Equities, Credits and Volatilities: A Multivariate Analysis of the European Market During the Sub-prime Crisis

Irene Schreiber, Gernot Müller, Claudia Klüppelberg, Niklas Wagner
Technische Universität München und Universität Passau

Version: 2009-10-07

Abstract

Motivated by recent developments in light of the sub-prime and subsequent financial crisis we fit two different vector autoregressive generalized conditional heteroscedastic (VAR-GARCH) models to three financial indices with the aim of understanding the development of dependency structures between credit spreads and other macroeconomic variables. Our analysis includes daily quotes from June 2004 to April 2009 of the iTraxx Europe index, the Dow Jones Euro Stoxx 50 index, and the Dow Jones VStoxx index. We propose a robust, time-varying modeling approach concerning the conditional mean, and a BEKK versus DCC-GARCH approach concerning the conditional covariance. Furthermore we allow for a parsimonious model specification by setting insignificant coefficients to zero. Our empirical results indicate that the autoregressive coefficients vary strongly with time and even change their signs. Well-known interrelations, such as the negative correlation between CDS' and stocks are lost through the financial crisis. The conditional covariance estimates in the BEKK and DCC model are fairly similar, given the difference in the number of model parameters. We found evidence of strongly varying conditional variances and correlations, with dependencies increasing after the outbreak of the financial crisis. This knowledge may help to improve decision tools in the financial industry, especially in areas such as asset pricing, portfolio selection, and risk management.

Keywords: credit risk, credit default swaps, iTraxx index, vector autoregression, multivariate GARCH, VAR-GARCH, BEKK, DCC.

1 Introduction

Recent developments in light of the U.S. sub-prime and the subsequent financial crisis in 2008 have shown, that credit derivatives may be more closely related to other asset classes than previously suspected. This has made the understanding of dependency structures a key issue in business-oriented finance and subject to extensive research. In particular, many authors have studied the relationship between financial derivatives and other asset classes, such as equities and bonds, in order to understand the contamination of the global capital markets by the preceding US sub-prime crisis. For this purpose, credit default swaps (CDS), which are understood to be at the core of the sub-prime crisis, have been extensively investigated. Examples of empirical studies of CDS' and their relationship with other macroeconomic variables include Houweling and Vorst (2005), Byström (2005, 2006), Alexander and Kaeck (2006), Sougné et al. (2008), Ericsson et al. (2009) and Norden and Weber (2009). However, these authors concentrate on modeling the conditional mean, e.g. by fitting regression type models, whereas our modeling approach goes one step further by including the conditional covariance structure in the model. More precicely, this paper presents an investigation into the dependency structure of equities, credit spreads and volatilities in the European market by means of the three time series iTraxx Europe, Euro Stoxx 50, and VStoxx index. We fit a VAR-BEKK and a VAR-DCC model to our three dimensional time series of daily quotes from June 2004 to April 2009. Main innovations of our approach are the robust time-varying multivariate modeling approach as well as the combination of two different models, namely the vector autoregressive (VAR) and the multivariate GARCH (MGARCH) model.

It is worth mentioning that the starting point of our analysis was the data set of the iTraxx Europe family itself, and the attempt to model dependencies between different classes of aggregated credit spreads, e.g. between the iTraxx Senior Financials and Sub Financials. This question could be addressed adequately by a co-integration type modeling approach. However, the sub-indices of the iTraxx Family are far less liquid than the benchmark index iTraxx Europe. Therefore, due to long passages of missing values in the time series as a consequence of the financial crisis and the subsequent drying-out of the CDS market, the data set is unsuitable for an empirical study with this purpose.

In our paper we focus on modeling dependencies between aggregated credit spreads by means of the iTraxx Europe index and other asset classes. In a first, preceding study, in addition to stocks and stock market volatility, we also included long and short term interest rates by means of the LIBOR three months interest rate and a European government bond index by Bloomberg in our analysis, i.e. in a multivariate VAR-GARCH model with five time series. However, the interest rates were found to deliver no additional explanatory information considering the conditional mean structure. Moreover, with these five time series, the conditional covariance structure was not able to be captured by the different MGARCH models.

The first part of our modeling framework is based on VAR models, which have proven very useful to capture the evolution and the interdependencies of multiple time series. Introduced by

Sims (1972, 1980), they have been used for a variety of purposes such as data description and forecasting, as well as structural inference and policy analysis. The theoretical background on VAR models has been extensively explored and discussed in the literature, see e.g. Hannan (1970), Brockwell and Davis (1991), Lütkepohl (1991, 2005) and Hamilton (1994). For our purpose we extend the classical VAR modeling approach by admitting time-varying coefficients and by following a robust iteratively re-weighted least squares approach (see Huber, 1981), thereby reducing the influence of outliers and enhancing the model adequacy.

When investigating financial time series, it is a well known fact that models based on the homoscedasticity assumption are not sufficient to grasp stylized features such as volatility clustering and time-varying correlations. The second part of our modeling approach therefore is based MGARCH models, which in this context provide an ideal setting for investigating dependency structures of different financial time series. The widespread use of MGARCH models as an extension of the univariate ARCH and GARCH models by Engle (1982) and Bollerslev (1986) dates back to the well known VEC or so called general MGARCH model, first proposed by Bollerslev et al. (1988). From this starting point the theoretical development branches out, and as a result a variety of different MGARCH models evolved over the past two decades (see Li et al., 2002 and Bauwens et al., 2006, for recent reviews). As opposed to the univariate case, a coherent theory valid for all MGARCH models has yet to be developed. To cover all models therefore goes beyond the scope of this paper. In contrast, we deliberately focus on the two most well known and frequently used models in practice, namely the BEKK and the DCC model, which will be discussed here.

Engle and Kroner (1995) introduce the famous BEKK model as a special case of the VEC model by Bollerslev et al. (1988), which entailed the development of various subclasses and similar modeling approaches. The main advantage of this model is, that the BEKK parametrization automatically guarantees the positive definiteness of H_t. Besides that, the number of parameters in comparison with the general VEC model is remarkably reduced.

The key idea behind another class of MGARCH models is the nonlinear combination of univariate GARCH models, thus enabling separate modeling of variances and correlations. Probably the most popular model in practice is the well known dynamic conditional correlation (DCC) model by Engle and Sheppard (2001), which was introduced as a generalization of the constant conditional correlation (CCC) model by Bollerslev (1990). One reason why the DCC model is very popular with practitioners is its parsimony, as the model enables keeping the number of parameters relatively low (in comparison with both the VEC and BEKK model). Another advantage of this model lies in its flexibility, i.e. the univariate GARCH equations for the conditional variances may be specified by any kind of univariate GARCH parametrization, thereby including special model classes such as nonlinear or exponential GARCH models (see Engle and Sheppard, 2001, 2002).

The remainder of this paper is organized as follows: in Section 2 we briefly introduce the necessary modeling framework. Section 3 is dedicated to a multivariate empirical study of the iTraxx, Euro Stoxx 50 and VStoxx. Section 4 then concludes this paper with a brief summary of the main steps and most important findings on the topic.

2 Model Specification and Estimation

2.1 Model

In this section we briefly introduce the necessary theory which we will apply to our data in Section 3. All stochastic objects in this paper are defined on the probability space $(\Omega, \mathfrak{F}, P)$. Consider a vector stochastic process $(y_t)_{t\in\mathbb{Z}}$, i.e. $y_t : \Omega \to \mathbb{R}^N$. As usual, we condition on the sigma field, denoted by $\mathfrak{F}_{t-1}$, generated by the past information until time $t-1$. Note that we will follow the convention of using lowercase letters to denote either a random variable or its realization as a time series. In this paper we will consider the following vector autoregressive generalized conditional heteroscedastic (VAR-GARCH) model:

$$y_t = c + \Phi_1 y_{t-1} + \cdots + \Phi_p y_{t-p} + \varepsilon_t, \tag{1}$$

$$\varepsilon_t = H_t^{1/2}(\theta) z_t, \quad z_t \sim WN(0, I_N)\ i.i.d., \tag{2}$$

where $c \in \mathbb{R}^N$ denotes a vector of constants, $\Phi_1, \ldots, \Phi_p \in \mathbb{R}^{N\times N}$ are matrices of autoregressive coefficients and $\theta \in \Theta$ contains all GARCH parameters. Furthermore, $(z_t)_{t\in\mathbb{Z}}$ is a multivariate white noise process, $I_N \in \mathbb{R}^{N\times N}$ as usual is the identity matrix and $H_t^{1/2}(\theta) \in \mathbb{R}^{N\times N}$ is a positive definite matrix, such that H_t is the conditional covariance matrix of y_t, e.g. $H_t^{1/2}$ may be obtained by the Cholesky factorization of H_t.

The conditional mean part of the model in (1) is given by a VAR model of order p, while the conditional covariance matrix H_t in (2) is specified by an MGARCH model. As mentioned above, in this work we focus on the two most prominent models, the BEKK and the DCC model, which will be briefly discussed here.

Assume $(y_t), (\varepsilon_t)$ and (z_t) to be vector stochastic processes as given by (1) and (2). The BEKK(p,q)[1] model by Engle and Kroner (1995) for the conditional covariance matrix $H_t \in \mathbb{R}^{N\times N}$ is defined as

$$H_t = C'C + \sum_{i=1}^{q} A_i' \varepsilon_{t-i}\varepsilon_{t-i}' A_i + \sum_{j=1}^{p} B_j' H_{t-j} B_j, \tag{3}$$

where $A_i, B_j, \in \mathbb{R}^{N\times N}$ are parameter matrices and $C \in \mathbb{R}^{N\times N}$ is an upper triangular matrix. As mentioned above the main advantage of this model is, that the BEKK parametrization automatically guarantees the positive definiteness of H_t. The number of parameters in the BEKK model is $N(N+1)/2 + N^2(p+q)$, i.e. $\mathcal{O}(N^2)$. In order to reduce the number of parameters, different simplifications of the model evolved, e. g. the diagonal BEKK model where A_i and B_j in (3) are diagonal matrices or the scalar BEKK model where A_i and B_j are each replaced

[1] The acronym BEKK stands for Baba, Engle, Kraft & Kroner who wrote an earlier version of the paper by Engle and Kroner (1995) (see Engle, Kroner, Baba, and Kraft, 1993).

by a scalar times a matrix of ones. To ensure the uniqueness of the parametrization, certain restrictions have to be imposed on the coefficient matrices. For instance, in the special case of the BEKK(1,1) model with $H_t = C'C + A'\varepsilon_{t-1}\varepsilon'_{t-1}A + B'H_{t-1}B$ and parameter matrices $A = (A_{ij})_{i,j=1}^N, B = (B_{ij})_{i,j=1}^N$ Engle and Kroner (1995) show, that uniqueness is achieved by requesting all diagonal elements of C to be positive, as well as $A_{11}, B_{11} > 0$. These conditions for the coefficient matrices can be extended to the general case when $p, q > 1$. Engle and Kroner (1995) also show, that the BEKK model as defined in (1), (2) and (3) is stationary if and only if all eigenvalues of the matrix $\sum_{i=1}^q A'_i \otimes A_i + \sum_{j=1}^p B'_j \otimes B_j$ are less than one in modulus.

The second model for the conditional covariance matrix in (2) which we will consider in this paper is the dynamic conditional correlation (DCC) model by Engle (2002). The key idea of this model is to specify the conditional covariance matrix H_t in two steps. First, a univariate GARCH model is chosen for each individual conditional variance $H_{ii,t}, i = 1, \dots, N$. Second, based on the individual conditional variances the conditional correlation matrix is specified, thereby imposing its positive definiteness. The DCC(p,q) model for the conditional covariance matrix $H_t \in \mathbb{R}^{N \times N}$ is defined as

$$H_t = D_t R_t D_t, \quad \text{with} \quad D_t = diag(H_{11,t}^{1/2}, \dots, H_{NN,t}^{1/2}). \tag{4}$$

The elements of D_t are defined as univariate GARCH models, i.e. $\forall\ i = 1, \dots, N$ we define

$$H_{ii,t} = \omega_i + \sum_{q=1}^{q_i} \alpha_{iq}\varepsilon_{i,t-q}^2 + \sum_{p=1}^{p_i} \beta_{ip} H_{ii,t-p}, \tag{5}$$

with the usual restrictions for non-negativity and stationarity being imposed $\forall\ i = 1, \dots, N$:

(i) $\omega_i > 0$,

(ii) $\forall\ p = 1, \dots, p_i, \quad \forall\ q = 1, \dots, q_i : \alpha_{iq}, \beta_{ip}$ are such that $H_{ii,t}$ will be positive with probability one,

(iii) $\sum_{q=1}^{q_i} \alpha_{iq} + \sum_{p=1}^{p_i} \beta_{ip} < 1$.

The dynamic correlation structure is given by

$$Q_t = (1 - \sum_{m=1}^{M} a_m - \sum_{n=1}^{N} b_n)\overline{Q} + \sum_{m=1}^{M} a_m(\nu_{t-m}\nu'_{t-m}) + \sum_{n=1}^{N} b_n Q_{t-n}, \tag{6}$$

$$R_t = \mathrm{diag}(Q_{11,t}^{1/2}, \dots, Q_{NN,t}^{1/2})^{-1} Q_t \,\mathrm{diag}(Q_{11,t}^{1/2}, \dots, Q_{NN,t}^{1/2})^{-1}, \tag{7}$$

with $\nu_t := D_t^{-1}\varepsilon_t$, and $\overline{Q}$ being the unconditional covariance matrix of $\nu_t : \nu_t \sim \mathcal{N}(0, \overline{Q})$. We refer to Engle and Sheppard (2001, Proposition 2) for sufficient conditions regarding the positive definiteness of H_t.

2.2 Estimation Method

Regarding the estimation of the model parameters in (1), (2) and (3), (4)–(5), respectively, we follow a two step approach, where in the first step the parameters of the VAR model, and in the second step the GARCH parameters are estimated.

Concerning the VAR model coefficients we pursue the robust iteratively re-weighted least squares (RLS) approach of Huber (1981). Assume that the sample size is T and that we are given p pre-sample values $y_{-p+1}, \ldots, y_0$. We define:

$$\begin{aligned} &Y := (y_1, \ldots, y_T) \in \mathbb{R}^{N\times T}, \\ &\Pi := (c, \Phi_1, \ldots, \Phi_p) \in \mathbb{R}^{N\times(Np+1)}, \quad \pi := \mathrm{vec}(\Pi) \in \mathbb{R}^{N(Np+1)}, \\ &x_t := (1, y'_{t-1}, \ldots, y'_{t-p})' \in \mathbb{R}^{Np+1}, \quad X := (x_1, \ldots, x_T) \in \mathbb{R}^{(Np+1)\times T}, \\ &E := (\varepsilon_1, \ldots, \varepsilon_T) \in \mathbb{R}^{N\times T}, \end{aligned}$$

where $\mathrm{vec}(\cdot)$ is the column stacking operator that stacks the columns of a $m \times n$ matrix as a vector of dimension mn. Using this notation we may then rewrite (1) as a linear model

$$Y = \Pi X + E, \quad \text{or equivalently} \quad \mathrm{vec}(Y) = (X' \otimes I_N)\pi + \mathrm{vec}(E), \tag{8}$$

where $\otimes$ denotes the Kronecker product or direct product of two matrices. Note that as opposed to classical linear modeling, the matrix of covariables contains lagged dependent variables.
The unknown parameters of the VAR model contained in π in (8) are then estimated by the RLS approach of Huber (1981), who introduces the class of M-estimates, in order to reduce the influence of outliers and achieve distributional robustness. We investigate the problem $\sum_{i=1}^{NT} \rho(x_i\,;\,\pi) = \min!$, or equivalently $\sum_{i=1}^{NT} \psi(x_i\,;\,\pi) = \sum_{i=1}^{NT} w_i x_i = 0$, where x_i is the i-th residual of the NT-dimensional linear model in (8), $\rho(x\,;\,\pi)$ is a weighting function, $\psi(x) := (\partial/\partial\theta)\rho(x\,;\,\pi)$ and $w_i := \psi(x_i\,;\,\pi)/x_i$. The weighting function $\rho(x\,;\,\pi)$ is assumed to be twice continuously differentiable in x almost everywhere, with nonnegative second derivative wherever defined. Huber (1981) proposes

$$\rho(x) = \begin{cases} x^2/2 & : \quad |x| \le c, \\ c|x| - c^2/2 & : \quad |x| > c, \end{cases} \tag{9}$$

which implies weights $w_i = 1$ if $|x_i| \le c$ and $w_i = c/x_i$ if $|x_i| > c$. In the context of VAR models strong consistency of the RLS estimator is e.g. shown by Campbell (1982), while asymptotic normality is derived by Li and Hui (1989).

We now briefly discuss estimation procedures for the BEKK and DCC-GARCH model. In the case of the BEKK model given by (2) and (3) we perform maximum likelihood (ML) estimation.

Assume we have a given sample size of $t = 1, \ldots, T$. The log likelihood function is then given by

$$L(\theta) = -\frac{1}{2}\sum_{t=1}^{T}(N\ln(2\pi) + \ln|H_t(\theta)| + \varepsilon_t' H_t(\theta)^{-1}\varepsilon_t), \tag{10}$$

where $\theta := \text{vec}(C, A_1, \ldots, A_q, B_1, \ldots, B_p) \in \Theta \subset \mathbb{R}^{N(N+1)/2+N^2(p+q)}$ contains all unknown GARCH parameters. The likelihood function is maximized with respect to θ by using numerical methods. A closed form solution does not necessarily exist, due to the nonlinearity of the likelihood function. For asymptotic properties of the ML estimator, see e.g. Comte and Lieberman (2003), who derive strong consistency and asymptotic normality.

According to Engle and Sheppard (2001), the DCC model as defined in (4)–(7) was designed to allow for a two-stage estimation procedure. They suggest decomposing the parameter vector θ into two disjoint parts, one for the individual conditional volatilities and one for the conditional correlations. Then in the first stage univariate GARCH models for each component of $\varepsilon_t = (\varepsilon_{1t}, \ldots, \varepsilon_{Nt})$ are estimated. In the second stage, using transformed residuals resulting from the first stage, an estimator for the conditional correlations is derived. As $H_t = D_t R_t D_t$ in the DCC model according to (4), the likelihood function in (10) may be rewritten in the following way:

$$\begin{aligned} L(\theta) &= -\frac{1}{2}\sum_{t=1}^{T}(N\ln(2\pi) + \ln|D_t R_t D_t| + \varepsilon_t' D_t^{-1} R_t^{-1} D_t^{-1}\varepsilon_t) \\ &= -\frac{1}{2}\sum_{t=1}^{T}(N\ln(2\pi) + 2\ln|D_t| + \varepsilon_t' D_t^{-1} D_t^{-1}\varepsilon_t - \nu_t'\nu_t + \ln|R_t| + \nu_t' R_t^{-1}\nu_t). \end{aligned} \tag{11}$$

Let $\theta = (\theta_1, \theta_2)$ denote the parameters for the conditional volatilities and conditional correlations, as given in (4)–(5) and (6)–(7), respectively. The likelihood in (11) is decomposed into two disjoint parts:

$$L(\theta) = L(\theta_1, \theta_2) = L_V(\theta_1) + L_C(\theta_2),$$

with a volatility part $L_V(\theta_1) := -\frac{1}{2}\sum_{t=1}^{T}(N\ln(2\pi) + 2\ln|D_t| + \varepsilon_t' D_t^{-1} D_t^{-1}\varepsilon_t)$ and a correlation part $L_C(\theta_2) := -\frac{1}{2}\sum_{t=1}^{T}(-\nu_t'\nu_t + \ln|R_t| + \nu_t' R_t^{-1}\nu_t)$. The volatility part then corresponds to the sum of the likelihood functions of N univariate GARCH models

$$L_V(\theta_1) := -\frac{1}{2}\sum_{t=1}^{T}\sum_{i=1}^{N}\left(\ln(2\pi) + \ln(H_{ii,t}) + \frac{\varepsilon_{it}^2}{H_{ii,t}}\right).$$

Now first solve $\hat{\theta}_1 = \arg\max L_V(\theta_1)$, and then subsequently $\hat{\theta}_2 = \arg\max L_C(\hat{\theta}_1, \theta_2)$. Note that Engle and Sheppard (2001) argue, that consistency and asymptotic normality of $\hat{\theta}_1$ and $\hat{\theta}_2$ hold due to results given by Newey and McFadden (1994) concerning consistency of an estimator in a two-step general method of moments problem, usually resulting in a loss of efficiency. However, this argumentation is recently being questioned by e.g. Caporin and McAleer (2008, 2009), who conclude that the properties of the DCC estimates as claimed by Engle and Sheppard (2001) cannot be derived by their course of argumentation.

3 Empirical Analysis

3.1 Data Set

Our data set consists of three time series, the Dow Jones Euro Stoxx 50 index, the Dow Jones VStoxx index, a volatility index based on options on the Euro Stoxx 50 and the CDS index iTraxx Europe. In our analysis we focus on the iTraxx Europe benchmark index with a maturity of five years, as this is the most liquid index within the iTraxx Europe index family. As the membership of the iTraxx is adjusted every six months by issuing a new index series, we construct a time series that contains the most recent series at any point in time. In this way we ensure that our analysis is always built on the most liquid names. The data period starts on June 23, 2004 and ends on April 30, 2009, i.e. the data set covers 1230 daily quotes for each of the time series. Basic characteristics of the data are summarized in Table 1. The data was transformed to

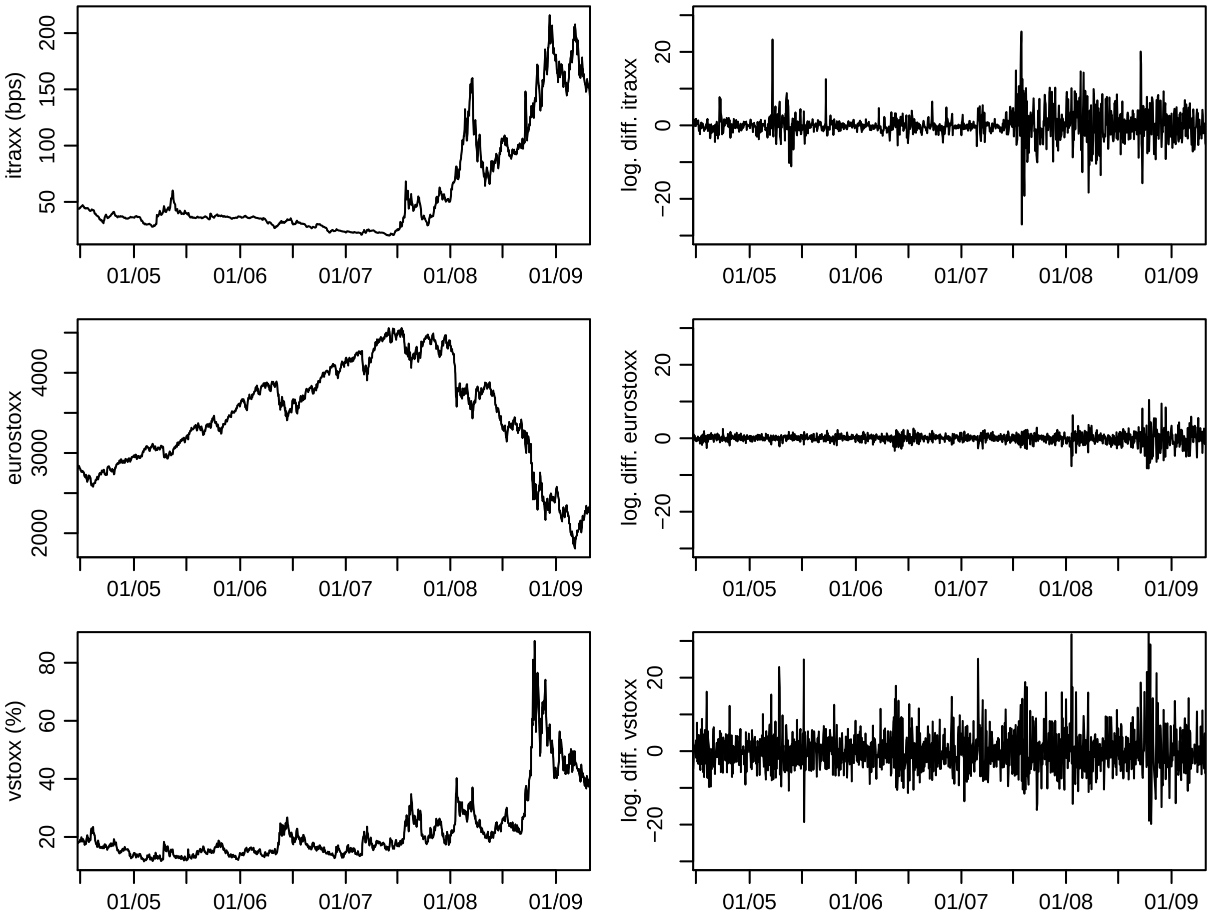

Figure 1: Daily quotes of the iTraxx Europe, the Euro Stoxx 50 and the VStoxx index between 2004-06-23 and 2009-04-30.

	2004-06-23 to 2009-04-30			2004-06-23 to 2007-08-15			2007-08-16 to 2009-04-30		
	itraxx	eurost.	vstoxx	itraxx	eurost.	vstoxx	itraxx	eurost.	vstoxx
min.	20.09	1809.98	11.60	20.09	2580.04	11.60	29.10	1809.98	17.24
1st qu.	31.00	2980.13	14.85	27.78	3055.85	14.03	68.28	2451.58	22.56
median	37.12	3524.58	17.72	35.20	3544.58	15.67	101.38	3429.58	27.32
mean	59.72	3463.51	22.13	33.38	3537.33	16.15	108.71	3326.18	33.26
3rd qu.	74.37	3987.13	24.00	37.19	3988.04	17.58	154.42	3881.09	42.05
max.	215.92	4557.57	87.51	68.20	4557.57	30.74	215.92	4489.79	87.51
std.	46.43	654.60	11.85	7.41	540.75	2.96	48.71	808.63	13.97
skewn.	1.57	-0.31	2.20	0.42	0.10	1.34	0.21	-0.24	1.18
kurt.	4.26	2.28	8.03	3.65	1.88	5.52	1.89	1.70	3.82

Table 1: Basic characteristics of the data set. Left: whole period, middle: first “tranquil” period, right: last “volatile” period.

logarithmic differences multiplied by one hundred (see Figure 1). Evidence of simple trends and seasonality was not found. Note, that on the whole iTraxx and Euro Stoxx show counter trends, whereas iTraxx and VStoxx indicate a positive interrelation. The three time series display typical stylized features such as volatility clustering and at least one structural break, which e.g. in mid 2007 is related to the rise of the sub-prime crisis. The characteristics of the time series, e.g. in terms of mean and volatility levels, change significantly before and after the outbreak of the crisis (see Table 1), thereby the biggest structural changes are visible in the iTraxx index. The estimated corresponding autocorrelation functions of the data and the squared data, as well as the corresponding cross correlations between the time series can be found in Figure 2. We see some autocorrelation in the time series, especially within the iTraxx at lag one, however the values are rather small. Cross correlations are perceivable only at lag zero. The autocorrelations and cross correlations in the squared data give rise to the hypothesis of stochastic volatility.

3.2 A VAR Model for the Conditional Mean

In a first step, in order to capture the weak autocorrelation in the data as seen in Figure 2, we model the conditional mean of the time series by fitting a VAR model as given by (1) to the data. In order to determine the model order p, we fit different models up to order $p = 10$ via ML estimation and calculate the associated information criteria AIC, HQ and SC (see Akaike 1973, 1974, Hannan and Quinn, 1979 and Schwarz, 1978). As displayed in Table 2, AIC suggests $p = 4$, whereas HQ and SC both recommend model order $p = 1$. We therefore fit a VAR(1) model to our data. When conducting ML estimation of the coefficient matrix Φ we find that six out of nine coefficients are insignificant. Precisely only the coefficients Φ_{11}, Φ_{21} and Φ_{33} are significant at a 90 % confidence level. We also observe, that the coefficient matrix contains mostly very small values. In order to gain deeper insight into the vector autoregressive structure of our data set we therefore conduct a rolling window analysis of the coefficient matrix. We use different windows from 25 to 300 days, finding that all coefficients vary over time, some very strongly

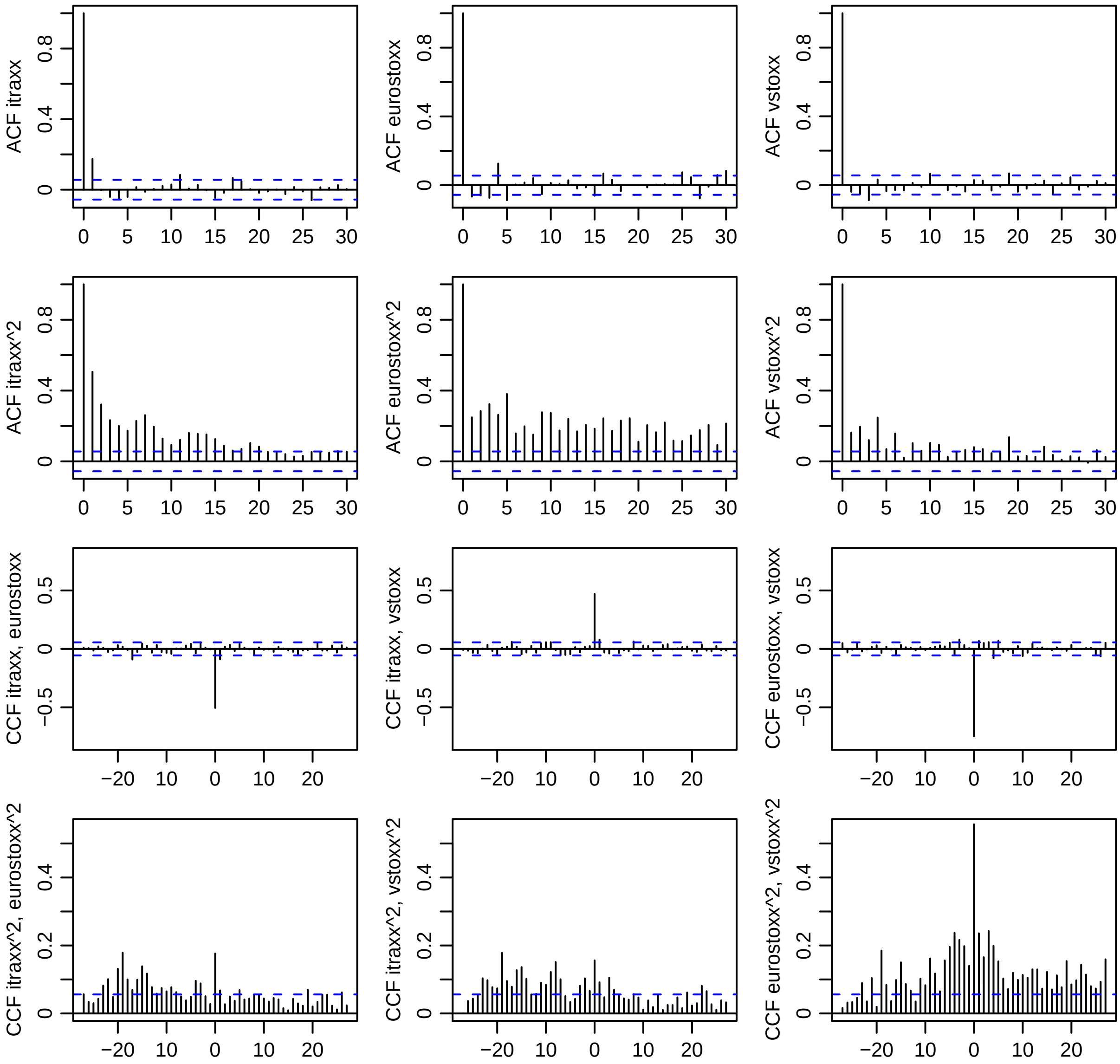

Figure 2: Autocorrelations and cross correlations of the data set with 95% confidence bounds.

and even changing their signs. This explains the large number of insignificant close to zero coefficients which we observed in the first place. Additionally we observe that the coefficients react sensitively to apparent outliers in the original data. For this reason we decide in favor of the RLS estimation procedure by Huber (1981), with the Huber weighting function as defined in (9).

We again conduct a rolling window analysis, this time estimating robustly (see Figure 3). In comparison with the ML estimates, the influence of outliers is remarkably reduced, however we find that both with ML and RLS estimation, the coefficients are time-varying and often change their signs. Consequently we will follow a robustly estimated VAR(1) approach with a time-varying coefficient matrix.

In the following we assess the question of which entries of the coefficient matrix in Figure 3

p	1	2	3	4	5	6	7	8	9	10
AIC(p)	5.735	5.737	5.735	5.723	5.730	5.731	5.736	5.736	5.741	5.746
HQ(p)	5.754	5.770	5.782	5.784	5.806	5.821	5.840	5.854	5.873	5.893
SC(p)	5.785	5.825	5.860	5.886	5.931	5.970	6.012	6.050	6.092	6.135

Table 2: Order selection criteria AIC, HQ and SC for our data set.

	weighted LS estimation								
	(2004-06-23 to 2009-04-30)			(2004-06-23 to 2007-08-15)			(2007-08-16 to 2009-04-30)		
	est.	std.error	t-stat.	est.	std.error	t-stat.	est.	std.error	t-stat.
c_1	0.077	0.112	0.688	0.041	0.151	0.272	0.165	0.092	1.793
c_2	-0.015	0.042	-0.357	0.055	0.041	1.341	-0.126	0.093	-1.355
c_3	0.052	0.151	0.344	0.034	0.161	0.211	0.022	0.199	0.111
Φ_{11}	0.190	0.022	8.832	0.141	0.019	7.495	0.309	0.058	5.303
Φ_{12}	-0.032	0.073	-0.435	-0.390	0.104	-3.755	0.064	0.169	0.380
Φ_{13}	0.021	0.019	1.109	0.001	0.018	0.038	-0.063	0.055	-1.143
Φ_{21}	-0.002	0.009	-0.208	0.014	0.010	1.362	-0.033	0.021	-1.566
Φ_{22}	-0.010	0.031	-0.324	-0.078	0.056	-1.404	-0.002	0.061	-0.027
Φ_{23}	0.006	0.008	0.767	-0.010	0.010	-1.096	0.030	0.020	1.501
Φ_{31}	0.084	0.044	1.899	-0.069	0.059	-1.175	0.194	0.072	2.689
Φ_{32}	0.008	0.151	0.053	0.461	0.324	1.421	0.045	0.209	0.217
Φ_{33}	-0.074	0.039	-1.924	-0.020	0.055	-0.359	-0.076	0.068	-1.109

Table 3: RLS estimation of the coefficients of the VAR(1) model for different time periods. Left: whole period, middle: "tranquil" period, right: "volatile" period.

may be set to zero and, as a consequence, will not be included in the further analysis. For this purpose, we simultaneously follow two criteria. For the first criterion we split the time series into two disjoint parts, namely the first 800 data points (2004-06-23 to 2007-08-15) and the last 430 data points (2007-08-16 to 2009-04-30). This partition splits our time series into a "tranquil" period preceding the sub-prime crisis, and a "volatile" period starting mid of 2007. Separately analyzing these two periods is self-evident given the apparent structural breaks in the original data (see Figure 1) and the coefficient matrix structure (see Figure 3). We then perform RLS estimation for each period separately. The results are displayed in Table 3. For the tranquil period, Φ_{11} and Φ_{12} are significant, whereas for the volatile period, Φ_{11} and Φ_{31} are significant on a 90% confidence level. As a first criterion for which coefficients to include in the analysis we follow the convention of admitting all coefficients that are significant on a 90% level at least in one of the two periods. According to this criterion Φ_{11}, Φ_{12}, Φ_{31}, Φ_{33} are included in our analysis. However, this criterion has the drawback that it excludes coefficients that are strongly time-varying and thus may only be significant for certain short time periods. As a second criterion for which coefficients to admit for the analysis we therefore decide to admit strongly time-varying coefficients, in addition to those being significant according to the first criterion. This means that

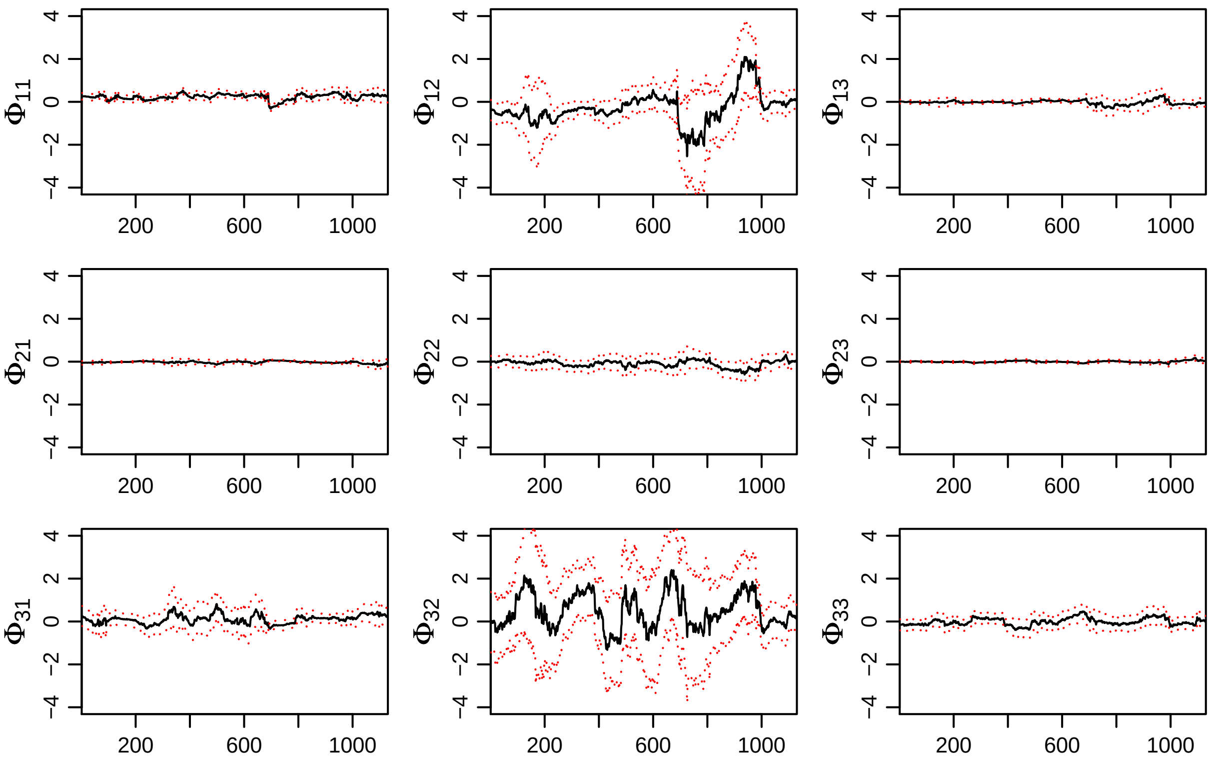

Figure 3: Robust weighted LS estimation of Φ, using 100 day rolling windows. 95% confidence bounds.

additionally Φ_{32} is included as well. Therefore, as a final model regarding the conditional mean, we propose a VAR(1) modeling approach with robustly estimated, time-varying coefficients and with $\Phi_{13} = \Phi_{21} = \Phi_{22} = \Phi_{23}$ set to zero. Due to the fact that we set some coefficients to zero, the five non-zero entries in the coefficient matrix have slightly different values from those displayed in Figure 3, yet the overall structure of the of the coefficients remains unchanged.

We find evidence of negative dependencies between the CDS spreads and the stock markets in terms of Φ_{12} during the tranquil period (see Table 3). Consequently, the stock market tends to lead the CDS market, our results being consistent with previous empirical studies, see e.g. (Byström, 2005, 2006) and Alexander and Kaeck (2006). However, from Figure 3 we observe, that the coefficients Φ_{12} and Φ_{32} vary strongly and even change their signs, while the other coefficients display less or no variations. We find, that Φ_{12} displays irregular variations, with several structural breaks which may be explained by credit events such as the downgrading of General Motors and Ford in May 2005 or the outbreak of the sub-prime crisis in July 2007. In contrast, Φ_{32} displays strong cyclical variations during the whole time period.

We now proceed with deriving the model residuals. The time-varying coefficient matrix Φ_t is estimated by the data points $t, t-1, \dots, t-99$. Following a forecasting perspective we set $\varepsilon_t = y_t - \Phi_t y_{t-1}, t = 101, \dots, T$, i.e. we have a new time series of residuals ε_t, with $t = 101, \dots, T$. In our case ($T = 1230$) we obtain a residual time series of 1130 data points. We find no evidence of remaining autocorrelation in the residuals. As this was the objective of our analysis so far, in this respect the model fit is very good. Cross correlations at lag zero are still perceivable, as they evidently cannot be captured by the VAR model. However, we still observe characteristic patterns and structural changes in the residuals, and thus the residual time series is obviously not generated by a white noise process. Furthermore, the autocorrelation and cross correlations plots of the squared residuals on the whole still resemble the ones in Figure 2, which emphasizes the need for an additional modeling of the covariance structure of our time series.

3.3 A BEKK Model for the Conditional Covariance

After fitting a VAR(1) model to our data we now proceed with the modeling of the conditional covariance structure. Portmanteau and Lagrange multiplier tests for potential ARCH effects (see e.g. Lütkepohl, 1991, 2005) in the residuals of the VAR model show strong evidence of ARCH effects and confirm the heteroscedasticity assumption. We therefore fit a BEKK-GARCH model as given by (3) to the residuals obtained from the VAR(1) model. The BEKK model is particularly compelling due to its parametrization that by definition guarantees the positive definiteness of the covariance matrix. Besides that, the number of parameters is notably reduced in comparison with the general MGARCH model. Furthermore, in comparison with other MGARCH

	BEKK(1,1)	BEKK(1,2)	BEKK(2,1)	BEKK(2,2)
AIC	6221.836	6189.417	6192.885	6187.527

Table 4: AIC criterion for the model order selection in the BEKK model.

	est.			std.error			t-stat.		
C	0.358	0.000	0.816	0.076	0.000	0.236	4.692	–	3.456
	0.000	0.252	-2.376	0.000	0.071	0.419	–	3.523	-5.671
	0.000	0.000	0.232	0.000	0.000	0.111	–	–	2.090
A_1	0.678	-0.033	0.215	0.043	0.011	0.061	15.741	-3.007	3.521
	0.000	0.133	0.000	0.000	0.028	0.000	–	4.766	–
	0.000	0.000	0.000	0.000	0.000	0.000	–	–	–
A_2	0.000	-0.015	0.157	0.000	0.009	0.052	–	-1.732	3.030
	0.000	0.381	0.000	0.000	0.034	0.000	–	11.055	–
	0.000	0.031	0.000	0.000	0.006	0.000	–	5.119	–
B_1	0.241	-0.026	0.000	0.073	0.012	0.000	3.294	-2.201	–
	0.686	-0.820	0.959	0.111	0.099	0.244	6.156	-8.307	3.927
	0.000	-0.082	0.000	0.000	0.014	0.000	–	-5.949	–
B_2	0.806	-0.079	0.271	0.028	0.017	0.077	28.887	-4.725	3.511
	0.971	-0.935	3.152	0.160	0.111	0.324	6.063	-8.410	9.719
	0.084	-0.183	1.110	0.021	0.018	0.071	4.003	-10.407	15.744
eigenvalues	14.817	3.003	2.895	0.874	0.781	0.341	0.105	0.092	0.021

Table 5: Coefficients and eigenvalues of the reduced BEKK(2,2) model. 28/28 significant parameters.

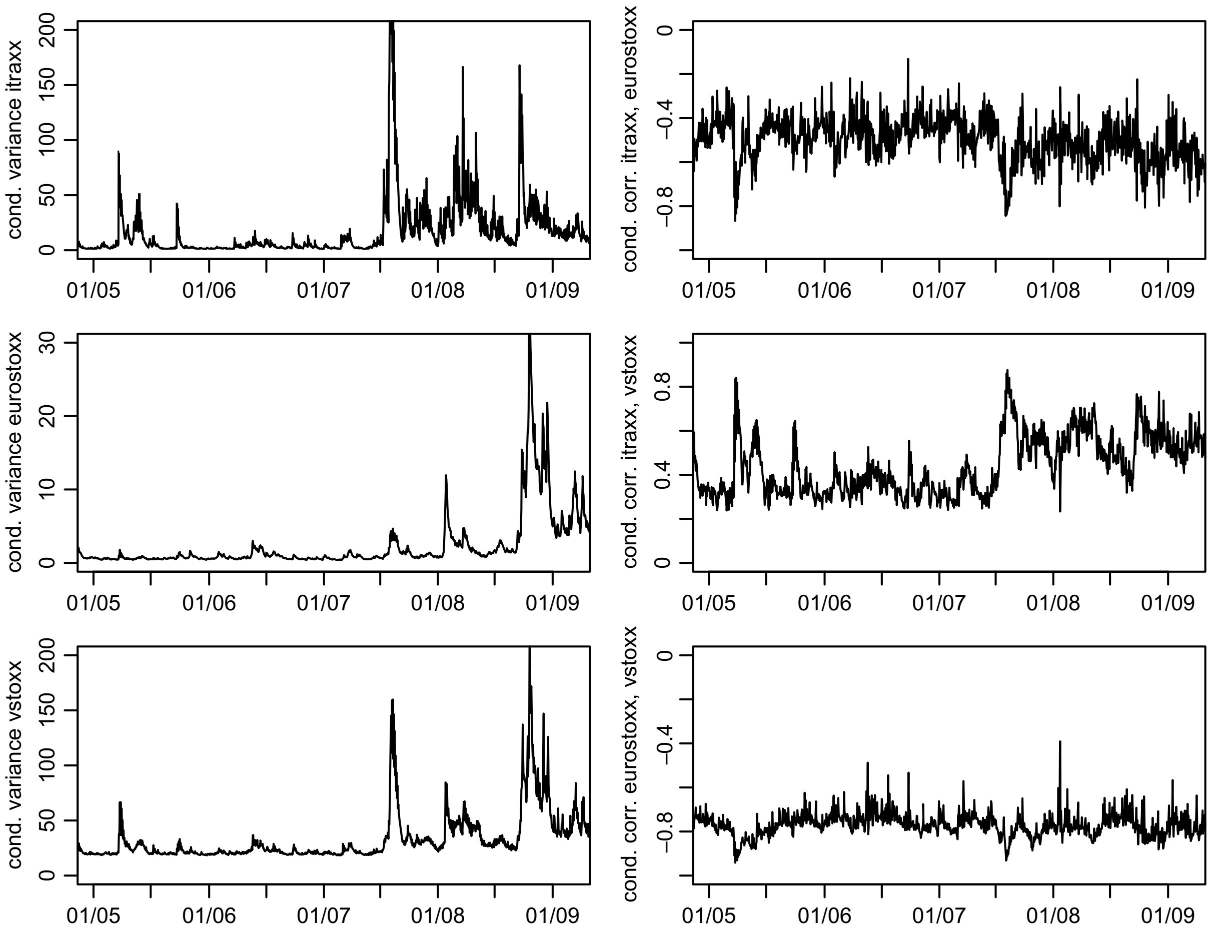

Figure 4: Conditional volatilities and correlations of the BEKK(2,2) model. Left side: individual conditional volatilities of the iTraxx, Euro Stoxx and VStoxx. Right side: correlations.

models (e.g. the DCC model), as mentioned above, the theoretical background regarding model characteristics and properties of estimators is rather sound. We use the AIC criterion for model order selection and compare orders of $p, q = 1, 2$ (see Table 4). The BEKK(1,1) model is clearly outperformed by the other three choices, which are very close to each other. As the model order $p = q = 2$ is best in terms of AIC, we decide in favor of the BEKK(2,2) model. We estimate the coefficients of the BEKK parametrization of H_t in (3) via ML estimation and obtain only 29 out of 42 significant coefficients at a confidence level of 90%. Following a consecutive multiple testing scheme we successively set insignificant coefficients to zero and finally obtain a model within which all remaining 28 coefficients are significant (see Table 5). The spectral radius of the estimated matrix $\sum_{i=1}^{2} A_i' \otimes A_i + \sum_{j=1}^{2} B_j' \otimes B_j \in \mathbb{R}^{9\times 9}$ is larger than one, therefore the process H_t is not stationary (see Engle and Kroner, 1995). Figure 4 displays the coefficients of H_t. We observe strongly varying conditional volatility and conditional correlations for all three time series. The volatility range is especially large for the iTraxx, with the lowest values close to zero and the peaks at 200. The structural break visible in the original time series (see Figure 1) at about mid of 2007 after the outbreak of the sub-prime crisis is clearly visible here as well.

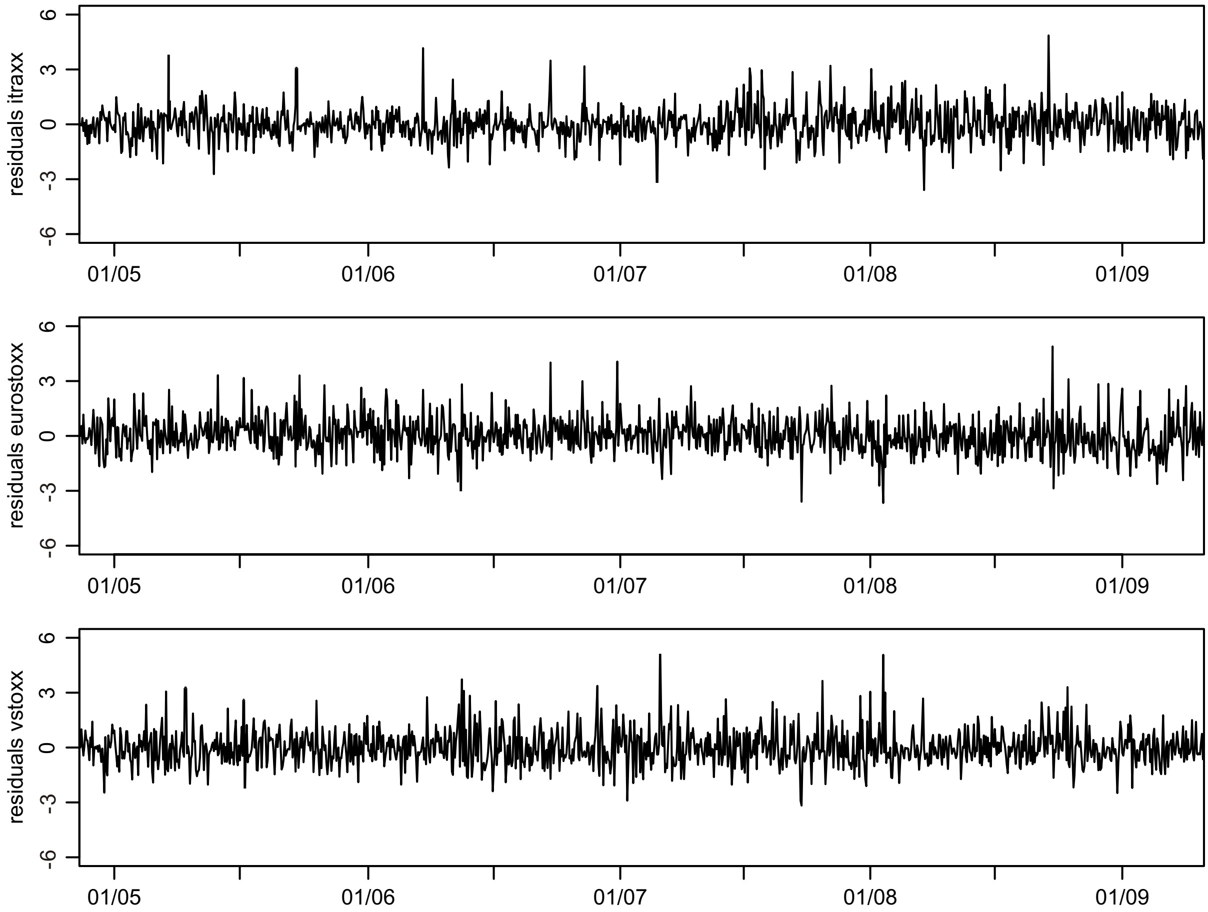

Figure 5: Residuals after fitting a BEKK(2,2) model to the residuals of the VAR(1) model.

After the break all three volatilities have a higher level on the whole, and vary more strongly. This, again, is particularly evident in the case of the iTraxx index. The correlation between the iTraxx and the Euro Stoxx as well as the Euro Stoxx and the VStoxx is negative, while the correlation between the iTraxx and the VStoxx is positive. The conditional correlations between the three time series fluctuate strongly over time. While the correlation between the iTraxx and the other two indices is stronger after the structural break, the correlation between the Euro Stoxx and the VStoxx stays on the same level, which is not surprising, as the values of the VStoxx are calculated on the basis of options on the Euro Stoxx. By its nature the VStoxx is therefore closely linked to the development of the Euro Stoxx.

Figure 5 shows the residuals after fitting the BEKK(2,2) model. In comparison with the original data in Figure 1, we can see that much of the previous volatility patterns have vanished, this implies that the BEKK model was able to capture the volatility structure of the data set. The autocorrelations and cross correlations of the residuals and the squared residuals are displayed in Figure 6. There are no significant auto- and cross correlations left. We conduct portmanteau and Lagrange multiplier tests and find no evidence of remaining ARCH effects. Overall, when considering the test results and the autocorrelations and cross correlations plots of all three time

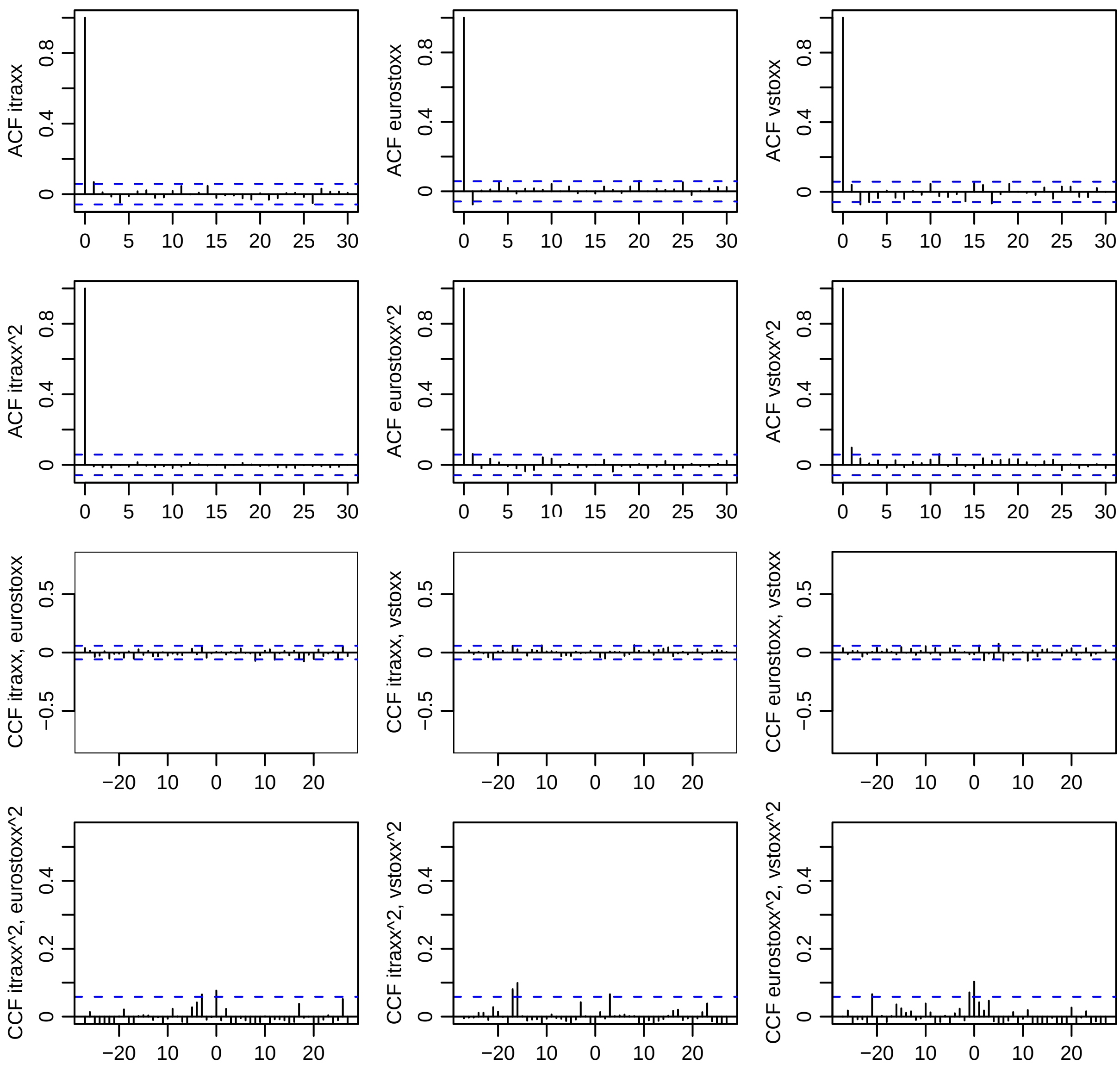

Figure 6: Autocorrelations and cross correlations with 95% confidence bounds of the residuals and the squared residuals after fitting a BEKK(2,2) model to residuals of the VAR(1) model.

series, the evidence that our model captures the structure in the second order moments of the time series well is very strong. When conducting the Jarque-Bera test for normality (see Jarque and Bera, 1987), the null hypothesis of normally distributed residuals is clearly rejected. Note, that this alone should not be viewed as a drawback of this modeling approach, as the model in (1) and (2) is based on white noise in contrast to normal innovations. However, in order to obtain consistency and asymptotic normality of the coefficient estimators strong requirements such as the existence of the eighth moments of the error distribution (Comte and Lieberman, 2003) are necessary. This limitation is obviously undesirable for financial time series, where the existence of higher order moments is regarded as problematic. When considering alternative heavy tailed error distributions such as the t-distribution these restrictions should be kept in

	est.	std.error	t-stat.		est.	std.error	t-stat.
ω_1	0.141	0.084	1.679	α_1	0.270	0.013	20.769
ω_2	0.027	0.087	0.310	α_2	0.133	0.037	3.595
ω_3	0.914	0.040	22.850	α_3	0.065	0.031	2.097
a	0.024	0.019	1.263	β_1	0.780	0.422	1.848
b	0.956	0.031	30.839	β_2	0.858	0.025	34.320
				β_3	0.909	0.032	28.406

Table 6: Coefficients of the DCC model with $p_i = q_i = 1 \forall\ i = 1, \ldots, N$ in (5) and $M = N = 1$ in (6).

mind. When the residuals are found to be skewed, the relevance of the Student distribution may be questioned. Therefore in this case, skewed distributions with fat tails, such as mixtures of multivariate normal densities or the generalized hyperbolic distribution are more suitable alternative error distributions.

3.4 A DCC Model for the Conditional Covariance

As an alternative modeling approach and comparison to the BEKK model we now fit the DCC model by Engle (2002) as given by (4)–(7) to the residuals of the VAR(1) model. As mentioned before, this model has become increasingly popular among practitioners, due to its flexibility and parsimony in combination with the simple estimation procedure. For our purpose we chose a relatively simple parametrization in (5) and (6) with a univariate GARCH(1,1) model for the individual conditional variances and the equivalent for the DCC parameters. The estimates are displayed in Table 6. In comparison to the BEKK model with 42 parameters in its non-reduced form, the DCC model only needs eleven parameters, which evidently simplifies matters a lot. The conditional covariance process is not stationary, as $\alpha_1 + \beta_1 = 1.050$ (see Engle and Sheppard, 2001).

The conditional volatilities and correlations in the DCC model are shown in Figure 7. The parameters of H_t resemble smooth versions of the ones in the BEKK model (see Figure 4). One reason for this can be seen in the significantly lower number of estimated parameters, which makes it difficult for the DCC model to capture all the variance in the data. The correlations of the iTraxx with the other two time series in the DCC model display a trend, while in the case of the BEKK model the correlations could be interpreted as local stationary time series with at least one structural break, e. g. in July 2007.

The residuals strongly resemble the residuals in the BEKK model (Figure 5), i.e. fat tailed white noise residuals. However, the autocorrelations and cross correlations in the DCC model (see Figure 8) differ slightly from those in the BEKK model. We observe that there is still some weak cross correlation at lag zero perceivable in the residuals, which contradicts the model assumption of white noise. Furthermore we see some cross correlations in the squared residuals, which implies that the model is a less good fit four our data set, when comparing with the BEKK

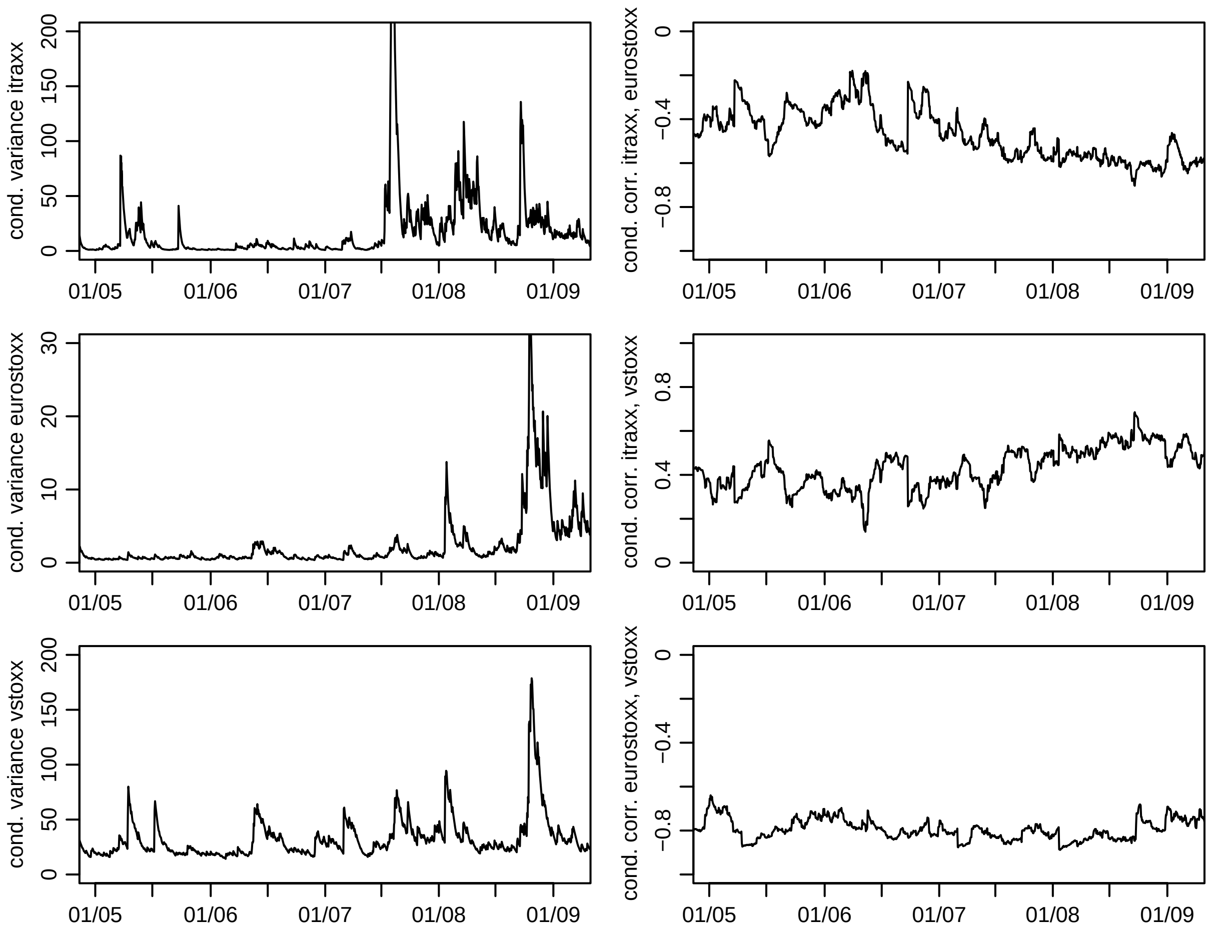

Figure 7: Conditional volatilities and correlations of the DCC model. Left side: individual conditional volatilities of the iTraxx, Euro Stoxx and VStoxx. Right side: correlations.

model in Figure 4. On the other hand the BEKK model has a total number of 42 parameters or 28 parameters in its reduced form, which by far exceeds the number of parameters of the DCC model. The portmanteau and Lagrange multiplier tests for remaining ARCH effects, as well as the tests for normality yield the same results as in the case of the BEKK model. Considering our empirical results when comparing the BEKK with the DCC model we find that the DCC model in its simplest form produces quite similar variance and correlation estimates, while having only one fourth of the parameters the BEKK model has. It is therefore not surprising that the DCC model enjoys widespread popularity with practitioners. The parsimony of the DCC approach is particularly compelling when high dimensional vector time series are involved, for instance when analyzing stock portfolios with many assets.

It is worth mentioning, that in our empirical study we also compared our VAR-GARCH modeling approach with a simple MGARCH model, where in a first step we subtracted the mean from the original time series, and then in a second step fitted a BEKK versus DCC model to the data. We found that this model clearly is a less good fit when comparing autocorrelations and cross correlations of the residuals with Figure 6 and Figure 8. This confirms our time-varying,

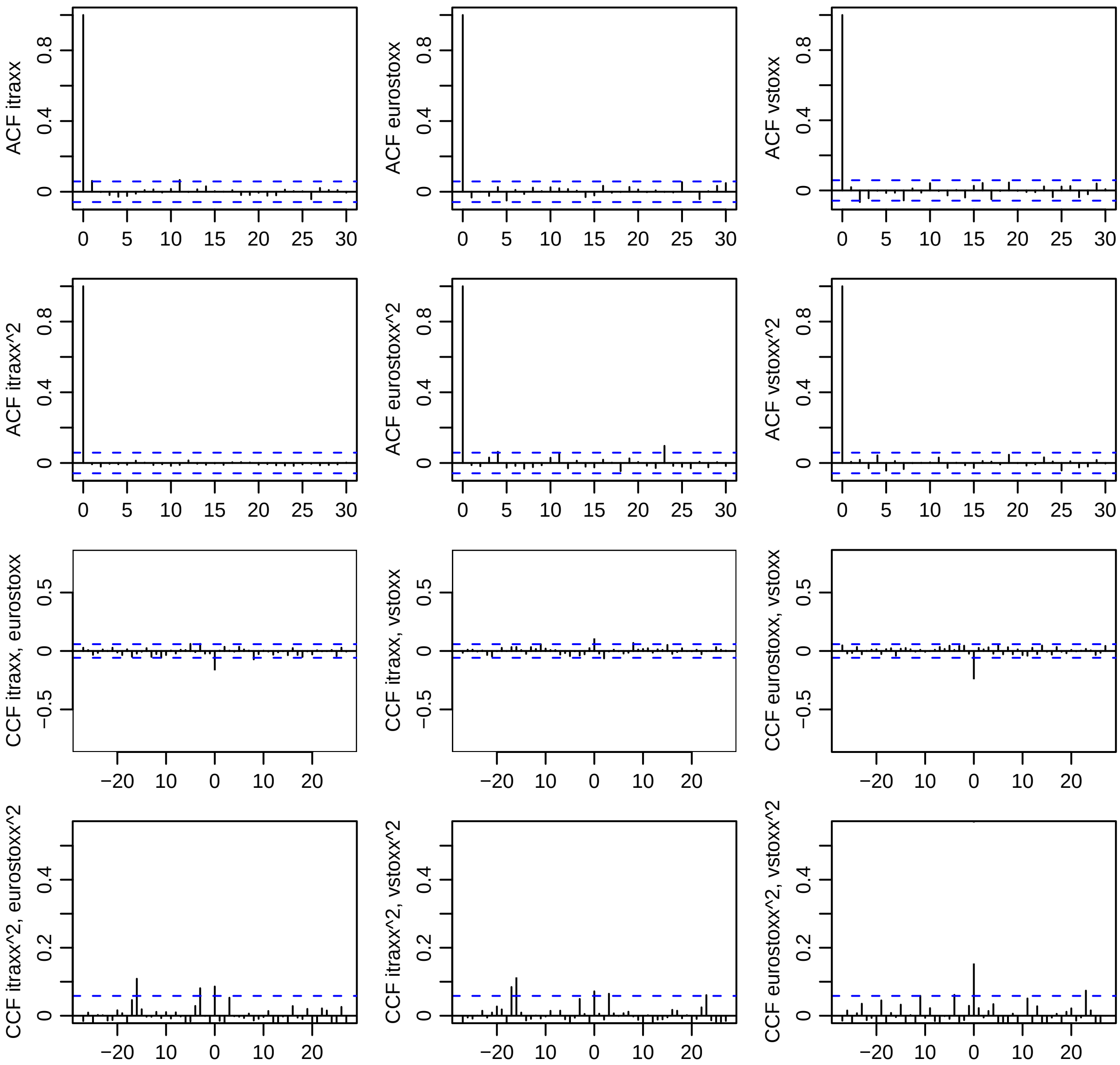

Figure 8: Autocorrelations and cross correlations with 95% confidence bounds of the residuals and the squared residuals after fitting the DCC-GARCH model to residuals of the VAR(1) model.

robustly estimated and combined VAR-GARCH modeling approach.

4 Conclusion

In our analysis we fitted two different VAR-GARCH models to a three-dimensional financial time series of daily quotes from June 2004 to April 2009, the iTraxx Europe index, the Euro Stoxx 50 index and the volatility index VStoxx. In an initial exploratory investigation we found evidence of weak autocorrelation in the time series, especially in the iTraxx and the Euro Stoxx index. Therefore, in order to capture the structure in the conditional mean, we fitted a VAR model to the data. We selected order one as recommended by the HQ and SC information criteria. In order to account for the apparent outliers in the data, and the strongly varying entries of the coefficient matrix, we robustly estimated a VAR(1) model with time-dependent coefficients using RLS estimation. By establishing two criteria for setting insignificant coefficients to zero we achieved a parsimonious model specification. Our empirical results indicate that the autoregressive coefficients vary strongly with time and even change their signs. Well-known interrelations, such as the negative correlation between CDS' and stocks are lost through the financial crisis. We found the model adequate in terms of the conditional mean, however most of the dependency structure of the time series was captured by the MGARCH models, which were fitted to the residuals subsequently.

From the variety of MGARCH modeling approaches developed up to date we chose the well known BEKK model, which is particularly compelling due to its parametrization that by definition guarantees the positive definiteness of the covariance matrix. We fitted a BEKK(2,2) model to the data, the model order being determined by AIC. As a model comparison we chose to fit a DCC model to the data, motivated by its widespread popularity among practitioners. For our purpose we chose a simple model class with a GARCH(1,1) model for the individual conditional variances and the equivalent for the DCC parameters.

We found that the conditional variances and correlations vary strongly with time. The correlations between the iTraxx and the Euro Stoxx and the Euro Stoxx and the VStoxx are negative, whereas the correlation between the iTraxx and the VStoxx is positive. The correlations increase significantly in absolute values after the outbreak of the financial crisis. The main difference between the two models lies in the smoother variance and correlation estimates in the DCC model. Besides that, the correlations in the DCC model display a trend, whereas in the BEKK model the correlations could be interpreted as local stationary time series. Both series of residuals are white noise yet not normally distributed.

In terms of the conditional mean, our results extend previous empirical studies, allowing for robustly estimated, time-varying coefficients. However, to the best of our knowledge, there are no existing studies of aggregated credit spreads, stocks and stock market volatility in which the conditional covariance structure is considered. Therefore our findings offer some of the first insights regarding the variance and correlation structure of this data set. We found evidence of strongly varying conditional variances and correlations, with dependencies increasing after the outbreak of the financial crisis. This knowledge opens the door to better decision tools in various areas, such as asset pricing, portfolio selection, and risk management. The dynamics of the financial crisis particularly with regard to the correlations between different asset classes may hence be understood from a new and more thorough view point.

References

Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In B. N. Petrov and F. Csáki (Eds.), *2nd International Symposium on Information Theory*, Budapest, pp. 267–281.

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control 19*(6), 716–723.

Alexander, C. and A. Kaeck (2006). Regimes in CDS spreads: a Markov switching model of iTraxx Europe indices. ICMA Centre Discussion Papers in Finance, Henley Business School, Reading University.

Bauwens, L., S. Laurent, and J. V. Rombouts (2006). Multivariate GARCH models: a survey. *Journal of Applied Econometrics 21*(1), 79–109.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics 31*(3), 307–327.

Bollerslev, T. (1990). Modelling the coherence in short-run nominal exchange rates: a multivariate generalized ARCH model. *The Review of Economics and Statistics 72*(3), 498–505.

Bollerslev, T., R. F. Engle, and J. M. Wooldridge (1988). A capital asset pricing model with time-varying covariances. *Journal of Political Economy 96*(1), 116–31.

Brockwell, P. J. and R. A. Davis (1991). *Time Series: Theory and Methods*. New York: Springer.

Byström, H. N. E. (2005). Credit default swaps and equity prices: the iTraxx CDS index market. Working Papers, Department of Economics, Lund University (Lund, Sweden).

Byström, H. N. E. (2006). CreditGrades and the iTraxx CDS index market. *Financial Analysts Journal 62*(6), 65–76.

Campbell, K. (1982). Recursive computation of M-estimates for the parameters of a finite autoregressive process. *The Annals of Statistics 10*(2), 442–453.

Caporin, M. and M. McAleer (2008). Scalar BEKK and indirect DCC. *Journal of Forecasting 27*(6), 537–549.

Caporin, M. and M. McAleer (2009). Do we really need both BEKK and DCC? A tale of two covariance models. Documentos del Instituto Complutense de Análisis Económico 0904, Universidad Complutense de Madrid, Facultad de Ciencias Económicas y Empresariales.

Comte, F. and O. Lieberman (2003). Asymptotic theory for multivariate GARCH processes. *Journal of Multivariate Analysis 84*(1), 61–84.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica 50*(4), 987–1007.

Engle, R. F. (2002). Dynamic conditional correlation: a simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *Journal of Business & Economic Statistics 20*(3), 339–350.

Engle, R. F. and K. F. Kroner (1995). Multivariate simultaneous generalized ARCH. *Econometric Theory 11*(1), 122–150.

Engle, R. F., K. F. Kroner, Y. Baba, and D. F. Kraft (1993). Multivariate simultaneous generalized ARCH. Economics Working Paper Series 89–57, Department of Economics, University of California at San Diego.

Engle, R. F. and K. Sheppard (2001). Theoretical and empirical properties of dynamic conditional correlation multivariate GARCH. NBER Working Paper 8554, National Bureau of Economic Research.

Ericsson, J., K. Jacobs, and R. A. Oviedo (2009). The determinants of credit default swap premia. *Journal of Financial and Quantitative Analysis 44*(1), 109–132.

Hamilton, J. (1994). *Time Series Analysis.* Princeton University Press.

Hannan, E. J. (1970). *Multiple Time Series.* New York: John Wiley & Sons.

Hannan, E. J. and B. G. Quinn (1979). The determination of the order of an autoregression. *Journal of the Royal Statistical Society, Series B* (41), 190–195.

Houweling, P. and T. Vorst (2005). Pricing default swaps: empirical evidence. *Journal of International Money and Finance 24*(8), 1200–1225.

Huber, P. J. (1981). *Robust Statistics.* New York: John Wiley & Sons.

Jarque, C. M. and A. K. Bera (1987). A test for normality of observations and regression residuals. *International Statistical Review 55*, 163–172.

Li, W. K. and Y. V. Hui (1989). Robust multiple time series modelling. *Biometrika 76*(2), 309–315.

Li, W. K., S. Ling, and M. McAleer (2002). Recent theoretical results for time series models with GARCH errors. *Journal of Economic Surveys 16*(3), 245–269.

Lütkepohl, H. (1991). *Introduction to Multiple Time Series Analysis.* Berlin: Springer-Verlag.

Lütkepohl, H. (2005). *New Introduction to Multiple Time Series Analysis.* Berlin: Springer-Verlag.

Newey, W. K. and D. McFadden (1994). Large sample estimation and hypothesis testing. In R. F. Engle and D. McFadden (Eds.), *Handbook of Econometrics*, Volume 4, pp. 2111–2245. Amsterdam: Elsevier Science.

Norden, L. and M. Weber (2009). The comovement of credit default swap, bond and stock markets: an empirical analysis. *European Financial Management 15*(3), 529–562.

Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics* (6), 461–464.

Sims, C. A. (1972). Money, income and causality. *American Economic Review 62*(4), 540–52.

Sims, C. A. (1980). Macroeconomics and reality. *Econometrica 48*(1), 1–48.

Sougné, C., C. Heuchenne, and G. Hübner (2008). The determinants of CDS prices: an industry-based investigation. In N. Wagner (Ed.), *Credit Risk*, pp. 85–96. Boca Raton: Chapman & Hall CRC.