov 2008; **3.b)** 4th Nov – 4th Dec 2008 and **3.c)** 2nd Dec-31st Dec 2008 where the vertices represent the various DAX30 companies, labelled using their stock symbols (please see Appendix A). Here the edge length does not relate to the correlation between the vertices. The edges identified as insignificant by the Bonferroni correction are the orange edges.



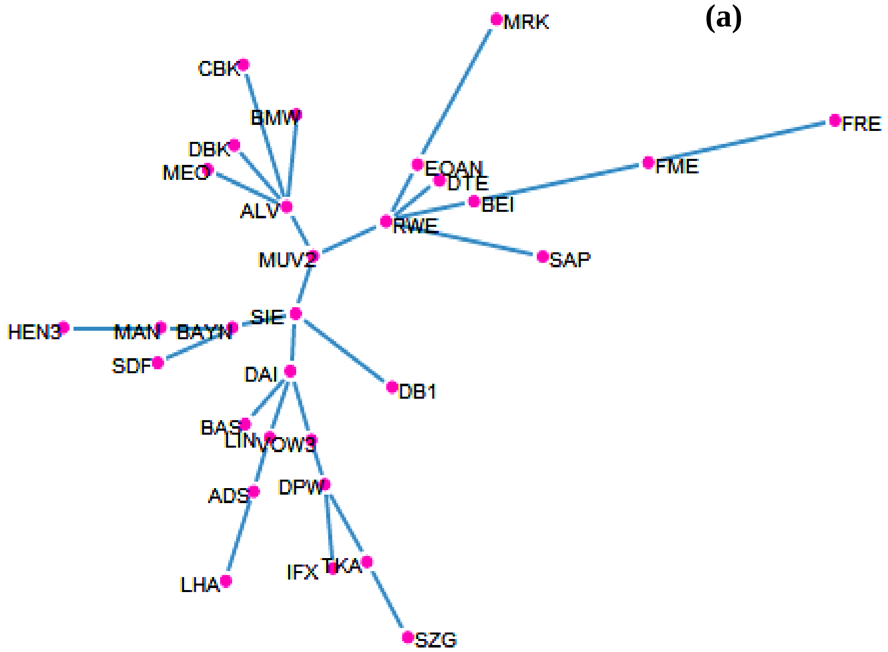

DPB
MRK
CBK
SAP
DBK
EOAN
BAS
BMW
ADS
IFX
DAI
ALV
MAN
HEN3
SDF
MUV2
TKA
BAYN
SIE
RWE
DB1
DPW
LIN
LHA
MEO
DTE
VOW3
FME

(c)

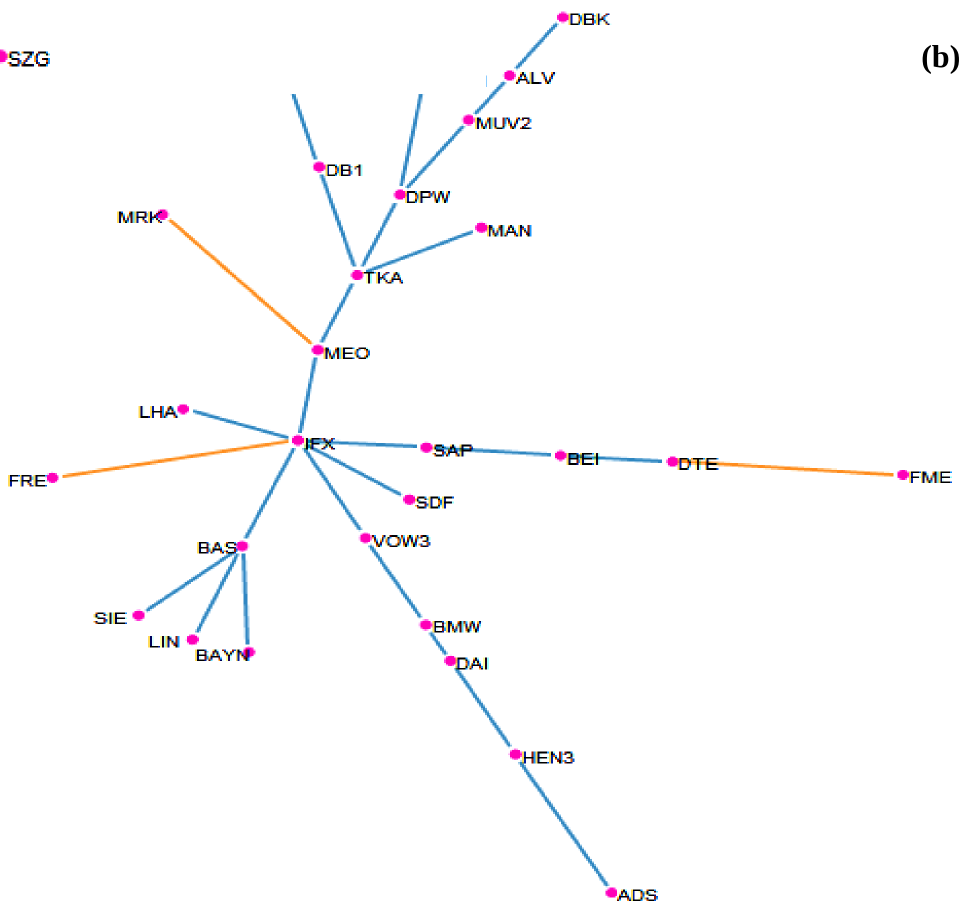
Figure 4

The MST for **4.a)** 7th May-8th Jun 2010; **4.b)** 4th Jun – 6th Jul 2010 and **4.c)** 2nd Jul-3rd Aug 2010 where the vertices represent the various DAX30 companies, labelled using their stock symbols (please see Appendix A). The edge length is determined by the corr-distance so that shorter edges correspond to higher positive correlations. The edges identified as insignificant by the Bonferroni correction are the orange edges.

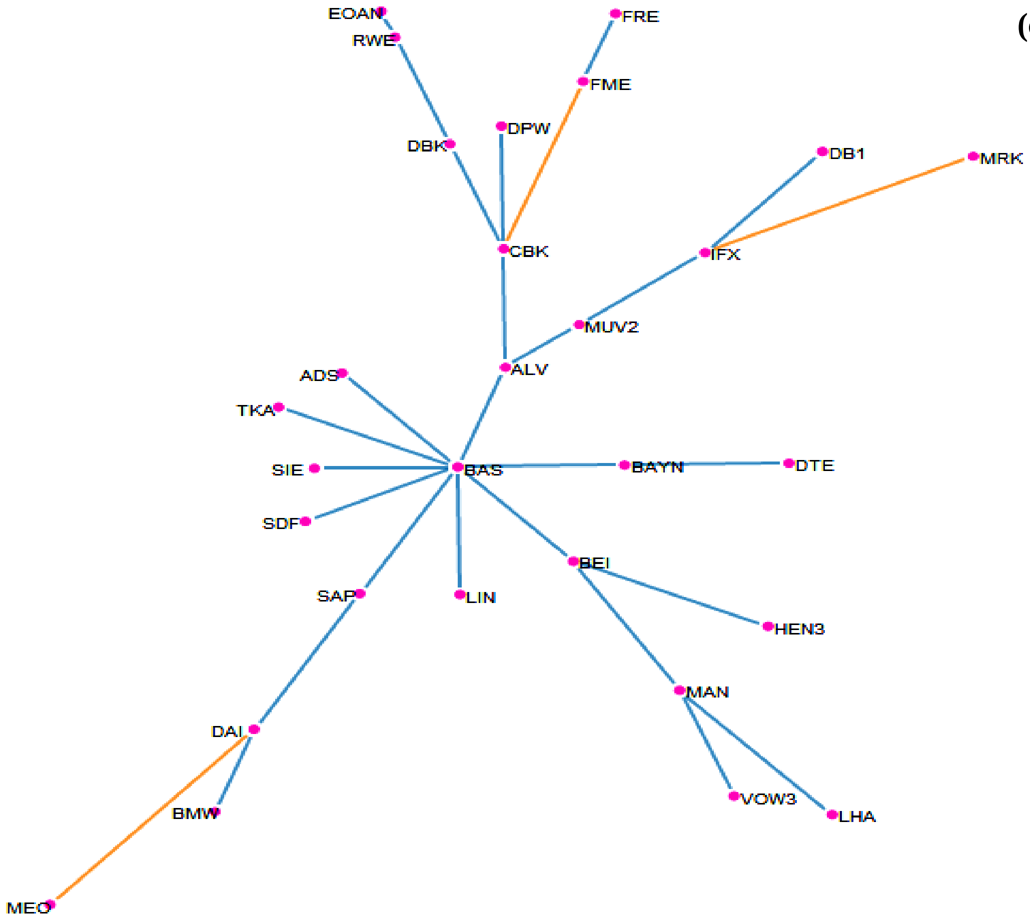
(a)



(b)



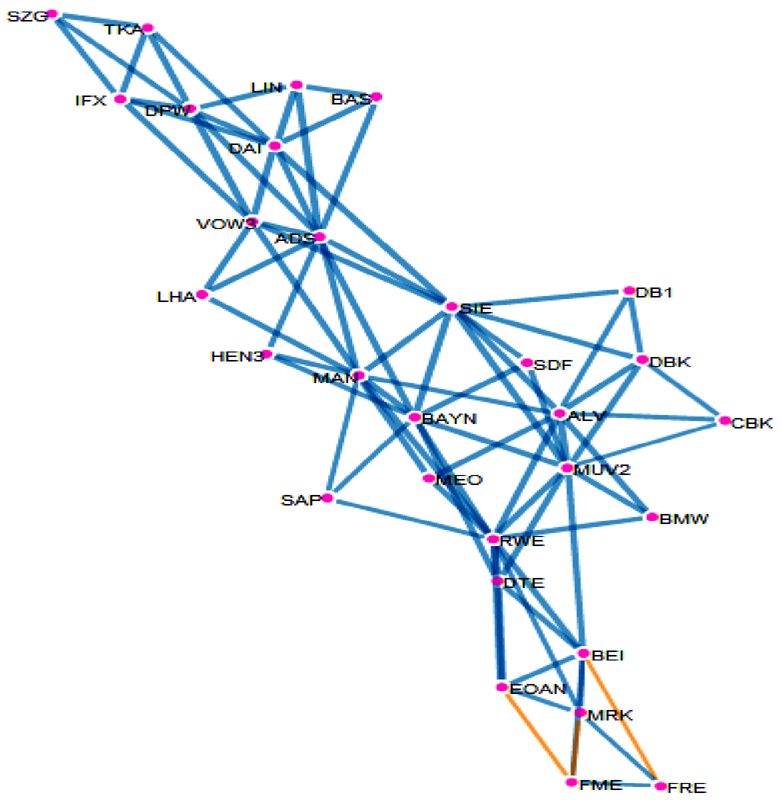
(c)



EOAN
RWE
FRE
FME
DPW
DBK
DB1
MRK
CBK
IFX
MUV2
ALV
ADS
TKA
SIE
BAS
BAYN
DTE
SDF
BEI
SAP
LIN
HEN3
MAN
DAI
BMW
VOW3
LHA
MEO

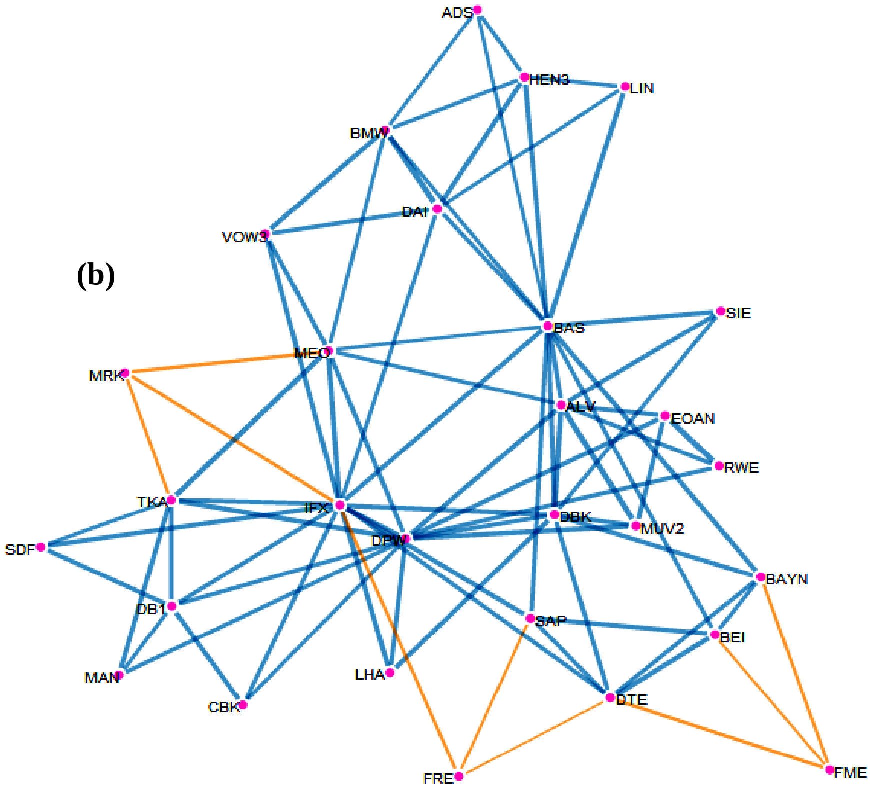
Figure 5

The PMFG for **5.a)** 7th May-8th Jun 2010; **5.b)** 4th Jun – 6th Jul 2010 and **5.c)** 2nd Jul-3rd Aug 2010 where the vertices represent the various DAX30 companies, labelled using their stock symbols (please see Appendix A). Here the edge length does not relate to the correlation between the vertices. The edges identified as insignificant by the Bonferroni correction are the orange edges.

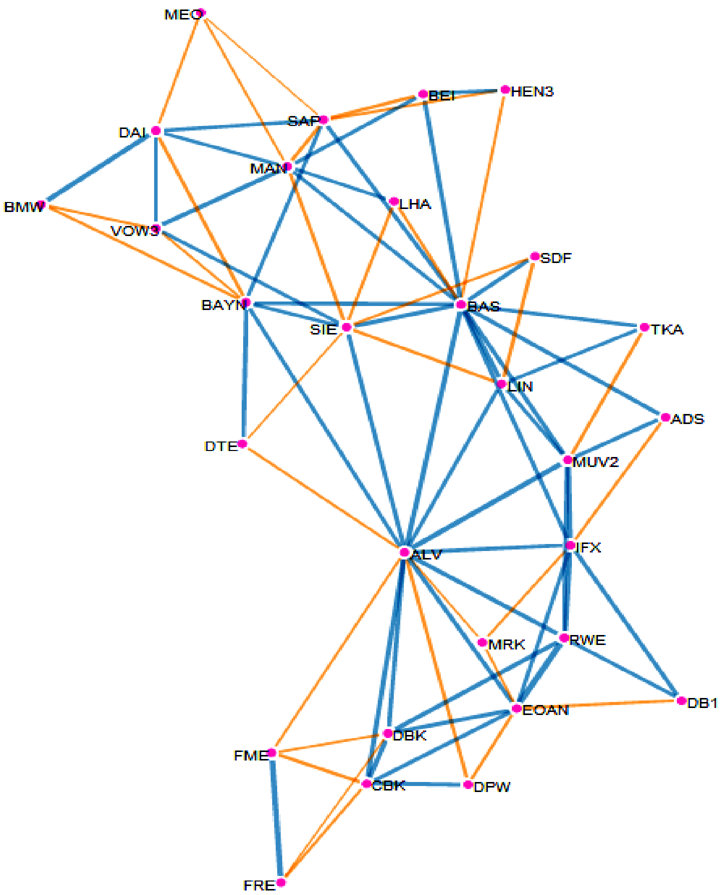


(a)

(b)


ADS
HEN3
LIN
BMW
DAI
VOW3
BAS
SIE
MEO
MRK
ALV
EOAN
RWE
TKA
IFX
DBK
MUV2
SDF
DPW
BAYN
DB1
SAP
BEI
MAN
LHA
CBK
DTE
FME
FRE

(c)


MEO
BEI
HEN3
DAI
SAP
MAN
BMW
LHA
VOW3
SDF
BAYN
BAS
SIE
TKA
LIN
ADS
DTE
MUV2
ALV
IFX
MRK
RWE
DB1
EOAN
DBK
FME
CBK
DPW
FRE