CQE Center for Quantitative Economics

Volatility Transmission in Overlapping Trading Zones

Andreas Masuhr[†]

67/2017

[†] Department of Economics, University of Münster, Germany

wissen.leben
WWU Münster

Volatility Transmission in Overlapping Trading Zones

Andreas Masuhr

Westfälische Wilhelms-Universität Münster

Abstract

Previous volatility spillover models (Engle et al. (1990), Clements et al. (2015)) use artificially non overlapping trading zones to identify sources of volatility transmission between these zones. The problem of non overlapping zones is overcome using a copula GARCH approach that allows for multiple overlaps between zones incorporating vine copulas to flexibly model the dependence structure and to meet stylized facts of return data. Stationarity conditions are examined and identifications problems concerning previous work, as well, are pointed out.
To handle the relatively large parameter space, the model is estimated by Bayesian methods using a differential evolution MCMC (Braak 2006) approach. Simulation studies are carried out in order to ensure robustness against copula or error term misspecification and in order to analyze the identification problem.

1 Introduction

Interest in volatility spillovers between different market places first aroused at the late 80's and early 90's of the past century. Beginning with the seminal work of Engle et al. (1990) channels of volatility spillovers for different markets and different assets, see for example the work of Fleming and Lopez (1999) among others, have been detected. Those papers following the GARCH approach of Engle et al. (1990) to model spillover effects generally construct trading days consisting of geographically distinct trading zones, such that e.g. London and New York do not share common trading hours. Thus the effects measured from these models are prone to misspecification and sensitive to changes of the trading day structure. One more recent example is the paper of Clements et al. (2015) who re-investigate the results of Engle et al. (1990) and additionally estimate a set of different model specifications using realized volatility, decompositions into good and bad news as well separating realized volatility into a continuous component and a jump component. On the downside Clements et al. (2015) neither give deep insights in their methodology nor question the artificial construction of a non overlapping trading day. Additionally, only normally distributed innovations are used although financial returns usually exhibit fat tails (see e.g. Rogalski and Vinso (1978) or Boothe and Glassman (1987) who both find that either t distributions with slightly less than 4 degrees of freedom or stable distributions with stability parameter $\alpha < 2$ fit best to daily foreign exchange rates).

Knowledge of volatility spillover channels is not only interesting on it's own but can be used to precisely forecast over night risk of assets or portfolios, especially if they either denote in foreign currencies or are most commonly traded abroad. To be able to create precise forecasts, a reliable estimation of both, the underlying process of the (conditional) volatility as well as the distribution of the asset's return are important. To that end, a multivariate copula GARCH (CMGARCH) model that meets those requirements is introduced.

The remainder of this paper organizes as follows: section 2 introduces copulas and vine copulas which will be used later on to connect the error terms. Section 3 covers the Heat Wave and Meteor Shower model of Engle et al. (1990) and one extension of Clements et al. (2015), whereas section 4 extends this model class to overlapping trading zones. Section 5 explains the algorithm that is used for estimation and section 6 supplies simulation studies and robustness checks. The final section briefly concludes.

2 Copulas and Vine Copulas

This section introduces concept of copulas and vine copulas, before turning the focus on modeling spillovers in the subsequent section. Consider a m-dimensional random vector $\boldsymbol{X} = (X_1, \ldots, X_m)$ with joint distribution function $F_{\boldsymbol{X}}$. If this joint distribution function is continuous, there is a unique decomposition into the marginal distribution functions F_{X_1} to F_{X_m} and a corresponding copula $C(u_1, \ldots, u_m)$, where $u_i = F_{X_i}(x_i)$, according to Sklar's famous theorem:

$$F_X(\boldsymbol{x}) = C\big(F_{X_1}(x_1), \ldots, F_{X_m}(x_m)\big) = C\big(u_1, \ldots, u_m\big). \tag{1}$$

Hence, it is possible to model marginal distributions and their dependency structure independently of each other. Although copulas add flexibility in modeling multivariate data the extend of additionally flexibility is limited since most copulas are symmetric in the sense that (at least) their functional form is identical in each dimension or strong restrictions apply when building multivariate Archimedean copulas (see Savu and Trede (2010)). The concept of regular vines -or pair copula constructions-, introduced by Bedford and Cooke (2002) and further developed by Aas et al. (2009), Dißmann (2010) among others offer a solution to this issue. A regular vine is a factorization of a multivariate density function into its marginal densities and (conditional) copula densities. This factorization can be conceived of a set of spanning trees with each node being a (conditional) random variable and each edge being a (conditional) copula to connect these nodes. Given a $d-$dimensional random variable, a regular vine $\mathcal{V}$ can be formally described as (see Brechmann (2010))

Definition 2.0.1. *Regular Vine*

1. $\mathcal{V} = \big(T_1, \ldots, T_{d-1}\big)$
2. $T_1 = \big(N_1, E_1\big)$ *is a tree with nodes* $N_1 = \{1, \ldots, d\}$. *For* $i = 2, \ldots, n-1$, $T_i = \big(N_i, E_i\big)$ *is a tree with nodes* $N_i = E_{i-1}$
3. *For* $i = 2, \ldots, n-1$, *if* $\{a, b\} \in E_i$, *where* $a = \{a_1, a_2\}$ *and* $b = \{b_1, b_2\}$, *then exactly one element of* a *equals one element of* b.

The third part of this definition is usually referred to as *proximity condition* and ensures that only pairs of variables are modeled together in one stage of the vine that share a node in the previous stage. As a simple example of a three dimensional regular vine consider the following factorization:

$$\begin{aligned} f_{X,Y,Z}(x,y,z) =& f_X(x) f_Y(y) f_Z(z) \cdot \\ & c_{X,Y}\big(F_X(x), F_Y(y)\big) \cdot c_{Y,Z}\big(F_Y(y), F_Z(z)\big) \cdot c_{X,Z|Y}\big(F_X(x|Y), F_Z(z|Y)\big) \end{aligned}$$

The conditional distributions $F_X(x|Y)$ are computed using (see Schmitz (2003))

$$\boldsymbol{P}(X \leq x|Y = y) = \frac{\partial C(u,v)}{\partial v}\bigg|_{F_X(x),F_Y(y)}.$$

Visually, this vine can be depicted as three trees where all nodes in each tree are connected by a single path (see figure 1) and belongs to the subclass of D-vines (drawable vines). D-vines are a natural starting point for modeling different trading zones within a trading day since time develops on a single path. Another subclass, the C-vines (canonical vines) is generate by placing one node in the center of each tree and directly connecting all other variables to build a star. If there is only a single layer of overlapping trading zones, it is possible to truncate the vine structure after the first stage (i.e. all subsequent stages consist of independence copulas) in order to save computation time (see Brechmann et al. (2012)) Using only bivariate copulas in each of the trees allows

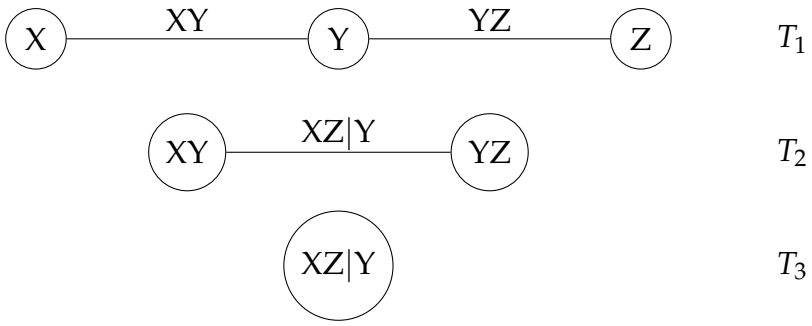

Figure 1: *Three dimensional regular vine*

for very flexible modeling of high dimensional data.
There are several possibilities to estimate regular vines either using a two stages approach by estimating the structure first and then fitting the pair copula parameters or by estimating both, structure and parameters at once (Dißmann et al. (2013)). A recent paper of Gruber and Czado (2015) introduces a fully Bayesian sampling algorithm to estimate the posterior distribution of regular vines. This approach gives posterior samples for both, the pair copula parameter and the structure of the vine. A comprehensive R-package for regular vines is supplied and maintained by Technical University of Munich (see Schepsmeier and Nagler (2017)).
In the remainder of this paper, the structure of the vine will be induced by the order of trading zones within a trading day and hence, only the copula parameters need to be estimated.

3 Heat Waves and Meteor Showers

In their seminal work Engle et al. (1990) detect two distinct effects of the transmission of volatility: one from adjacent markets (the so called meteor shower) and the other one from previous trading days in the same market (the so called heat wave). In a model with more than one trading zone, the meteor shower pattern reflects effects from neighboring trading zones on the current market's volatility, i.e. effects that move around the globe within the trading day, whereas the heat wave pattern describes local (in terms of geographically distinct markets) transmission effects. Canonical univariate GARCH models are thus only capable of modeling heat wave patterns and the approach of Engle et al. hence is an extension to multivariate GARCH models although only one asset is examined at once.

The original model of Engle et al. (1990) contains $i = 1, ..., m$ independent, i.e. non-overlapping trading zones and describes returns r_i in zone i as (the notation is taken from Clements et al. (2015)) as:

$$r_t^i = \varepsilon_t^i; \quad \varepsilon_t^i \sim N(0, h_t^i) \tag{2}$$

$$h_t^i = \kappa_i + \alpha_{ii} h_{t-1}^i + \sum_{j=1}^{i-1} \beta_{ij} \epsilon_{j,t}^2 + \sum_{j=i}^{n} \gamma_{ij} \epsilon_{j,t-1}^2 \tag{3}$$

Using matrix notation this model can be rewritten following Clements et al. (2015)

$$h_t = K + A h_{t-1} + B \epsilon_t^2 + G \epsilon_{t-1}^2, \tag{4}$$

where h_t and ϵ_t are m-element vectors of returns and innovations at time t, respectively. Both, Engle et al. (1990) and Clements et al. (2015) use data samples from the three trading zones USA (New York), Japan (Tokyo) and Europe (London) to estimate their models. I follow their choice of trading zones throughout this paper and hence, the parameter matrices corresponding to (4) are given by:

$$A = \begin{bmatrix} \alpha_{11} & 0 & 0 \\ 0 & \alpha_{22} & 0 \\ 0 & 0 & \alpha_{33} \end{bmatrix} \quad B = \begin{bmatrix} 0 & 0 & 0 \\ \beta_{21} & 0 & 0 \\ \beta_{31} & \beta_{32} & 0 \end{bmatrix} \quad G = \begin{bmatrix} \gamma_{11} & \gamma_{12} & \gamma_{13} \\ 0 & \gamma_{22} & \gamma_{23} \\ 0 & 0 & \gamma_{33} \end{bmatrix} \tag{5}$$

According to the matrix structure, the hypothesis of meteor showers would be supported by nonzero coefficients in matrix B and nonzero off diagonal coefficients in matrix$G$ and the heat wave hypothesis would be supported by

nonzero diagonal elements of G and A.
Up to now, only effects from current and lagged innovations on volatility in the current zone as well as effects from previous volatility have been modeled, but no intra day effects from volatility of preceding zones. To model such volatility spillovers the specification is changed according to Clements et al. (2015) into:

$$h_t = K + Ah_{t-1} + \tilde{A}h_t + Br_t^2 + Gr_{t-1}^2 \tag{6}$$

with

$$A = \begin{bmatrix} \alpha_{11} & 0 & \alpha_{13} \\ 0 & \alpha_{22} & 0 \\ 0 & 0 & \alpha_{33} \end{bmatrix} \quad B = \begin{bmatrix} 0 & 0 & 0 \\ \beta_{21} & 0 & 0 \\ 0 & \beta_{32} & 0 \end{bmatrix}$$
$$G = \begin{bmatrix} \gamma_{11} & 0 & \gamma_{13} \\ 0 & \gamma_{22} & 0 \\ 0 & 0 & \gamma_{33} \end{bmatrix} \quad \tilde{A} = \begin{bmatrix} 0 & 0 & 0 \\ \tilde{\alpha}_{21} & 0 & 0 \\ 0 & \tilde{\alpha}_{32} & 0 \end{bmatrix}. \tag{7}$$

Now, A contains all effects of previous day volatility in the same zone (on the diagonal) plus the effect of preceding volatility in the last zone on current volatility in the first zone, i.e. in the setting of three zones (London, New York and Tokyo) the effect of preceding volatility in Tokyo on current volatility in London. Throughout the remainder of the paper, specification (6) with parameter matrices (7) will be used.
One key assumption of the models of Clements et al. (2015) and Engle et al. (1990) is that trading zones are non overlapping, i.e. the trading times of two distinct trading zones must not share common time. If trading zones are non overlapping, then returns from different zones can be expected to be independent, given previous information. Thus it is sufficient to model a diagonal conditional covariance matrix $H_t = diag(h_t)$ for the return series. Empirically this assumption does not hold as there are overlaps between major trading zones, e.g. there is a three hour overlap between trading in Europe (London: 8am to 4pm GMT) and the United States (New York: 1pm to 9pm GMT). Additionally, a model that allows for overlapping trading zones can incorporate a higher number of trading zones without the need of setting their trading times to crudely short intervals. Ignoring overlapping trading zones can lead to biased estimators of spillover effects, both intra day and between trading days. The bias is strongest if a lot of volatility is created within the overlapping trading time and high shares of the overlap are absorbed by only one of the trading zones. In the previous example of three trading zones with

a three hour overlap between London and New York all parameter matrices can be expected to be biased if trading times are set accordingly to Clements et al. (2015) or Fleming and Lopez (1999), i.e. almost all overlap is absorbed by trading in New York (2.5 hours) and only a small share is added to trading in London (0.5 hours).

4 Volatility transmission in overlapping trading zones

4.1 Modeling overlapping trading zones

To examine volatility transmissions in overlapping trading zones those zones need to be modeled in advance. To that end, consider a world that consists of only two trading zones Z_1 and Z_2. Trading in the first zone starts at z_{1a} and ends at z_{1b} and trading times in zone two are from z_{2a} to z_{2b}. Usually (see models of Engle et al. (1990), Fleming and Lopez (1999) or Clements et al. (2015)) the restriction that those zones don't overlap is imposed, i.e. $z_{1b} \leq z_{2a}$. This assumption is set aside here and the duration that both zones overlap is denoted by $t_{12} = z_{1b} - z_{2a}$, the duration in which the first (second) trading zone is the only active zone is given by $t_1 = z_{2a} - z_{1a}$ ($t_2 = z_{2b} - z_{1b}$) and their total trading hours are denoted by $T_1 = z_{1b} - z_{1a}$ ($T_2 = z_{2b} - z_{2a}$). Let η_i denote the standardized news occurring in zone i with $\eta \sim \mathcal{F}$ and $\mathcal{F}$ being an arbitrary distribution with mean zero and variance equal to one. Further denote by $\eta_1^\star$, $\eta_2^\star$ and $\eta_{12}^\star$ the non standardized (i.e. their varaince is proportional to the trading time in the respective zone) news occurring exclusively in the first, second and overlapping zone. Then the former can be expressed as sums of the latter:

$$\eta_1 = \frac{\eta_1^\star + \eta_{12}^\star}{\sqrt{T_1}} \quad \text{and} \quad \eta_2 = \frac{\eta_{12}^\star + \eta_2^\star}{\sqrt{T_2}}. \tag{8}$$

Accordingly, the covariance between the news of two overlapping trading zones can be computed as

$$Cov(\eta_1, \eta_2) = E\left[\frac{\eta_1^\star + \eta_{12}^\star}{\sqrt{T_1}} \cdot \frac{\eta_{12}^\star + \eta_2^\star}{\sqrt{T_2}}\right] = \frac{t_{12}}{\sqrt{T_1 T_2}}, \tag{9}$$

which equals their correlation. Obviously, the correlation between news of overlapping trading zones is only based on the share of overlapping trading time compared to the total trading times. Since correlation only roots in identical information from both zones it is in line with the market efficiency hypothesis.

4.2 Volatility Transmission

The approach by Clements et al. (2015) reveals two major shortcomings: first, stylized facts of stock returns (or other financial instruments) are not met by the model as it only incorporates normally distributed innovations and second, the model does not include cross correlated returns as beginning with Engle et al. (1990), most subsequent papers constructs artificial trading days with non overlapping trading zones. To overcome both issues a model is proposed that follows the idea of copula-based multivariate GARCH models introduced by Lee and Long (2009). The framework can incorporate both, non normal error term distributions as well as arbitrary dependency structures between return series.
Consider the following model for m potentially overlapping trading zones:

$$r_t = \boldsymbol{H}_t^{\frac{1}{2}} \varepsilon_t \tag{10}$$

$$H_t = diag(h_t)^{\frac{1}{2}} \, \boldsymbol{M_t} \, diag(h_t)^{\frac{1}{2}} \tag{11}$$

$$h_t = \boldsymbol{k} + \boldsymbol{A} h_{t-1} + \tilde{\boldsymbol{A}} h_t + \boldsymbol{B} r_t^2 + \boldsymbol{G} r_{t-1}^2, \tag{12}$$

$$\varepsilon_t = \boldsymbol{\Sigma}_t^{-\frac{1}{2}} \eta_t \tag{13}$$

$$\eta_t \sim \mathcal{F}_{1,\dots,m}(\eta_1, \dots, \eta_m; \boldsymbol{\theta}_t) \tag{14}$$

where the parameter matrices $\mathbf{A}$, $\tilde{\mathbf{A}}$, $\mathbf{B}$ and $\mathbf{G}$ are defined as above and are collected in $\boldsymbol{\alpha} = \{\mathbf{A}, \tilde{\mathbf{A}}, \mathbf{B}, \mathbf{G}\}$. The observed returns r_t are modeled as product of a uncorrelated set of innovations ε_t and the conditional covariance matrix H_t. The uncorrelated innovations are computed from another set of error terms, η_t whose distribution should not be restricted here. Instead, the distribution $\mathcal{F}$ of the innovations η_t is constructed from m potentially different marginal distributions with $E(\eta_t) = 0$ and $E(\eta_t \eta_t') = \Sigma_t$, according to the previous section, and a copula $\mathcal{C}$ in order to model the dependence between innovations within a trading day (i.e. between trading zones). To allow both, multiple overlaps and high flexibility, the copula linking the individual trading zones will be constructed using a D-vine structure. Let $\boldsymbol{\theta}_t$ collect all marginal distribution parameters as well as all copula parameters. For identification assume that $diag(\Sigma_t) = \mathbf{1}_m$. Accordingly, $E(\varepsilon_t) = 0$ and $E(\varepsilon_t \varepsilon_t') = \mathbf{I}_m$ and consequently the innovations ε_t are uncorrelated but still dependent as the covariance matrix Σ_t can only partially account for the dependence of the innovations (at least if a non Gaussian copula is chosen). For the rest of this paper the parameter vector of the copula and the marginal distributions are assumed to be constant over time, i.e. $\boldsymbol{\theta}_t = \boldsymbol{\theta}$ and consequently $\Sigma_t = \Sigma$. As $\boldsymbol{M_t}$ is the correlation matrix of

the returns r_t it is set equal to $\mathbf{\Sigma}$ and is assumed to be constant over time, as well.

4.3 Stationarity

For the non overlapping specification with $\tilde{\boldsymbol{A}} = \mathbf{0}$ and $\boldsymbol{A}$ being diagonal, stationarity conditions can be found in Engle et al. (1990). The stationarity conditions for the above specification will now be evaluated. The one step ahead forecast of h_t is given by $E\left[h_t|\mathcal{F}_{t-1}\right]$ where $\mathcal{F}_{t-1}$ denotes all information up to $t-1$, i.e $\mathcal{F}_{t-1} = \left\{r_0, \ldots, r_{t-1}, h_0, \ldots, h_{t-1}\right\}$. Starting at equation (12) and taking expectations yields:

$$E\Big[h_t|\mathcal{F}_{t-1}\Big] = E\Big[\boldsymbol{k} + \boldsymbol{A}h_{t-1} + \tilde{\boldsymbol{A}}h_t + \boldsymbol{B}r_t^2 + \boldsymbol{G}r_{t-1}^2|\mathcal{F}_{t-1}\Big],$$

and applying iterated expectations to $E\left[r_t^2|\mathcal{F}_{t-1}\right]$ results in

$$E\Big[h_t|\mathcal{F}_{t-1}\Big] = \boldsymbol{k} + \boldsymbol{A}h_{t-1} + \tilde{\boldsymbol{A}}E\Big[h_t|\mathcal{F}_{t-1}\Big] + \boldsymbol{B}E\Big[h_t|\mathcal{F}_{t-1}\Big] + \boldsymbol{G}r_{t-1}^2,$$

since $diag(\boldsymbol{H}_t) = h_t$ and hence

$$E\Big[h_t|\mathcal{F}_{t-1}\Big] = (\boldsymbol{I} - \tilde{\boldsymbol{A}} - \boldsymbol{B})^{-1}(\boldsymbol{k} + \boldsymbol{A}h_{t-1} + \boldsymbol{G}r_{t-1}^2). \tag{15}$$

Accordingly, the $n \geq 1$ step ahead forecast is given by

$$E\Big[h_t|\mathcal{F}_{t-n}\Big] = (\boldsymbol{I} - \tilde{\boldsymbol{A}} - \boldsymbol{B})^{-1}\Big(\boldsymbol{k} + (\boldsymbol{A} + \boldsymbol{G})E\Big[h_{t-1}|\mathcal{F}_{t-n}\Big]\Big), \tag{16}$$

applying iterated expectations on $E\left[r_t^2|\mathcal{F}_{t-n}\right]$ and $E\left[r_{t-1}^2|\mathcal{F}_{t-n}\right]$.
Evaluating $E\Big[h_{t-1}|\mathcal{F}_{t-n}\Big]$ in (16) recursively for $n \to \infty$ reveals that stationarity requires all eigenvalues of $(\boldsymbol{I} - \tilde{\boldsymbol{A}} - \boldsymbol{B})^{-1}(\boldsymbol{A} + \boldsymbol{G})$ to lie inside the unit circle. If the process is stationary, the unconditional variance is given by

$$\boldsymbol{h} = \lim_{n\to\infty} E\Big[h_t|\mathcal{F}_{t-n}\Big] = (\boldsymbol{I} - \boldsymbol{A} - \tilde{\boldsymbol{A}} - \boldsymbol{B} - \boldsymbol{G})^{-1}\boldsymbol{k}. \tag{17}$$

Equation (17) allows to control a given parametrization for non-negativity of the unconditional variance $\boldsymbol{h}$.

4.4 Simulation

To derive a simulation scheme for overlapping trading zones, first Cholesky type decompositions are applied on the conditional covariance matrix $\boldsymbol{H}_t$ and the correlation matrix $\boldsymbol{M}$, i.e

$$\boldsymbol{H}_t = diag(h_t)^{\frac{1}{2}}\boldsymbol{M}diag(h_t)^{\frac{1}{2}} = \boldsymbol{L_t}\boldsymbol{L_t'}$$
$$\boldsymbol{M} = \boldsymbol{D}\boldsymbol{D}'$$

and hence, $\boldsymbol{L_t} = diag(h_t)^{\frac{1}{2}}\boldsymbol{D}$, where both, $\boldsymbol{L}$ and $\boldsymbol{D}$ are lower triangular matrices. Then the daily returns of all trading zones are given by $r_t = diag(h_t)^{\frac{1}{2}}\boldsymbol{D}\varepsilon$. Now, the model can be simulated by an inner loop iterating through trading zones and an outer loop iterating over trading days.

4.5 Identification

Starting at equation (6) and solving the system in a way that only lagged volatilities and past returns remain on the right hand side, one obtains:

$$\begin{aligned}
h_{3,t} = k_3 + \alpha_{32}k_2 + \alpha_{32}\alpha_{21}k_1 &+ h_{3,t-1}(\alpha_{33} + \alpha_{13}\alpha_{21}\alpha_{32}) \\
&+ h_{2,t-1}(\alpha_{32}\alpha_{22}) \\
&+ h_{1,t-1}(\alpha_{32}\alpha_{21}\alpha_{11}) \\
&+ r^2_{3,t-1}(\gamma_{33} + \gamma_{13}\alpha_{21}\alpha_{32}) \\
&+ r^2_{2,t-1}(\gamma_{22}\alpha_{32}) + r^2_{2,t}(\beta_{32}) \\
&+ r^2_{1,t-1}(\gamma_{11}\alpha_{21}\alpha_{32}) + r^2_{1,t}(\beta_{21}\alpha_{32})
\end{aligned}$$

$$\begin{aligned}
h_{2,t} = k_2 + \alpha_{21}k_1 &+ h_{2,t-1}(\alpha_{22}) + h_{1,t-1}(\alpha_{21}\alpha_{11}) + h_{3,t-1}(\alpha_{13}\alpha_{21}) \\
&+ r^2_{3,t-1}(\gamma_{13}\alpha_{21}) \\
&+ r^2_{2,t-1}(\gamma_{22}) \\
&+ r^2_{1,t-1}(\gamma_{11}\alpha_{21}) + r^2_{1,t}(\beta_{21})
\end{aligned}$$

$$\begin{aligned}
h_{1,t} = k_1 &+ h_{1,t-1}(\alpha_{11}) + h_{3,t-1}(\alpha_{13}) \\
&+ r^2_{3,t-1}(\gamma_{13}) \\
&+ r^2_{1,t-1}(\gamma_{11})
\end{aligned}$$

The model is identified if it is impossible to generate identical time series of $\boldsymbol{h}$ with different parameter constellations. For h_1, this is the case since each parameter is multiplied with the respective variable, individually. The equation for h_2 reveals that potential for identification problems is given since $h_{1,t-1}$ and $h_{3,t-1}$ are multiplied with products of parameters. Fortunately increasing/decreasing α_{21} demands an increase/decrease of both α_{11} and α_{13} and hence affects h_1. Turning to the equation for h_3 and bearing in mind that α_{21}, α_{11} and α_{13} must not be increased/decreased there are no more free parameters to change, resulting in identical values for h_3 without affecting h_1 and/or h_2. Note that identification of the GARCH parameters does not depend on whether or not the model includes overlapping trading zones.

Now, consider an alternative specification in which the data generating process is such that in the first of the above equations $h_{3,t}$ is a linear function of $h_{1,t}$, ie. $h_{3,t} = \nu h_{1,t}$, $\nu \neq 0$. Then, multiplying α_{11} by δ_{11} can be fully compensated by multiplying α_{13} by $\frac{\delta_{11}}{\nu}$ in the second and third equation. Consequently, the model is no longer identified. Unfortunately, many parameter constellations that seem to be quite innocent at first glance result in high correlations between intra day volatilities and hence lead to poorly identified parameters. An example of such parameterizations is given in one of the subsequent simulation studies. Additionally, intra day correlation between volatilities is noch (substantially) driven by intra day correlation between returns, i.e. the problem is also present in the non overlapping specifications used by Clements et al. (2015).

5 Bayesian Inference

5.1 Differential Evolution MCMC

Differential Evolution MCMC (DE-MCMC) is a combination of the Differential Evolution optimizer and Markov Chain Monte Carlo Methods that offers two major advantages: first, DE-MCMC does not rely on a precise specification of the proposal distribution and thus can handle large parameter spaces easily and second, DE-MCMC profits from running many Markov Chains in parallel and thus is perfectly suited to be used on a large scale cluster computer.
The general idea of DE-MCMC is rather simple[1]: N markov chains are run in parallel with $\theta_{i,j}$ denoting the parameter vector of chain j in iteration i. Then a value for $\theta_{i+1,j}$ is proposed by

$$\theta_{p,j} = \theta_{ij} + \gamma(\theta_{i,a} - \theta_{i,b}) + \epsilon \tag{18}$$

where a and b are two different randomly drawn elements from $\{1, ...N\} \backslash j$ and ϵ is drawn from a symmetric distribution with unbounded support to ensure irreducibility of the chains. In other words, DE-MCMC randomly chooses two different chains at each step, computes the difference of the parameter vector of those chains, scales this difference with a factor γ, adds some random noise ϵ and finally adds everything to the parameter vector of the current chain. Then, the proposed new parameter vector is accepted with probability

[1] For an almost textbook like treatment of the algorithm, see the paper of Braak (2006), that also contains some pseudo code for a basic DE-MCMC sampler.

$\alpha = min\left\{1, \frac{p(\theta_{p,j}|\boldsymbol{y})}{p(\theta_{i-1,j}|\boldsymbol{y})}\right\}$[2] with $p(\cdot|\boldsymbol{y})$ being the posterior distribution of the parameter vector θ given the data $\boldsymbol{y}$.
For a large number of chains N and a small variance of the noise ϵ, the proposal asymptotically (in this context, asymptotically refers to the number of chains and not to sample size) looks like $\theta_{p,j} = \theta_{ij} + \gamma\varepsilon$ as $N \to \infty$ with $E(\varepsilon) = \mathbf{0}$ and $Cov(\varepsilon) \to 2\Omega$, the covariance matrix of the posterior distribution (see Braak (2006)).
The given scheme results[3] in N Markov chains with unique stationary distribution that has density $p(\cdot|\boldsymbol{y})^N$. Thus, once all chains converged, all draws can be used as posterior samples leading to a large reduction in computation time. In a simulation study on a 100-dimensional normal distribution, Braak shows that convergence speed strongly depends on the number of chains and can be substantially increased by choosing high numbers of chains. This underlines that DE-MCMC is perfectly tailored for cluster computers with the ability of evaluating many chains in parallel. This argument gets even stronger the more complex and time consuming the evaluation of the likelihood is as the fraction of overhead reduces, the longer each individual processor needs for computation.
Convergence of the Markov chains can be assessed by computing the between- and within-sequence variances B and W (see Gelman et al. (2004)):

$$B = \frac{T}{N-1}\sum_{j=1}^{N}(\bar{\theta}_{\cdot j} - \bar{\theta})^2, \text{ where } \bar{\theta}_{\cdot j} = \frac{1}{T}\sum_{i=1}^{T}\theta_{i,j}, \ \bar{\theta} = \frac{1}{N}\sum_{j=1}^{N}\bar{\theta}_{\cdot j}$$

$$W = \frac{1}{N}\sum_{j=1}^{N}s_j^2, \text{ where } s_j^2 = \frac{1}{T-1}\sum_{i=1}^{T}(\theta_{i,j} - \bar{\theta}_{\cdot j})^2.$$

Then $\hat{R}$ measures the potential scale reduction of the current distribution of θ as $T \to \infty$ with:

$$\hat{R} = \sqrt{\frac{\frac{T-1}{T}W + \frac{1}{T}B}{W}}. \tag{19}$$

Usually values of $\hat{R}$ lower than 1.2 indicate convergence of the chains of a single parameter (see Gelman et al. (2004).).

[2]The acceptance probability is just the ratio of the posterior distributions as the DE-MCMC algorithm is used with a random walk proposal and thus just behaves like a random walk Metropolis-Hastings algorithm. For a textbook treatment see Greenberg (2008).
[3]For a proof, see Braak (2006).

5.2 Estimation

The previously described DE-MCMC algorithm is used for estimation by altering the simulation scheme in two ways: first a block wise sampler is used and second, the scaling parameter γ is fitted adaptively in order to achieve block wise acceptance ratios between 30% and 40% in each chain for a rolling window of 200 observations. The noise, added to the difference of two chains is distributed according to a normal distribution with expectation 0 and standard deviation equal to γ^2 ensuring relatively small noise compared to the difference of chains. Each full simulation step is split up into six blocks of which the first three sample all GARCH parameters (except the constant) belonging to the three trading zones. The fourth block samples a new level of the stationary volatility using the moment conditions implied by equation (17). The fifth block draws all marginal distribution parameters and is skipped for normally distributed errors and the final block samples all vine copula parameters. All prior distributions are set to uniform distributions either within the boundaries imposed by stationarity of the GARCH model or within natural boundaries, e.g. all correlation parameters have priors uniformly on $(-1, 1)$. In total, 64 chains are run in parallel on a single node of Palma-NG[4].

The correlation matrix $\mathbf{\Sigma}$ of the innovations η_t is approximated by monte carlo simulation of the underlying distribution using a sample size of 20,000. The log-likelihood is given by the sum of the log-densities of the marginal distributions and the log-copula-density plus the sum of the Jacobian of the transformation (see Lee and Long (2009)) from innovations η_t to returns r_t, i.e. the sum of $ln\left(\left|\Sigma^{0.5}(\boldsymbol{\theta})H_t^{-0.5}(\boldsymbol{\alpha})\right|\right)$ for all observations $t = 1, \ldots, T$ with $\boldsymbol{\alpha}$ denoting all GARCH parameters and $\boldsymbol{\theta}$ denoting the parameters of the marginal distributions and the copula. Then,

$$ln(\mathcal{L}) = \sum_{t=1}^{T} ln\big(f_1(\eta_{1,t})\big) + ... + ln\big(f_m(\eta_{m,t})\big) + ln\big(c(F_1(\eta_{1,t}), ..., c(F_m(\eta_{m,t}))\big) + \\ + \sum_{t=1}^{T} ln\left(\left|\Sigma^{0.5}(\boldsymbol{\theta})H_t^{-0.5}(\boldsymbol{\alpha})\right|\right).$$

[4]Palma-NG is an extension of PALMA (Paralleles Linux-System für Münsteraner Anwender) the high performance computer of WWU. It can be used by every person that is member of the group u0clustr (the group can be joined through meinZIV) and allows for small scale parallelization on a single node with low effort in terms of adapting code as each node of the cluster supports parallel computing on 64 cores. For more information see ZIV (2017) or consider the helpful examples on I:/Jan/HPC.

is the log-likelihood. As η_t needs to be recomputed, whenever either $\boldsymbol{\alpha}$ or $\boldsymbol{\theta}$ are changed, each block of the DE-MCMC sampler requires a full evaluation of the log-likelihood.

6 Simulation Studies

In this section, two simulation studies are carried out in order to investigate the previously addressed identification problem. All estimates are based on simulated samples consisting of three non overlapping trading zones for a period of 2,000 trading days.
The first simulation uses the parameterization that was estimated in Clements et al. (2015) Table 4 for a Euro-USD exchange rate data set:[5]

$$\boldsymbol{A} = \begin{bmatrix} 0.6478 & 0 & 0.1163 \\ 0 & 0.8983 & 0 \\ 0 & 0 & 0.8299 \end{bmatrix}, \boldsymbol{B} = \begin{bmatrix} 0 & 0 & 0 \\ 0.0452 & 0 & 0 \\ 0 & 0.0000 & 0 \end{bmatrix},$$
$$\boldsymbol{G} = \begin{bmatrix} 0.0683 & 0 & 0.0000 \\ 0 & 0.0176 & 0 \\ 0 & 0 & 0.0587 \end{bmatrix}, \tilde{\boldsymbol{A}} = \begin{bmatrix} 0 & 0 & 0 \\ 0.1352 & 0 & 0 \\ 0 & 0.1131 & 0 \end{bmatrix},$$
$$\boldsymbol{\rho} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \boldsymbol{k} = \begin{bmatrix} 6.4437 \cdot 10^{-9} \\ 2.0239 \cdot 10^{-8} \\ -3.3753 \cdot 10^{-9} \end{bmatrix}. \tag{20}$$

Innovations are assumed to be multivariate normal. This parameterization causes cross correlations between conditional volatilities to be around 95%[6]. As previously mentioned, this correlation can lead to poorly identified parameters which is the case here and can easily be seen from the joint posterior of α_{22} and α_{21} as depicted in Figure 2. Note that although the values of α_{21} are substantially larger than one, the stationarity condition still holds. From this posterior sample its clear that once there are significant volatility transmissions from one trading

[5] As the constant $\boldsymbol{k}$ is omitted in the tables of Clements et al. (2015), it has been fitted in order to match the unconditional volatilities, given in table 1 of Clements et al. (2015).

[6] Computing this correlation, one has to bare in mind that every trading zone has two different time spans between a neighboring zone and itself: The time from the end of zone A to the beginning of zone B and the time from the end of zone B to the beginning of zone A. However, storing the conditional variances in a matrix considers only one of these times for each zone. Hence the correlation between zones i and j with i being the earlier trading zone can be computed by: $Cor(h_i, h_j) = \frac{1}{2}\Big(Cor(h_{i,t}, h_{j,t}) + Cor(h_{i,t+1}, h_{j,t})\Big)$

zone to another within a trading day the model fails to precisely quantify the magnitude of transmission effects.

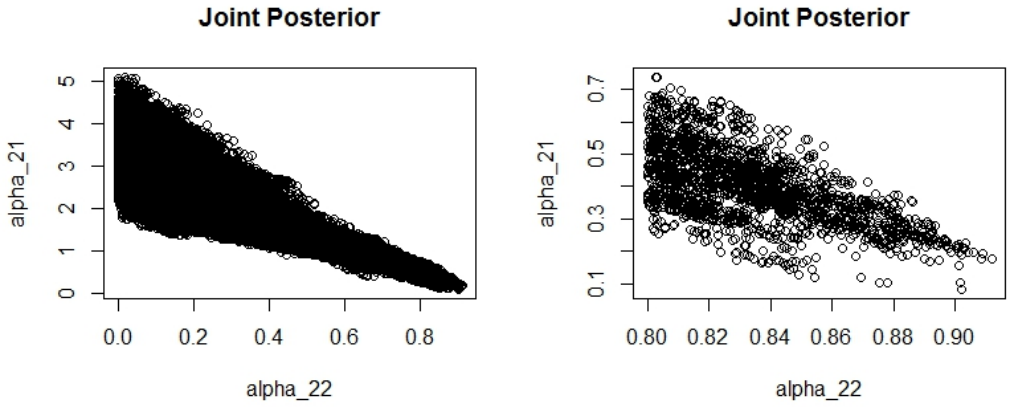

Figure 2: *Left panel: sample from the joint posterior distribution of α_{22} and α_{21}. Right panel: sample from the joint posterior with $\alpha_{22} \geq 0.8$*

The second simulation study is carried out with the parametrization estimated by Clements et al. (2015) for the bond market[7]. Again, $\boldsymbol{k}$ has been fitted to match unconditional volatilities of the three trading zones.

$$\boldsymbol{A} = \begin{bmatrix} 0.6178 & 0 & 0.0000 \\ 0 & 0.9027 & 0 \\ 0 & 0 & 0.9502 \end{bmatrix}, \boldsymbol{B} = \begin{bmatrix} 0 & 0 & 0 \\ 0.0192 & 0 & 0 \\ 0 & 0.0112 & 0 \end{bmatrix},$$
$$\boldsymbol{G} = \begin{bmatrix} 0.2819 & 0 & 0.0000 \\ 0 & 0.0818 & 0 \\ 0 & 0 & 0.0402 \end{bmatrix}, \tilde{\boldsymbol{A}} = \begin{bmatrix} 0 & 0 & 0 \\ 0.0000 & 0 & 0 \\ 0 & 0.0000 & 0 \end{bmatrix},$$
$$\boldsymbol{\rho} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \boldsymbol{k} = \begin{bmatrix} 4.4129 \cdot 10^{-8} \\ 1.2900 \cdot 10^{-9} \\ 4.531 \cdot 10^{-9} \end{bmatrix}. \tag{21}$$

This setting yields much lower cross sectional correlations between the conditional variances of roughly 30% and hence, estimation does not suffer from identification problems. Consequently, posterior means of all parameters do not deviate much from the values chosen for simulation. Posterior histograms

[7] They also estimated a specification for the stock market, but it is not possible to simulate data from their parameter constellation since conditional variances become negative, frequently!

and trace plots for all parameters are given in the appendix.
Additionally, robustness checks are carried out for the properly identified parametrization (21) using different error distributions and different copulas for an overlapping scenario. Simulating the model with t distributed margins with four degrees of freedom and estimating using normal margins reveals little to none problems. Some parameters are slightly biased (α_{22}, γ_{22} and α_{33}, γ_{33}) but most are unproblematic. This picture changes substantially if the wrong copula is chosen for estimation. Simulating with a vine built of t-copula with four degrees of freedom and a single overlap ($\rho_{21} = 0.3$) and estimating a Gaussian copula, instead, yields some severely biased posterior distributions. The posterior of α_{11} far below the value used for simulation, the posterior of α_{22} is not affected but the posterior of α_{33} is bimodal, now, with the uttermost mass between 0 and 0.2 whereas the true value is 0.95. All $\tilde{\boldsymbol{A}}$ and $\boldsymbol{B}$ coefficients are either biased (α_{21} and β_{21}) or bimodal and biased (α_{32} and β_{32}). All $\boldsymbol{G}$ and copula parameters are however in line with the parametrization used for simulation.
Summarizing, if the model is properly identified, the model is quite robust against error term misspecification. On the other hand, there are much stronger robustness issues if the copula is chosen incorrectly.

7 Concluding Remarks

The paper at hand is an extension to the volatility transmission literature. The artificial assumption of non overlapping trading zones is set aside and a vine copula GARCH model is put in place. Stationarity conditions for the GARCH specifications are developed and both, identification and robustness issues are examined. The introduced model is highly flexible as no assumptions are made on error term distributions or copula families and has been estimated using a Bayesian sampler that profits from being run in parallel on a cluster computer. The results of Clements et al. (2015) need to be reviewed, in particular against the background of potential identification issues concerning their specification and the assumption of non overlapping trading zones.

8 Appendix

8.1 Posterior Distributions and Trace Plots of the first simulation study

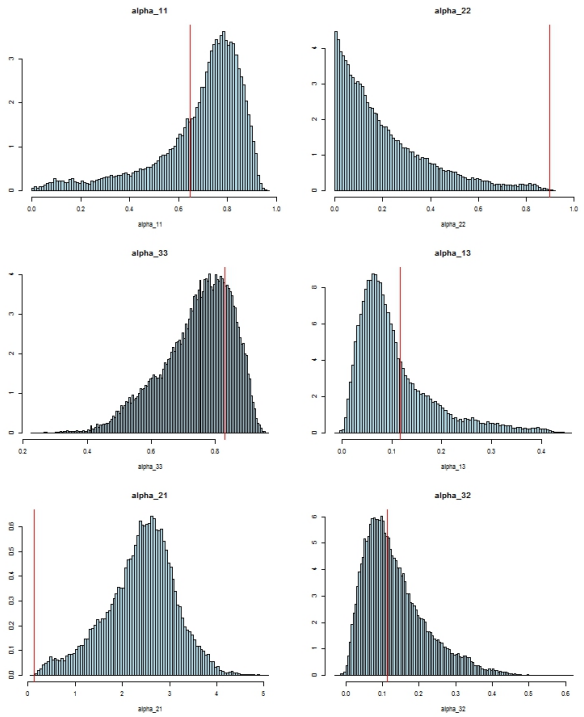

Figure 3: *Posterior distributions of all* $\boldsymbol{A}$ *and* $\tilde{\boldsymbol{A}}$ *parameters. The red lines display the true values.*

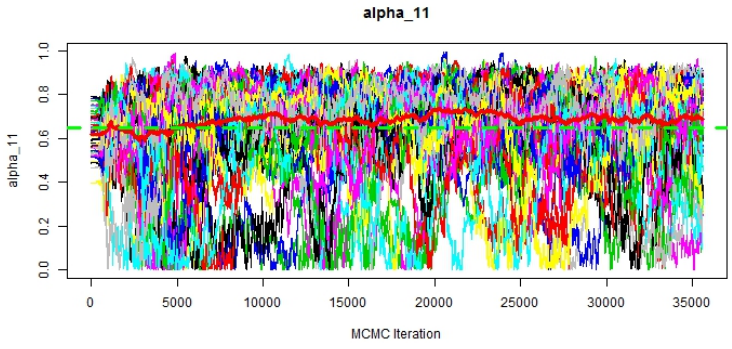

Figure 4: *Trace plot of* α_{11} *using 64 chains in parallel.*

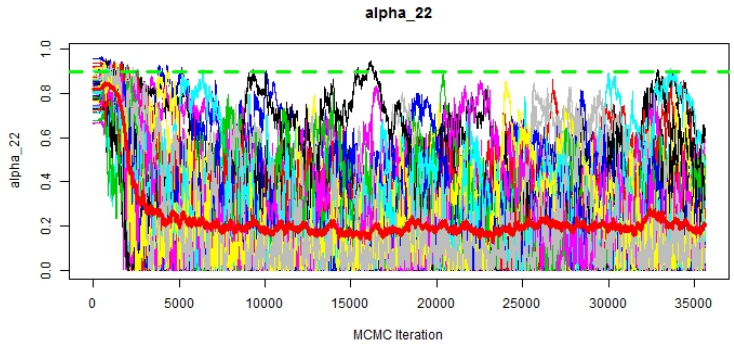

Figure 5: *Trace plot of α_{22} using 64 chains in parallel.*

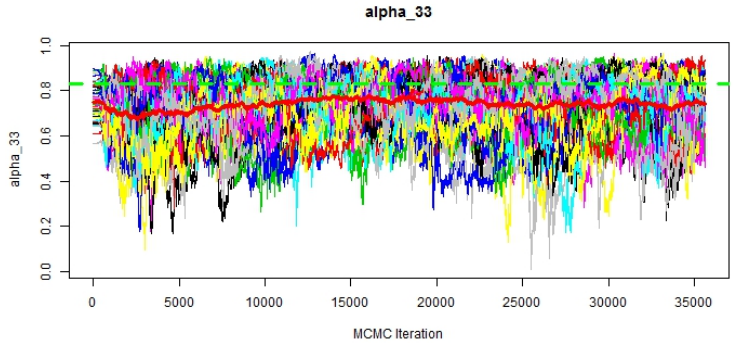

Figure 6: *Trace plot of α_{33} using 64 chains in parallel.*

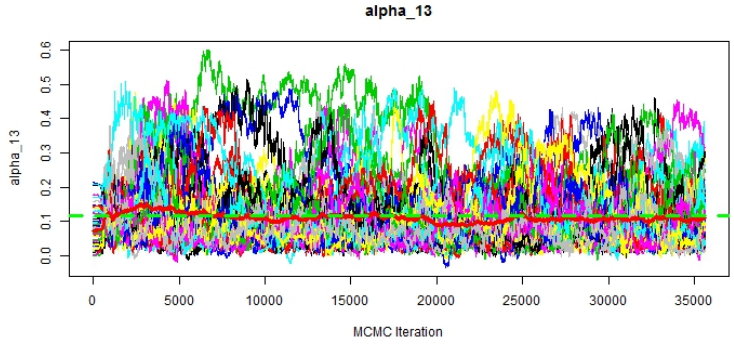

Figure 7: *Trace plot of α_{13} using 64 chains in parallel.*

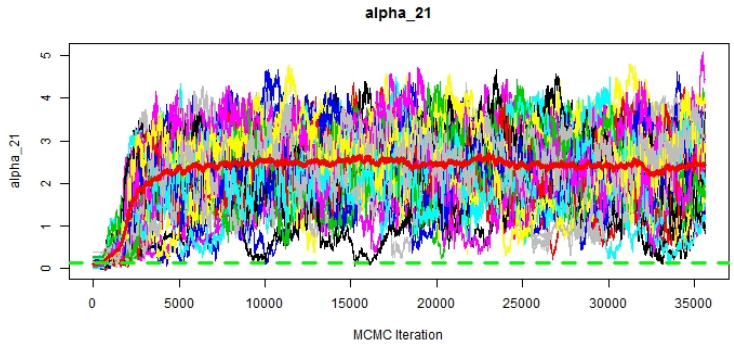

Figure 8: *Trace plot of α_{21} using 64 chains in parallel.*

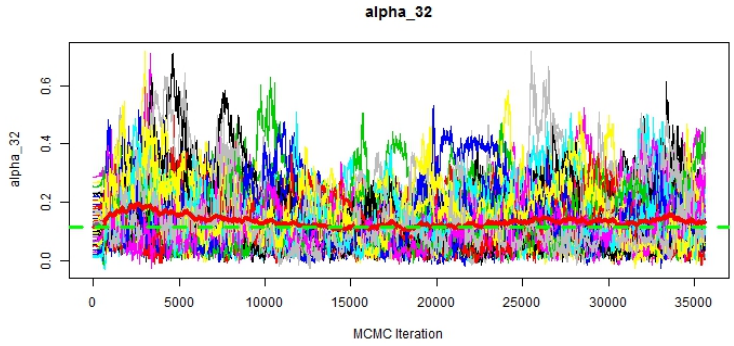

Figure 9: *Trace plot of α_{32} using 64 chains in parallel.*

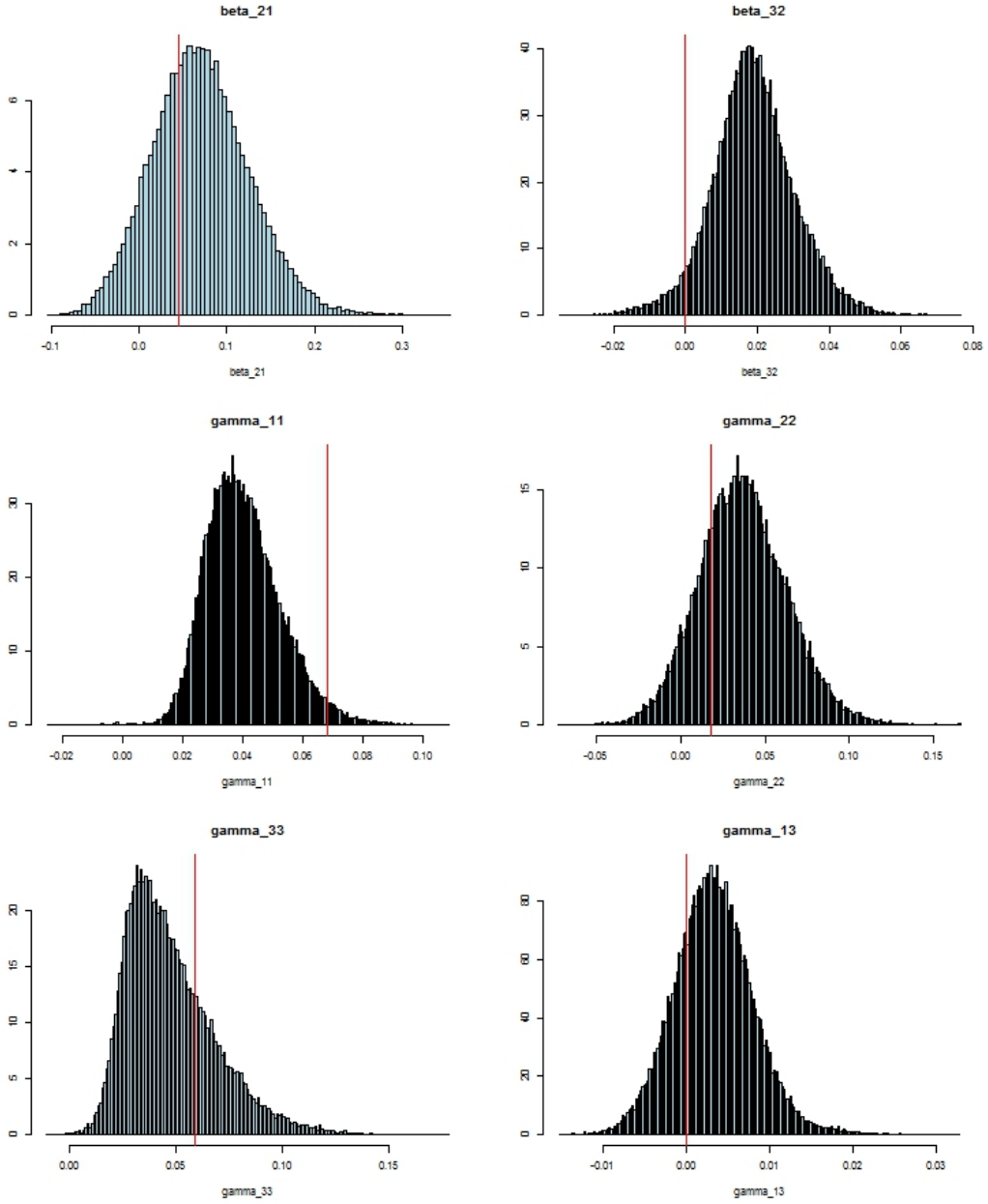

Figure 10: *Posterior distributions of all* ***B*** *and* ***G*** *parameters. The red lines display the true values.*

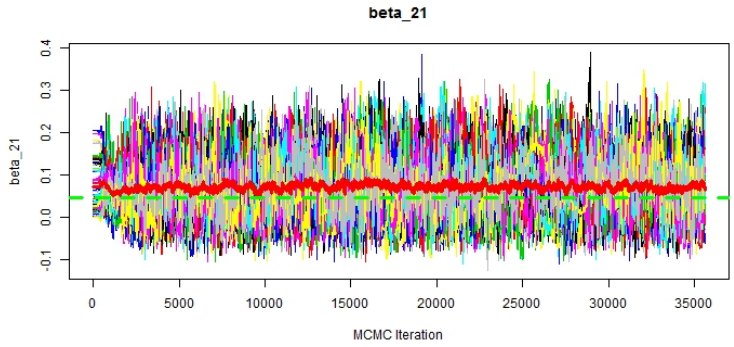

Figure 11: *Trace plot of $\beta 21$ using 64 chains in parallel.*

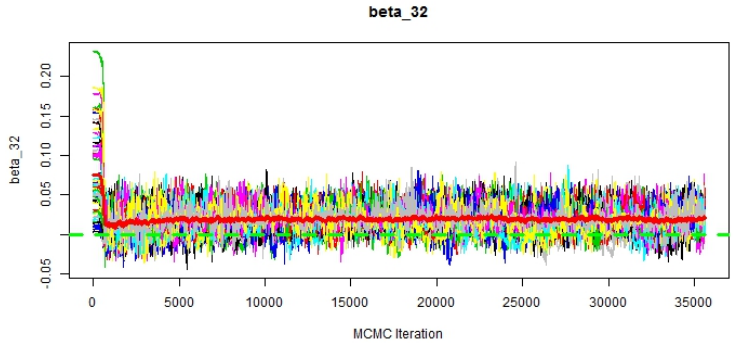

Figure 12: *Trace plot of $\beta 32$ using 64 chains in parallel.*

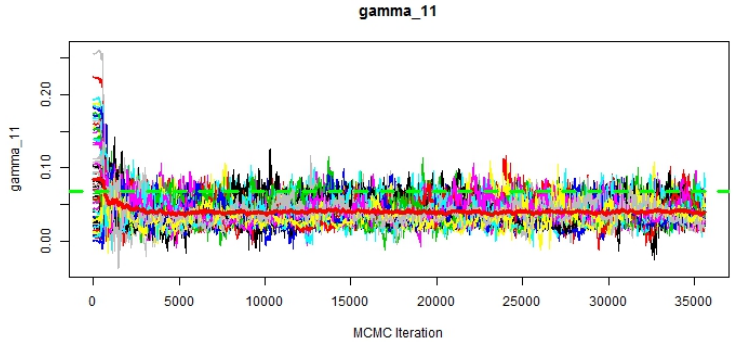

Figure 13: *Trace plot of $\gamma 11$ using 64 chains in parallel.*

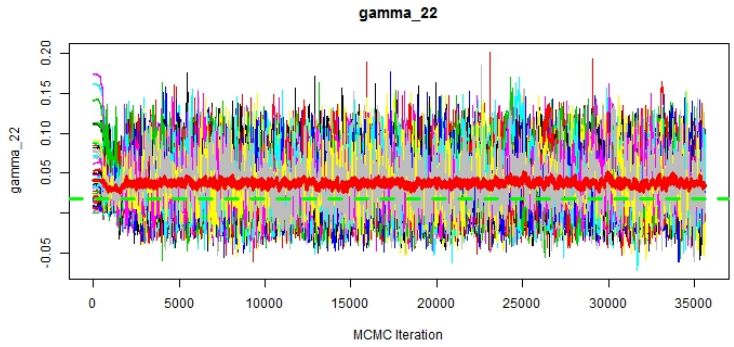

Figure 14: *Trace plot of $\gamma 22$ using 64 chains in parallel.*

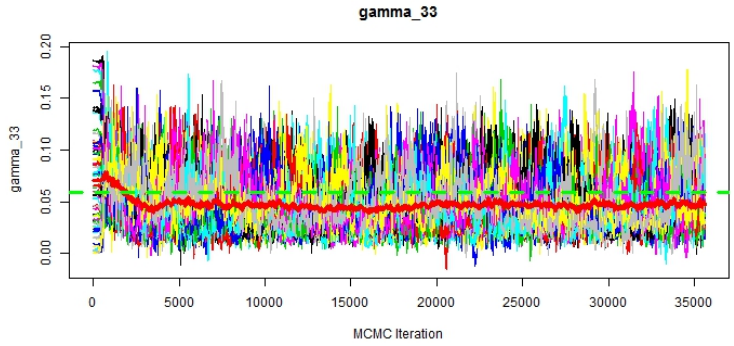

Figure 15: *Trace plot of $\gamma 33$ using 64 chains in parallel.*

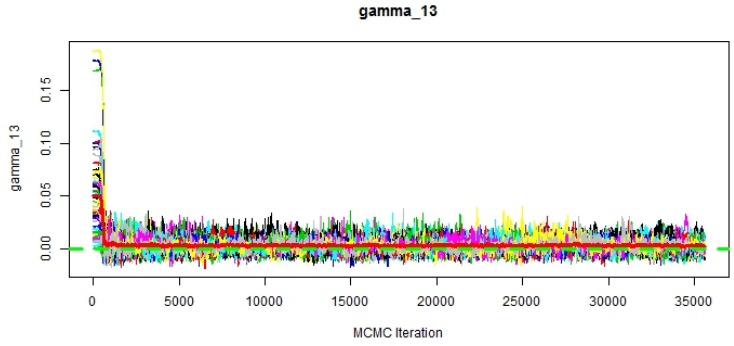

Figure 16: *Trace plot of $\gamma 13$ using 64 chains in parallel.*

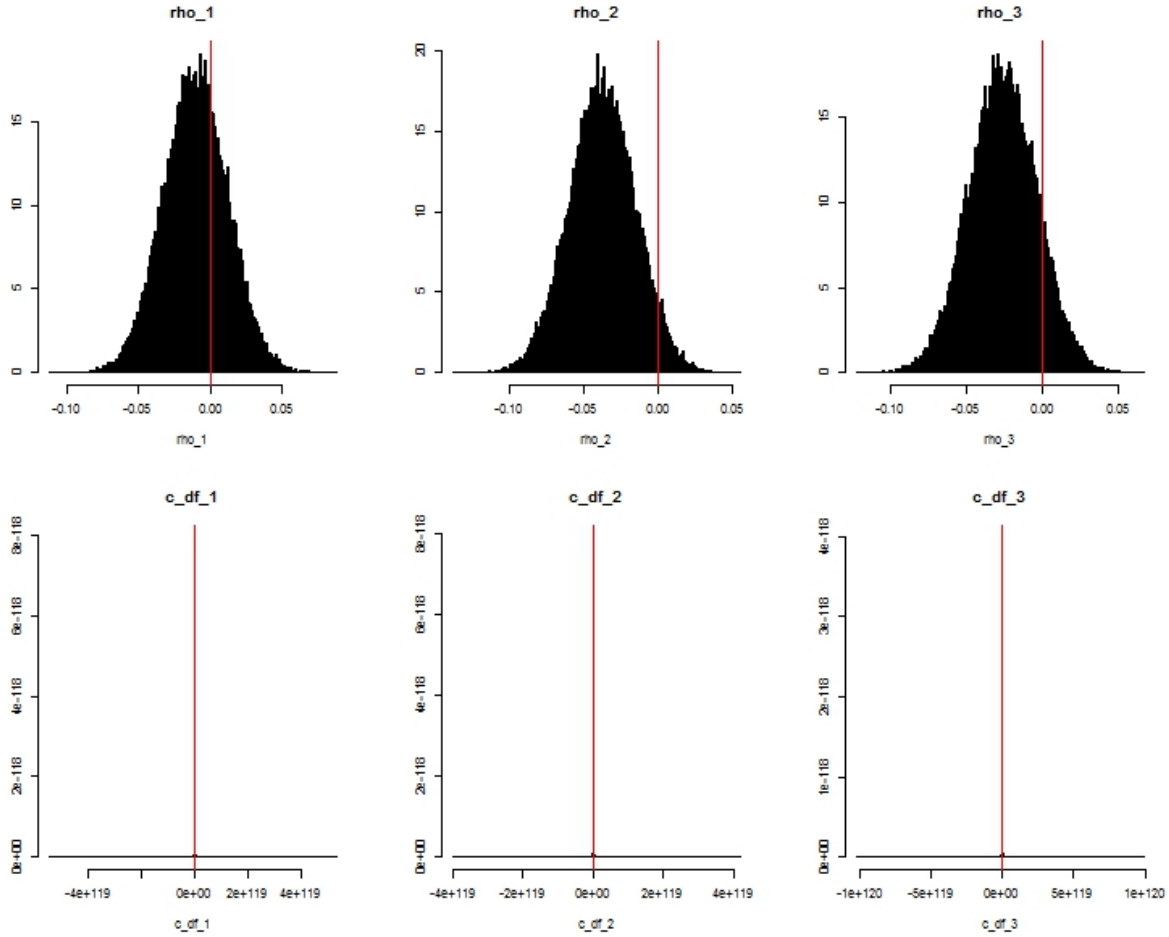

Figure 17: *Posterior distributions of all copula parameters. The red lines display the true values.*

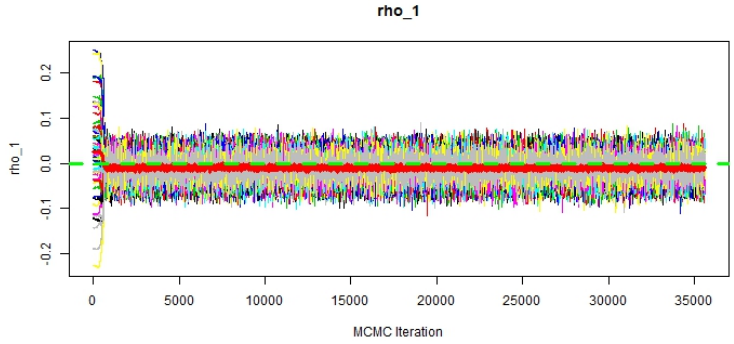

Figure 18: *Trace plot of ρ_1 using 64 chains in parallel.*

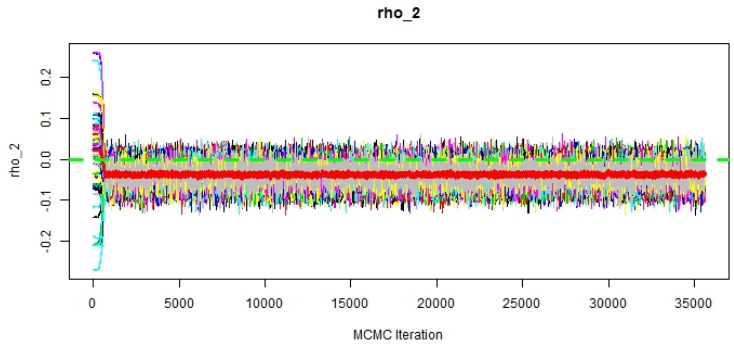

Figure 19: *Trace plot of ρ_2 using 64 chains in parallel.*

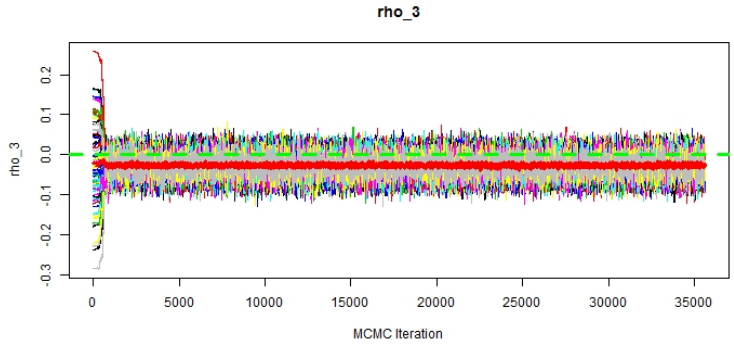

Figure 20: *Trace plot of ρ_3 using 64 chains in parallel.*

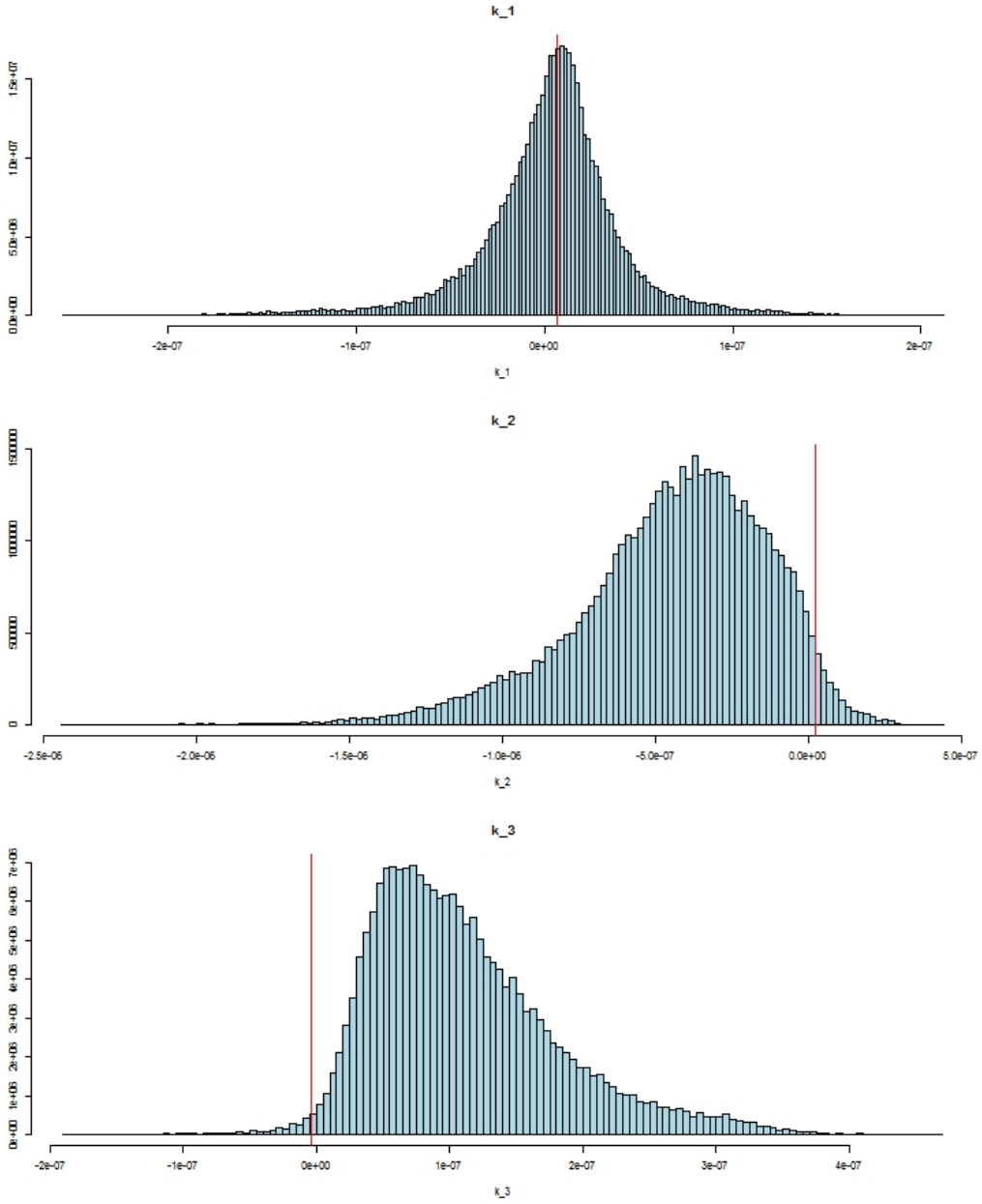

Figure 21: *Posterior distributions of all* $\mathbf{k}$ *parameters. The red lines display the true values.*

8.2 Posterior Distributions and Trace Plots of the second simulation study

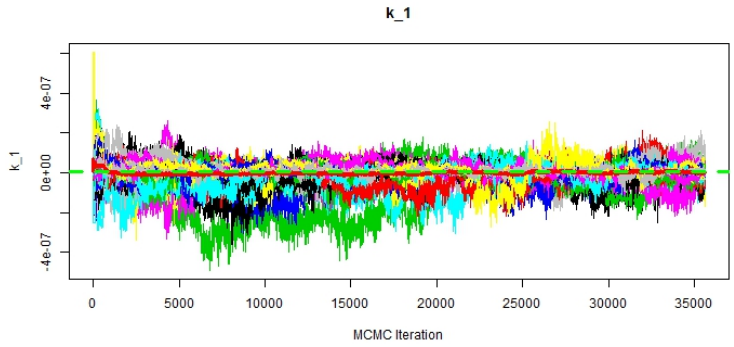

Figure 22: *Trace plot of k_1 using 64 chains in parallel.*

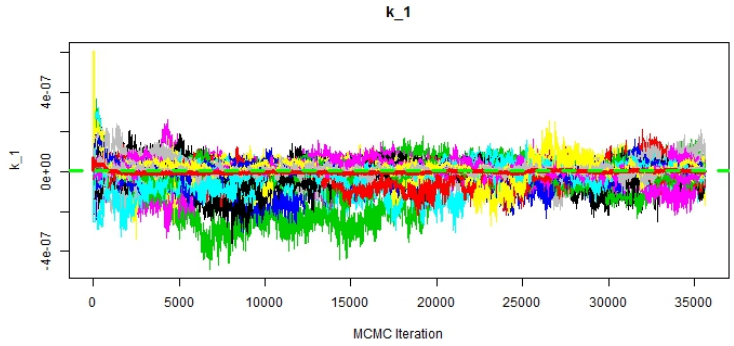

Figure 23: *Trace plot of k_2 using 64 chains in parallel.*

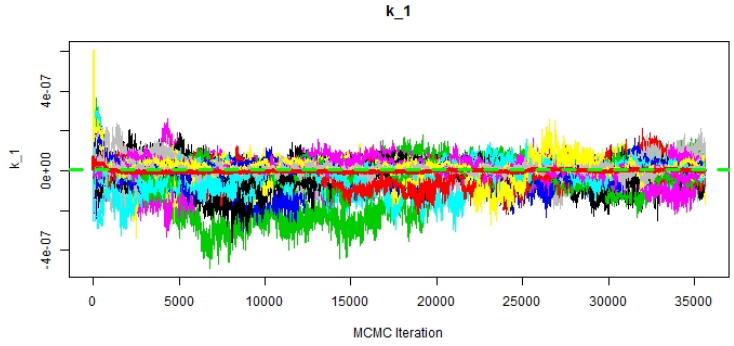

Figure 24: *Trace plot of k_3 using 64 chains in parallel.*

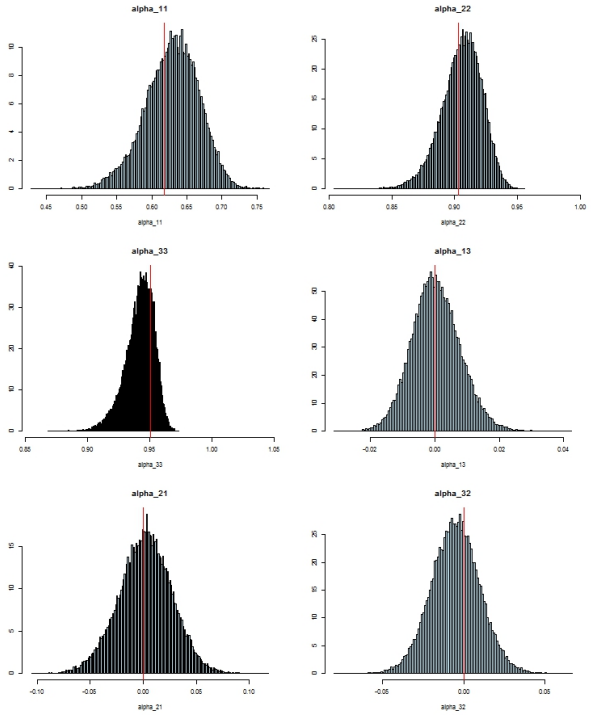

Figure 25: *Posterior distributions of all* $\boldsymbol{A}$ *and* $\tilde{\boldsymbol{A}}$ *parameters. The red lines display the true values.*

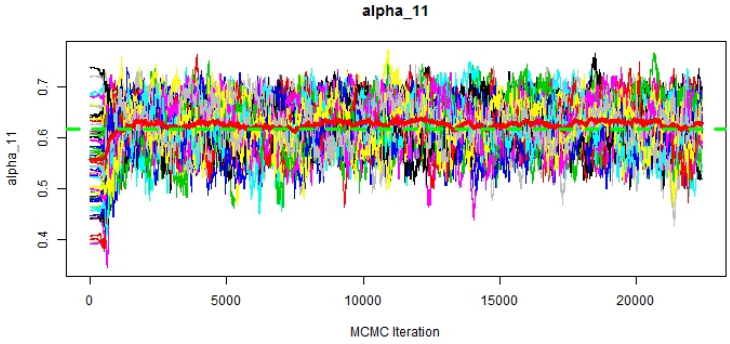

Figure 26: *Trace plot of* α_{11} *using 64 chains in parallel.*

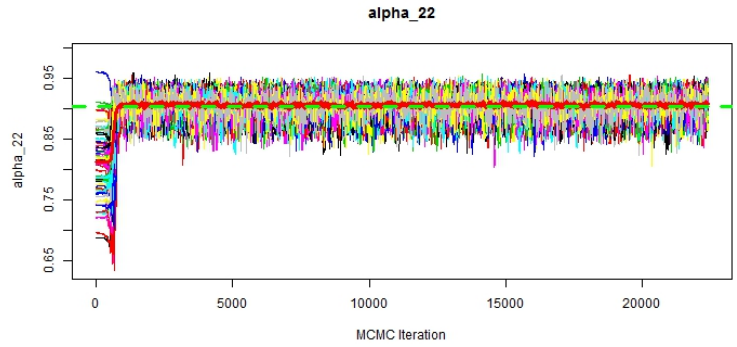

Figure 27: *Trace plot of α_{22} using 64 chains in parallel.*

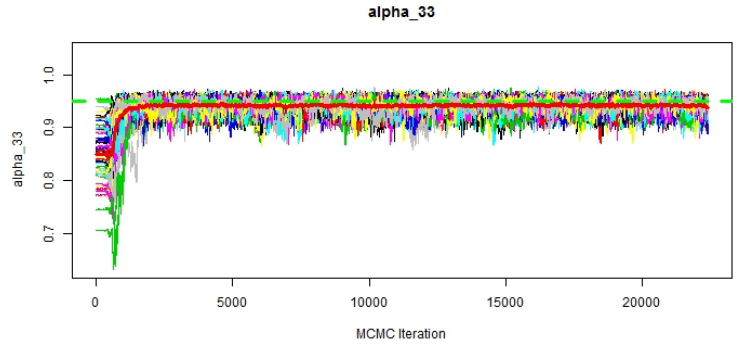

Figure 28: *Trace plot of α_{33} using 64 chains in parallel.*

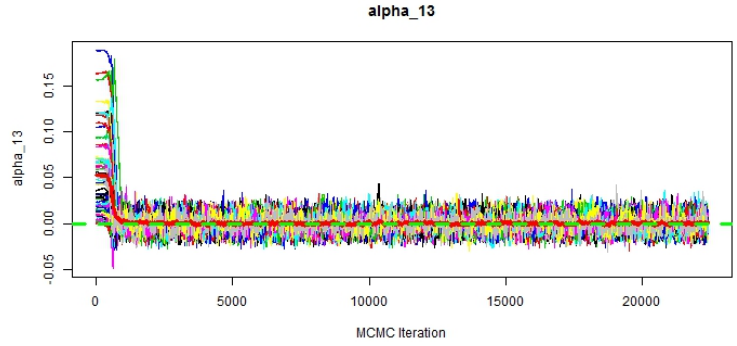

Figure 29: *Trace plot of α_{13} using 64 chains in parallel.*

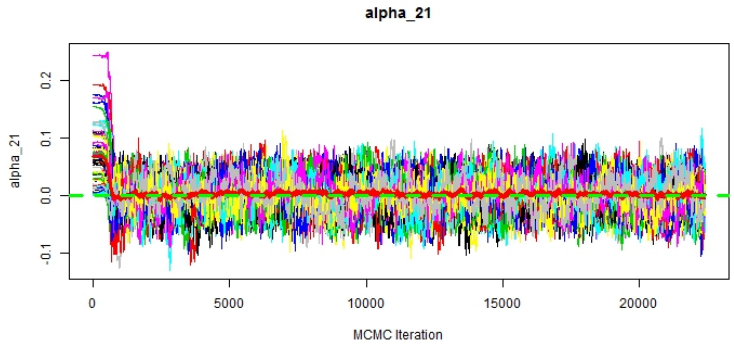

Figure 30: *Trace plot of α_{21} using 64 chains in parallel.*

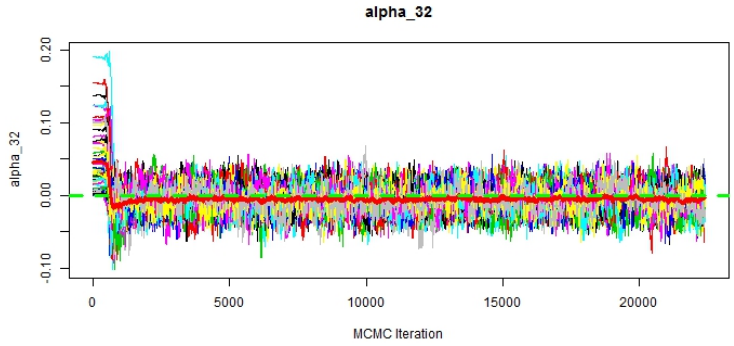

Figure 31: *Trace plot of α_{32} using 64 chains in parallel.*

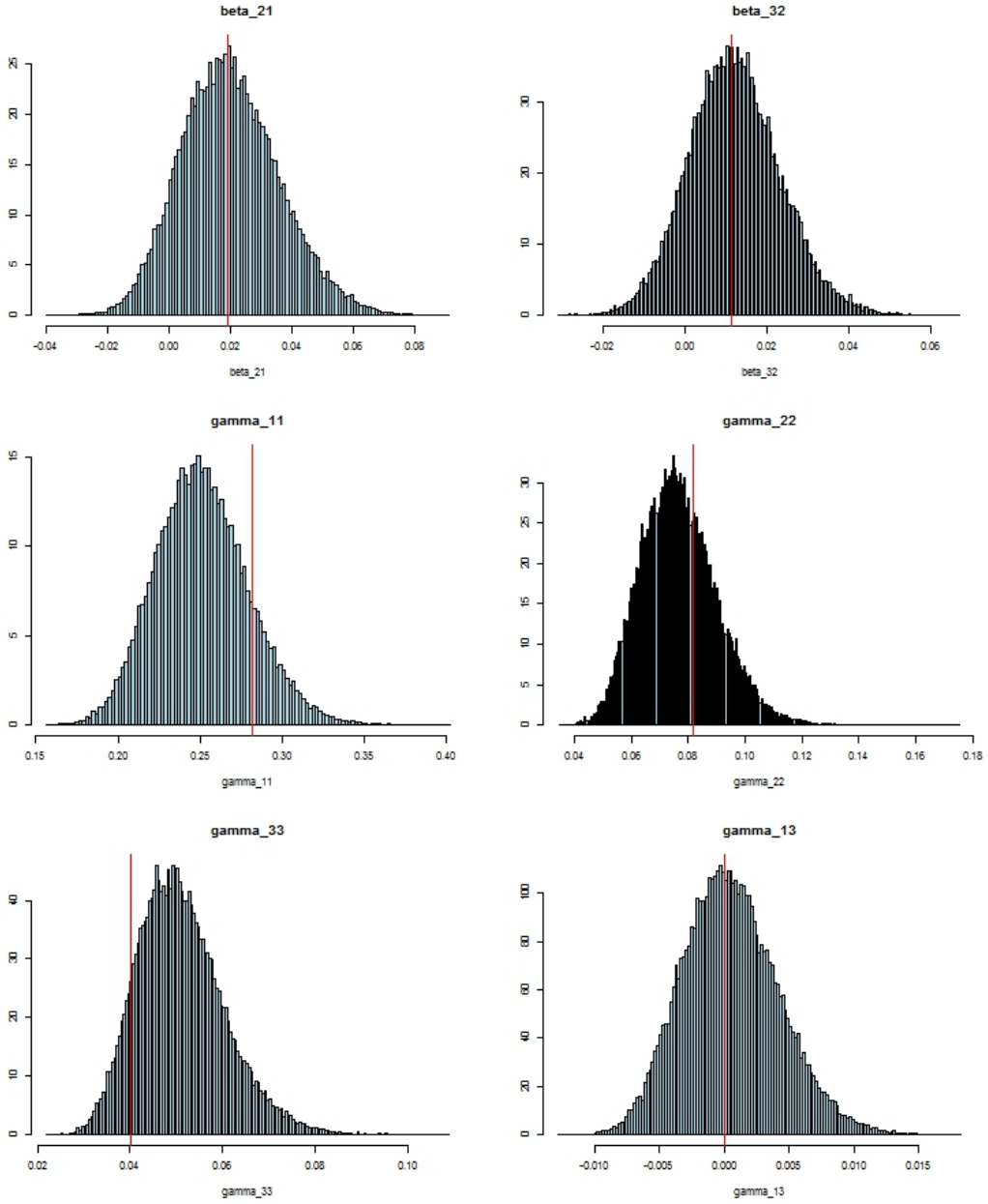

Figure 32: *Posterior distributions of all **B** and **G** parameters. The red lines display the true values.*

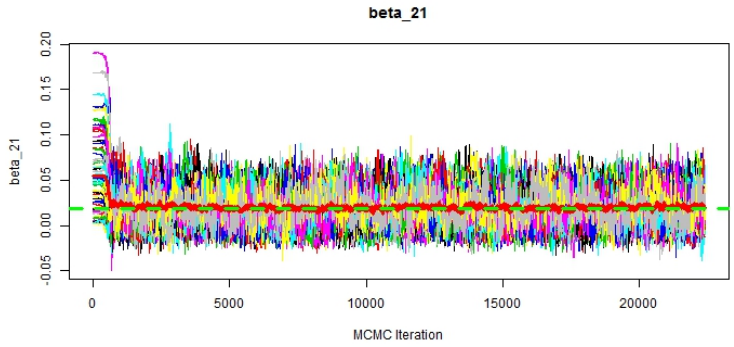

Figure 33: *Trace plot of $\beta 21$ using 64 chains in parallel.*

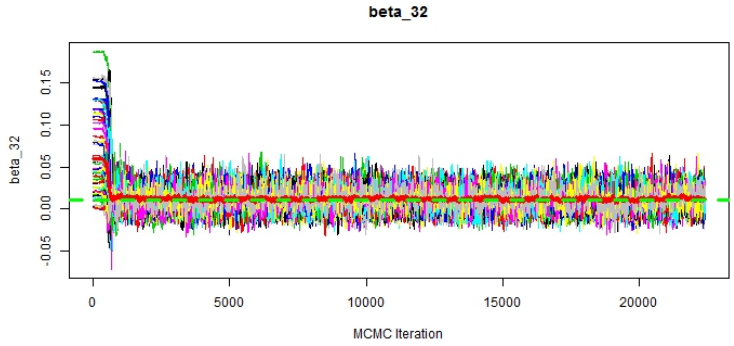

Figure 34: *Trace plot of $\beta 32$ using 64 chains in parallel.*

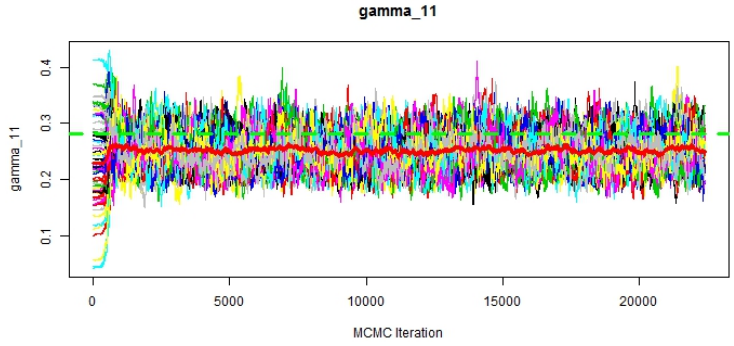

Figure 35: *Trace plot of $\gamma 11$ using 64 chains in parallel.*

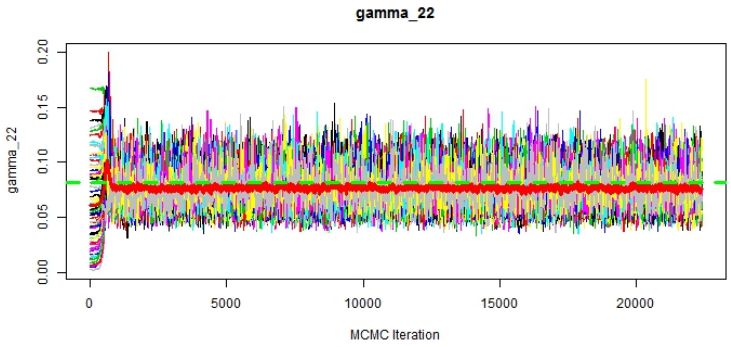

Figure 36: *Trace plot of $\gamma 22$ using 64 chains in parallel.*

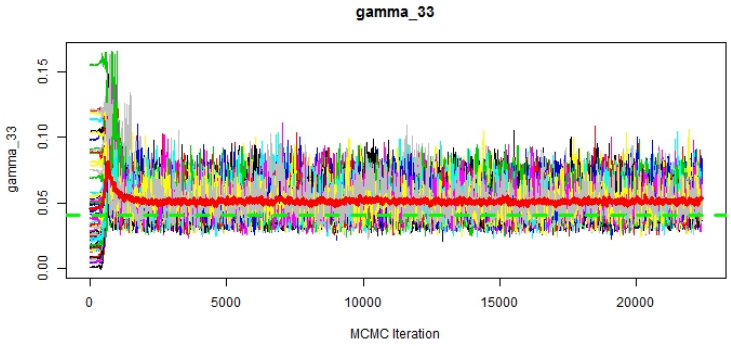

Figure 37: *Trace plot of $\gamma 33$ using 64 chains in parallel.*

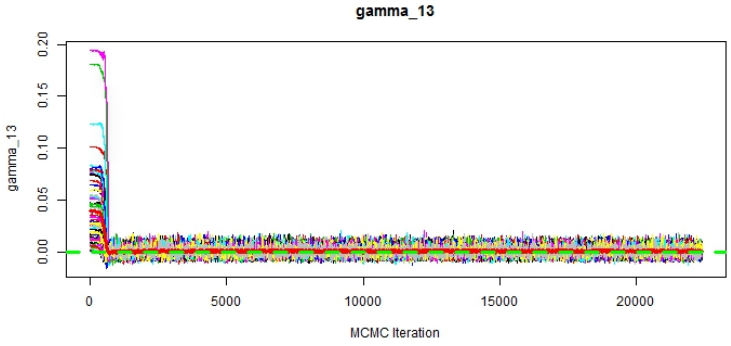

Figure 38: *Trace plot of $\gamma 13$ using 64 chains in parallel.*

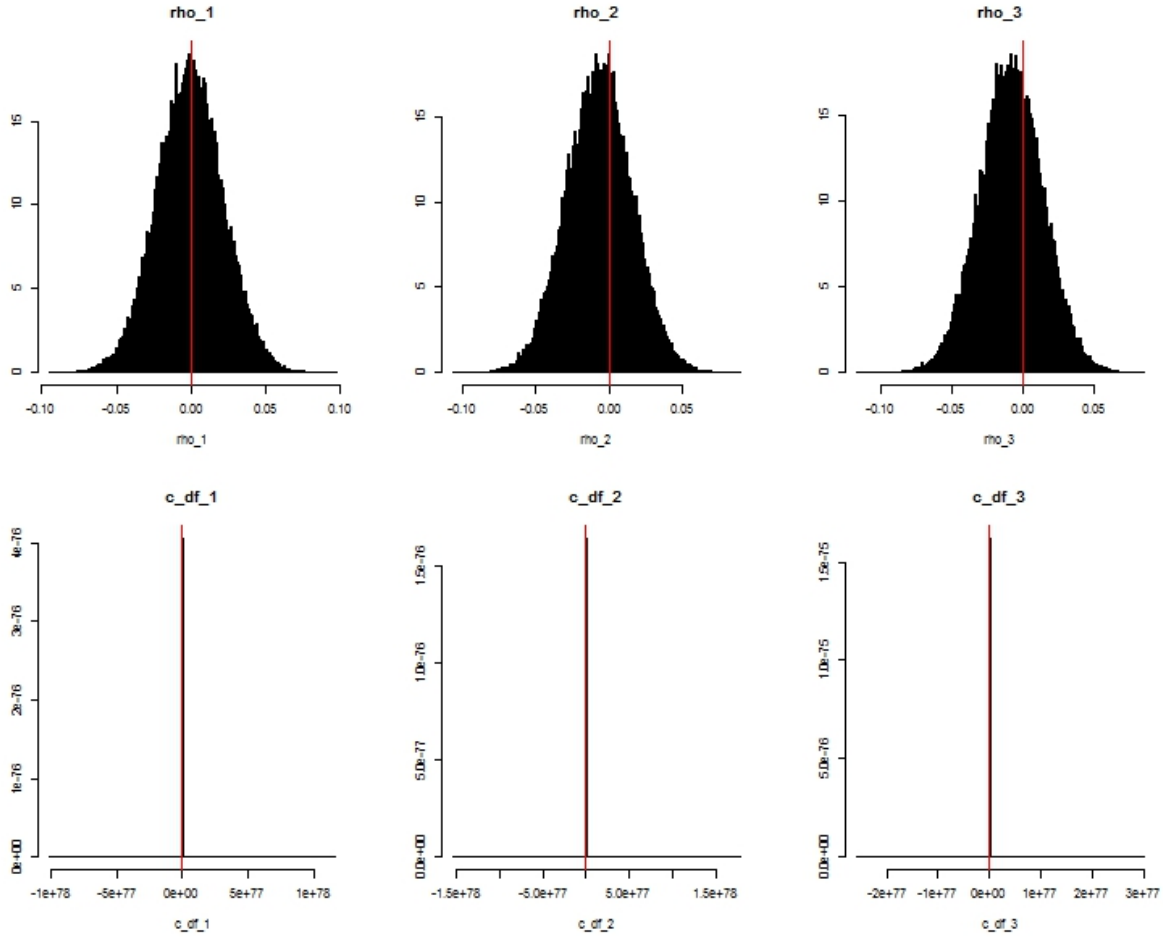

Figure 39: *Posterior distributions of all copula parameters. The red lines display the true values.*

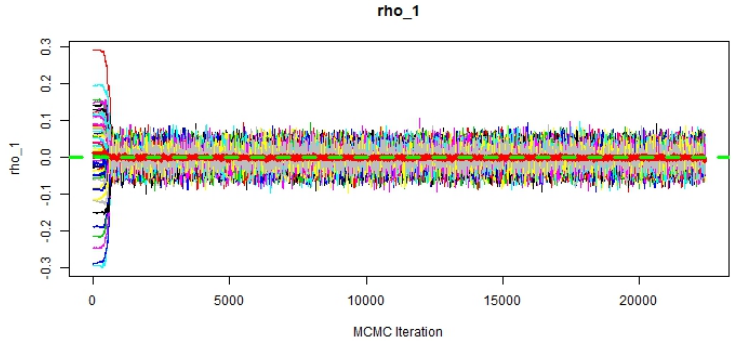

Figure 40: *Trace plot of ρ_1 using 64 chains in parallel.*

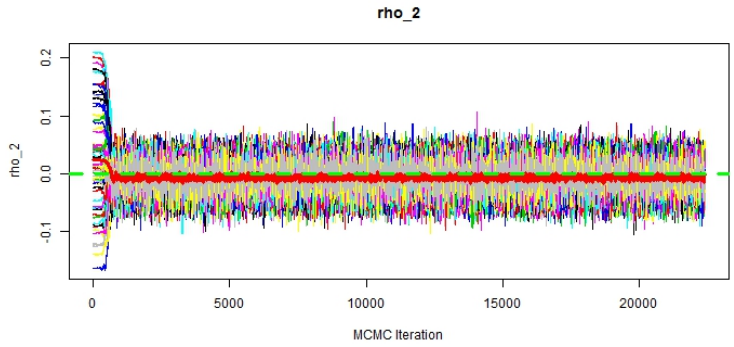

Figure 41: *Trace plot of ρ_2 using 64 chains in parallel.*

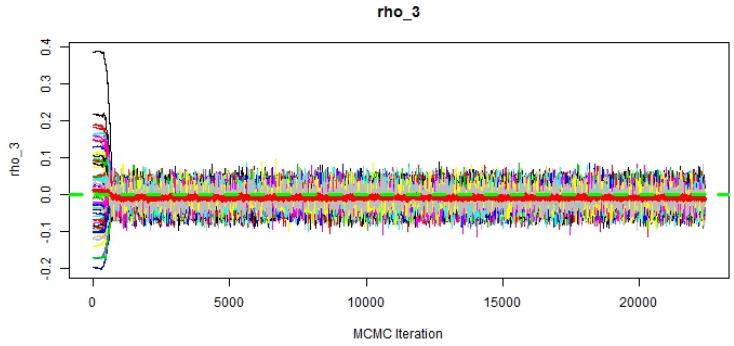

Figure 42: *Trace plot of ρ_3 using 64 chains in parallel.*

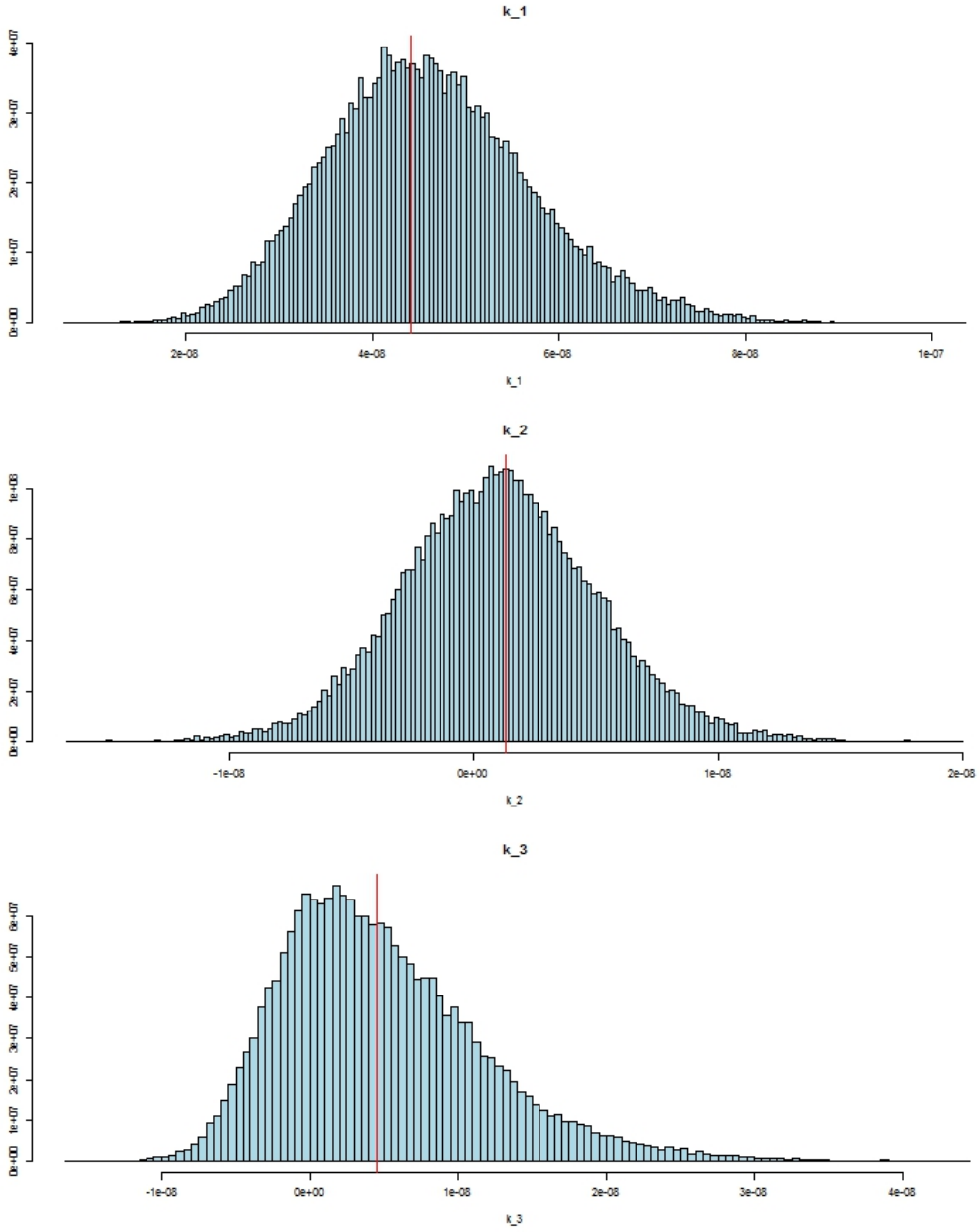

Figure 43: *Posterior distributions of all* ***k*** *parameters. The red lines display the true values.*

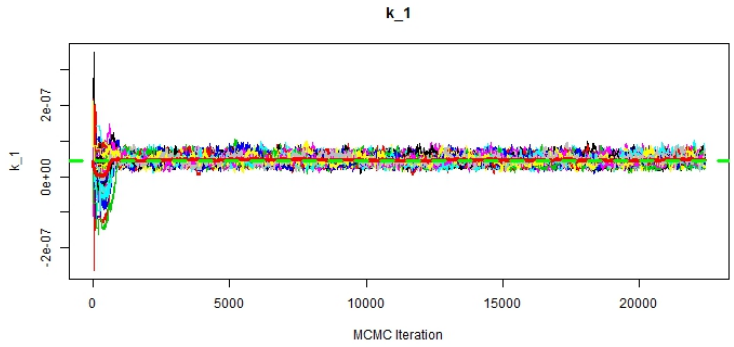

Figure 44: *Trace plot of k_1 using 64 chains in parallel.*

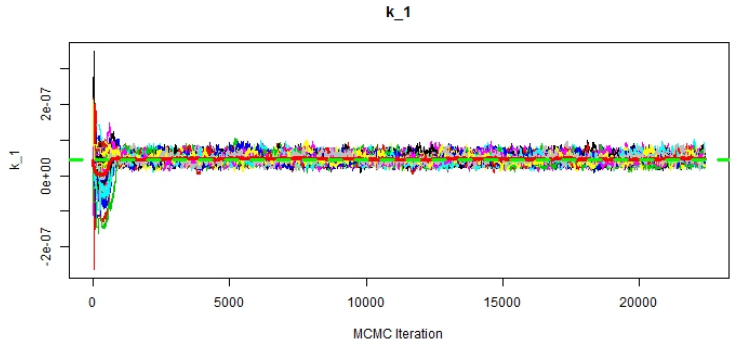

Figure 45: *Trace plot of k_2 using 64 chains in parallel.*

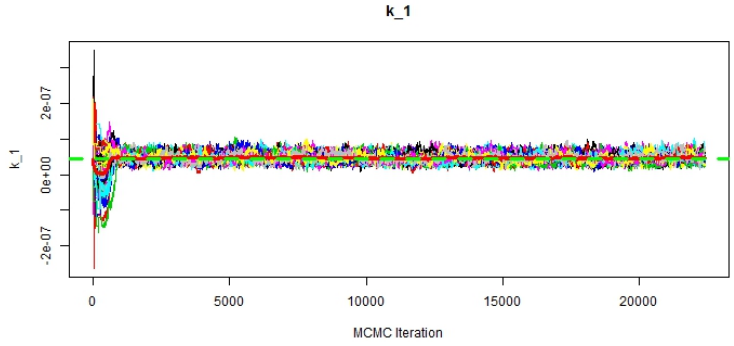

Figure 46: *Trace plot of k_3 using 64 chains in parallel.*

References

Aas, K., Czado, C., Frigessi, A., Bakken, H., 2009. Pair-copula constructions of multiple dependence. Insurance: Mathematics and Economics 44, 182–198. doi:10.1016/j.insmatheco.2007.02.001.

Bedford, T., Cooke, R.M., 2002. Vines–a new graphical model for dependent random variables. The Annals of Statistics 30, 1031–1068. doi:10.1214/aos/1031689016.

Boothe, P., Glassman, D., 1987. The statistical distribution of exchange rates: Empirical evidence and economic implications. Journal of International Economics 22, 297–319. doi:10.1016/S0022-1996(87)80025-9.

Braak, C.J.F.T., 2006. A markov chain monte carlo version of the genetic algorithm differential evolution: Easy bayesian computing for real parameter spaces. Statistics and Computing 16, 239–249. doi:10.1007/s11222-006-8769-1.

Brechmann, E.C., 2010. Truncated and simplified regular vines and their applications: Diplomarbeit. URL: https://mediatum.ub.tum.de/doc/1079285/1079285.pdf.

Brechmann, E.C., Czado, C., Aas, K., 2012. Truncated regular vines in high dimensions with application to financial data. Canadian Journal of Statistics 40, 68–85. doi:10.1002/cjs.10141.

Clements, A.E., Hurn, A.S., Volkov, V.V., 2015. Volatility transmission in global financial markets. Journal of Empirical Finance 32, 3–18. doi:10.1016/j.jempfin.2014.12.002.

Dißmann, J., 2010. Statistical Inference for RegularVines and Application. Ph.D. thesis. Technische Universität München.

Dißmann, J., Brechmann, E.C., Czado, C., Kurowicka, D., 2013. Selecting and estimating regular vine copulae and application to financial returns. Computational Statistics & Data Analysis 59, 52–69. doi:10.1016/j.csda.2012.08.010.

Engle, R.F., Ito, T., Lin, W.L., 1990. Meteor showers or heat waves? heteroskedastic intra-daily volatility in the foreign exchange market. Econometrica 58, 525. doi:10.2307/2938189.

Fleming, M.J., Lopez, J.A., 1999. Heat waves, meteor showers, and trading volume: An analysis of volatility spillovers in the u.s. treasury market. Federal Reserve Bank of New York Staff Reports URL: https://www.newyorkfed.org/medialibrary/media/research/staff_reports/sr82.html.

Gelman, A., Carlin, J.B., Stern, H.S., Rubin, D.B., 2004. Bayesian data analysis. Texts in statistical science. 2. ed. ed., Chapman & Hall, Boca Raton, Fla.

Greenberg, E., 2008. Introduction to Bayesian econometrics. Cambridge University Press, Cambridge.

Gruber, L., Czado, C., 2015. Sequential bayesian model selection of regular vine copulas. Bayesian Analysis 10, 937–963. doi:10.1214/14-BA930.

Lee, T.H., Long, X., 2009. Copula-based multivariate garch model with uncorrelated dependent errors. Journal of Econometrics 150, 207–218. doi:10.1016/j.jeconom.2008.12.008.

Rogalski, R.J., Vinso, J.D., 1978. Empirical properties of foreign exchange rates. Journal of International Business Studies 9, 69–79. URL: http://www.jstor.org/stable/154168.

Savu, C., Trede, M., 2010. Hierarchies of archimedean copulas. Quantitative Finance 10, 295–304. doi:10.1080/14697680902821733.

Schepsmeier, U., Nagler, T., 2017. Vinecopula: Statistical inference of vine copulas. URL: https://CRAN.R-project.org/package=VineCopula.

Schmitz, V., 2003. Copulas and stochastic processes: Zugl.: Aachen, Techn. Hochsch., Diss., 2003. Berichte aus der Mathematik, Shaker, Aachen.

ZIV, 2017. Palma.
https://www.uni-muenster.de/ziv/technik/palma/index.html. URL: https://www.uni-muenster.de/ZIV/Technik/PALMA/index.html.