Matrix Inequality Constraints for Vector Asymmetric Power HEAVY (or GARCH) (or MEM) Models and some New Mixture Formulations

M. Karanasos[†*], Y. Xu[‡]
[†]Brunel University London, UK;[‡] Cardiff University

July 4, 2016

Abstract

In this paper we review and generalize results on the derivation of tractable non-negativity constraints for N-dimensional Heavy/MEM/GARCH systems. These non-negativity conditions are expressed in terms of matrix inequalities which can be solved easily. Numerical examples are included to show the effectiveness of the proposed method.

Keywords:
JEL Classification:

[*]Address for correspondence: Menelaos Karanasos, Economics and Finance, Brunel University London, UB3 3PH, UK; email: menelaos.karanasos@brunel.ac.uk, tel: +44(0)1895265284, fax: +44 (0)1895269770.

1 Introduction

The non-negativity constraints are easily modified, that is they expressed in terms of matrix inequalities which can be solved easily.

We present a useful method for constructing tractable counterparts of the non-negativity constraints derived in Conrad and Karanasos (2010).

It is of considerable interest to investigate whether or not a number of reported estimated N-dimensional HEAVY/MEM/GARCH models satisfy these matrix inequality constraints. Indeed we find that an alarming number of seminar papers report estimated coefficients whose values violate the non-negativity conditions.

We give many examples where well known papers report

Most importantly, we also derive new tractable constraints for the asymmetric versions of these N-dimensional systems.

These allows new tractable constraints to be constructed for asymmetric multivariate processes and thus extends the results in Conrad and Karanasos (2010).

Finally, we propose a new mixture formulation, which is an effective way to relax some of these constraints.

These matrix inequalities are tractable in theory and practice.

Our results are applicable to all three type of multivariate models, that is HEAVY, MEM and GARCH.

The research by Nelson and Cao (1992), He and Teräsvirta (1999), Gourieroux (2007), Tsai and Chan (2007, 2008), Nakatani and Teräsvirta (2008, 2009), underline the theoretical interest in the derivation of such necessary and sufficient conditions. This strand of the literature originated with the seminal work of Nelson (1991a,b).

As we indicated at the outset the standard multivariate extended model requires all the variables to be non-negative. Thus, we need a mechanism to ensure that ....

The matrix inequalities can be practically checked with ease and some of these even effortlessly enforced in estimation.

New theoretical results on multivariate HEAVY/MEM/GARCH models including conditions which ensure the non-negativity of the variables (conditional means/variances).

We propose simple matrix inequalities which incorporate these constraints.

The relevance and the importance of the proposed method is demonstrated with three empirical examples on three different real datasets.

The more parameters there are in the conditional variance equation, the more likely it is that one or more of them will take negative values.

He and Teräsvirta (1999) show that the less severe non-negativity constraints allow more flexibility in the shape of autocorrelation function than the constraints restricting the parameters to be non-negative.

The non-negativity constraint on the coefficients in multivariate GARCH models is only a sufficient condition and may be weakened in certain cases (References).

As noted by Rabemananjara and Zakoian (1993) and Knight and Satchell (2007), non-negativity constraints on the parameters maybe a source of important difficulties in running estimation procedures. With a shock in the past, regardless of the sign, always has a positive effect on the current volatility: the impact increases with the magnitude of the shock. Therefore, cyclical or any non-linear behavior in volatility cannot be taken into account.

We show that these non-negativity constraints are translated into simple matrix inequalities, which are easily handled.

are dealt without difficulty.

these conditions are not only sufficient but necessary as well

rewrite (or reformulate) these set of constraints in a more (compact) intuitive visual form, as matrix inequalities

express the constraints in a more efficient way

allows us to communicate such constraints in a more user friendly way

the implications of these constraints can be seen explicitly.

We hope that our paper will create more awareness of the existence of these constraints and increase their usage

inequality constraints that require the power transformed conditional means or variances to be almost surely non-negative at all t.

A critical question is: "In Italics"

To answer this we first ....

One main concern is that the existence of such constraints is often ignored by researchers.

As a results many of the aforementioned seminal papers (i.e., Cipollini et al.) report estimated parameters with negative values, which frequently violate the inequality constraints.

Researchers should recognise that the existence of these constraints impose severe limitations

In practice, these constraints are difficult to be satisfied. In other words, and as already mentioned earlier they are commonly violated.

In practice these constraints are rarely fullfilled

We propose a new mixture formulation in order to eliminate some of these constraints. This can be done by modelling a number of the conditional variances (but not all) using exponential models.

the more general assymetric setting increases considerable the number of constraints

by replacing power transformations with logarithmic ones we cutting down on the dimensions of the inequality constraints

impose severe restrictions on the parameter space

2 The Model

2.1 Notation

Throughout the paper will adhere to the following notation. $\mathcal{F}_{t-1}^{XF}$ is the filtration generated by all available information through time $t-1$. We will use $\mathcal{F}_{t-1}^{HF}$ $(X = H)$ for the high frequency past data, i.e., for the case of the realized measure, or $\mathcal{F}_{t-1}^{LoF}$ $(X = Lo)$ for the low frequency past data, i.e., for the case of the close-to-close returns. Hereafter, for notational convenience we will drop the superscript XF.

We will use upper(lower) case boldface symbols to refer to Matrices(vectors). That is, $\mathbf{Y} = [y_{ij}]_{i,j=1,\ldots,N}$ is an $N \times N$ matrix, $\mathbf{y} = [y_i]_{i=1,\ldots,N}$ is an $N \times 1$ column vector and $diag[\mathbf{y}]$ denotes a diagonal matrix with elements y_i, $i = 1, \ldots, N$. The jth column and ith row of $\mathbf{Y}$ are written as $\mathbf{y}^j = [y_{ij}]_{i=1,\ldots,N}$ and $\mathbf{y}_i'$ (where $\mathbf{y}_i = [y_{ij}]_{j=1,\ldots,N}$) respectively. Let $Y_{ij}(L)$ be a polynomial of order N. Then $\mathbf{Y}(L) = [Y_{ij}(L)]_{i,j=1,\ldots,N}$ indicates an $N \times N$ matrix polynomial in the lag operator L. Also $\mathbf{I}_N$ is the identity matrix of order N and 0_N is a square null matrix of order N (hereafter, we will drop the subscript N). $\mathbf{J}_i$ ($\mathbf{J}^j$) is an $N \times N$ binary matrix that has ones in its ith row (jth column) and zeros elsewhere.

Using standard notation, $\mathbf{Y}'$ and $\mathbf{Y}^{-1}$ are the transpose and the inverse of the square matrix $\mathbf{Y}$. $\det[\mathbf{Y}(L)]$ and $adj[\mathbf{Y}(L)]$ denotes the determinant and the adjoint of the matrix polynomial $\mathbf{Y}(L)$, respectively. That is, $adj[\mathbf{Y}(L)] = [Y_{ij}^{\{a\}}(L)]_{i,j=1,\ldots,N}$ with $Y_{ij}^{\{a\}}(L) = (-1)^{i+j}\det[\mathbf{Y}_j^i(L)]$ where $\mathbf{Y}_j^i(L)$ is the $\mathbf{Y}(L)$ matrix without its jth row and ith column. In other words, $Y_{ij}^{\{a\}}(L)$ is the cofactor of the jith element of $\mathbf{Y}(L)$.

The elementwise expectation operator is denoted by $\mathbb{E}$, i.e., $\mathbb{E}(\mathbf{Y}) = [\mathbb{E}(y_{ij})]_{i,j=1,\ldots,N}$, whereas $\mathbf{Y}^{\wedge k} = [y_{ij}^k]_{i,j=1,\ldots,N}$ is the elementwise exponentiation. $abs[\mathbf{Y}] = [|y_{ij}|]_{i,j=1,\ldots,N}$ refers to the elementwise absolute value of $\mathbf{Y}$. The inequality $\mathbf{Y} > \mathbf{0}$ means that all elements of $\mathbf{Y}$ are positive real numbers. Further, $\max[\mathbf{Y}]$ indicates the largest element of the matrix $\mathbf{Y}$ and $\mathbf{Y}^k = \prod_{i=1}^k \mathbf{Y}$ means that the matrix $\mathbf{Y}$ is raised to the power of k. ($\mathbb{E}(\mathbf{Y} \,|\mathcal{F}_{t-1}\,)$ denotes the elementwise (conditional on time $t-1$) expectation operator). Moreover, let $\mathbf{Y}_{2\otimes} = \mathbf{Y} \otimes \mathbf{Y}$, $\mathbf{Y}_{\otimes I} = \mathbf{Y} \otimes \mathbf{I}$ and $\mathbf{Y}_{I\otimes} = \mathbf{I} \otimes \mathbf{Y}$ where $\otimes$ is the Kronecker product of two matrices. $vec(\mathbf{Y})$ is a vector in which the columns of the matrix $\mathbf{Y}$ are stacked one underneath the other

2.2 The Asymmetric Power Specification

Consider the zero conditional mean, $\mathbb{E}(\boldsymbol{\varepsilon}_t \,|\mathcal{F}_{t-1}\,) = 0$ N-dimensional vector process, $\boldsymbol{\varepsilon}_t = [|\varepsilon_{it}|^{\delta_i}]_{i=1,\ldots,N}$ where $\mathcal{F}_{t-1} = \sigma(\boldsymbol{\varepsilon}_{t-1}, \boldsymbol{\varepsilon}_{t-2}, \ldots)$ and $\delta_i \in (0, \infty)\ \forall\ i$. We assume that the noise vector $\boldsymbol{\varepsilon}_t$ is characterized by the relation

$$\boldsymbol{\varepsilon}_t = \mathbf{E}_t \boldsymbol{\sigma}_t \text{ or } \varepsilon_{it} = e_{it}\sigma_{it},\ i = 1, \ldots, N, \tag{1}$$

where $\boldsymbol{\sigma}_t = [\sigma_{it}^{\delta_i}]_{i=1,\ldots,N}$ is $\mathcal{F}_{t-1}$ measurable, $\mathbf{e}_t = [|e_{it}|^{\delta_i}]_{i=1,\ldots,N}$ (hereafter we will drop the subscript for notational convenience) and $\mathbf{E}_t = diag\{\mathbf{e}_t\}$. Let also $\boldsymbol{\varepsilon}_t^* = [\varepsilon_{it}]$, $\mathbf{e}_t^* = [e_{it}]$ with $\mathbf{E}_t^* = diag\{\mathbf{e}_t^*\}$ and $\boldsymbol{\sigma}_t^* = [\sigma_{it}]$. Then $\boldsymbol{\varepsilon}_t^* = \mathbf{E}_t^* \boldsymbol{\sigma}_t^*$ or $\varepsilon_{it} = e_{it}\sigma_{it}$, $i = 1, \ldots, N$. The stochastic vector $\mathbf{e}_t^* = [e_{it}]$ is independent and identically distributed (i.i.d) with mean zero and positive definite time varying correlation matrix $\mathbf{R}_t = [\rho_{ij,t}]_{i,j=1,\ldots,N}$ with $\rho_{ii} = 1$. From the above equation it follows that $\boldsymbol{\Sigma}_t = \mathbb{E}[\boldsymbol{\varepsilon}_t^*(\boldsymbol{\varepsilon}_t^*)' \,|\mathcal{F}_{t-1}] = diag[\boldsymbol{\sigma}_t^\star]\mathbf{R}_t diag[\boldsymbol{\sigma}_t^\star]$, where $\boldsymbol{\sigma}_t^\star$ is the value of $\boldsymbol{\sigma}_t$ when $\delta_i = 0.5\ \forall\ i$. Notice that $\rho_{ij,t} = \mathbb{E}(e_{it}e_{jt} \,|\mathcal{F}_{t-1}\,)$, for $i \neq j$, is the conditional correlation of ε_{it} and ε_{jt}, that is $\rho_{ij,t} = \frac{\sigma_{ij,t}}{\sqrt{\sigma_{it}}\sqrt{\sigma_{jt}}}$ with $\sigma_{ij,t} = \mathbb{E}(\varepsilon_{it}\varepsilon_{jt} \,|\mathcal{F}_{t-1}\,)$.

2.3 The Asymmetric Power Extension

As pointed out by Conrad and Karanasos (2010) a major problem in specifying a valid multivariate process lies in choosing appropriate parametric specifications for $\boldsymbol{\sigma}_t^*$ such that $\boldsymbol{\Sigma}_t$ is positive definite almost surely for all t. Positive definiteness of $\boldsymbol{\Sigma}_t$ follows if, in addition to the correlation matrix $\mathbf{R}_t$ being positive definite, the conditional variances of ε_{it} or their power transformations, $\sigma_{it}^{\delta_i/2}$, $i = 1, \ldots, N$, are positive as well.

In what follows we will consider two 'different' asymmetric power models. The first type of asymmetry was introduced by Glosten et. al. (1993) and we will refer to it as model 1 whereas the second type of asymmetry was introduced by Ding et. al. (1993) and we will refer to it as model 2.

The N-dimensional semi-unrestricted extended asymmetric power (SUE-AP) model of order $(1,1)$[1] consists of the following equations:

$$\sigma_{it}^{\delta_i/2} = \omega_i + \sum_{j=1}^{N}(\alpha_{ij} + \gamma_{ij}s_{j,t-1})\,|\varepsilon_{j,t-1}|^{\delta_j} + \sum_{j=1}^{N}\beta_{ij}\sigma_{j,t-1}^{\delta_j/2},\ i = 1,\dots,N, \tag{2}$$

where L is the lag operator and s_{jt} is a dummy variable that takes the value $s_{jt} = 1$ if $e_{jt} < 0$, 0 otherwise, that is $s_{jt} = [1 - sign(e_{jt})1]/2$. This can be either a HEAVY or a GARCH model or a MEM[2] It can be expressed/interpreted as an N-dimensional system with shocks and volatility spillovers:

$$(\mathbf{I} - \mathbf{B}L)\boldsymbol{\sigma}_t = \boldsymbol{\omega}+L\mathbf{A}_t\boldsymbol{\varepsilon}_t, \tag{3}$$

where $\mathbf{B}$ is an $N \times N$ full matrix: $\mathbf{B} = [\beta_{ij}]$, and its cross diagonal elements capture the volatility spillovers; $\boldsymbol{\omega} = [\omega_i]$ is an $N \times 1$ vector of the drifts; $\mathbf{A}_t= \mathbf{A} + \mathbf{\Gamma}_t$, where $\mathbf{A} = [\alpha_{ij}]$ and $\mathbf{\Gamma}_t= [\boldsymbol{\gamma}_{ij}s_{jt}]$, are $N \times N$ full matrices: the cross diagonal elements of $(\mathbf{\Gamma}_t)\mathbf{A}$ capture the (asymmetric) shock spillovers. Note that $\mathbf{\Gamma}_t$ can be written as $\mathbf{\Gamma}_t= \mathbf{\Gamma}diag\{\mathbf{s}_t\}$where $\boldsymbol{\xi}_t = [s_{it}]$. The above model is termed semi-unrestricted extended because as we will see below some of the elements of the $\mathbf{B}$ matrix (including some of the off-diagonal ones) but not all are allowed to take not only positive but negative values as well. That is, as in Conrad and Karanasos (2010) we consider a formulation of the extended model that allows for feedback effects between the volatilities, which can be of either sign, positive or negative.[3] However, on the other hand all the elements of the $\mathbf{A}$ and $\mathbf{\Gamma}$ matrices are restricted to be non-negative (see Theorem 4 below). Therefore, we will also introduce another model which allows for negative (asymmetric as well) shock effects either own or cross ones. As pointed out by Conrad and Karanasos (2010) a crucial problem concerns the identification of necessary and sufficient conditions for the unrestricted model to have a positive definite conditional covariance matrix. Note that, if there are no asymmetries, that is $\mathbf{\Gamma} = 0$, then the model reduces to the symmetric power one:

$$(\mathbf{I} - \mathbf{B}L)\boldsymbol{\sigma}_t = \boldsymbol{\omega} + \mathbf{A}L\boldsymbol{\varepsilon}_t, \tag{4}$$

which is further reduced to the benchmark model examined in Conrad and Karanasos (2010) if $\delta_i = 2$, for all i.

3 Matrix Inequality Constraints

To keep this article relatively self-contained we briefly review the main theoretical results of Conrad and Karanasos (2010) on the derivation of necessary and sufficient conditions, which ensure that the time varying covariance matrix $\mathbf{\Sigma}$ of an N-dimensional symmetric HEAVY/MEM/GARCH system is positive definite almost surely for all t. In this Section we give an outline of the second main step in the derivation of such (necessary and sufficient) conditions. The first step, that is the 'univariate' representations, which each conditional variance (or mean in the case of the MEM) admits, is given in the Additional Appendix as Lemma **??**. The second step are the infinite expansions, in terms of convolutions of infinite-order kernels and corresponding power transformed errors, of the aforementioned 'univariate' representations.The latter

[1] In what follows for notational simplicity we will drop the order of the model if it is $(1,1)$. We will refer to this model with the acronym SUE-AP

[2] The acronym HEAVY (High FrEquency bAsed VolatilitY) was introduced by Shephard and Sheppard (2010). For example, in the bivariate context the two variables can be the conditional variances of stock returns and of the signed square rooted (SSR) realized measure (realised variance). The HEAVY formulation parallels the GARCH formulation. It is also very similar to the bivariate multiplicative error model (MEM). In this case the two variables are the conditional expectations of the squared returns and of the realised measure. Therefore we will use the three terms, HEAVY, MEM, GARCH, interchangeably.

[3] In the symmetric restricted extended formulation (see Jeantheau, 1998 and Ling and McAleer, 2003) all the elements of the $\mathbf{B}$ and $\mathbf{A}$ matrices are allowed to take only positive values. As pointed out by Conrad and Karanasos (2010) the assumption that only positive feedback is allowed for its tempting because positive constants and parameter matrices with non-negative coefficients are a sufficient condition for the positive definiteness of the conditional covariance matrix in the extendend formulation.

two steps constitute the main steps in the derivation of Theorem 1 in Conrad and Karanasos (2010), which we state as Proposition **??** in the Additional Appendix.

After the two main steps in the derivation of the aforementioned Proposition are presented for completeness, then our more general result follows on formal grounds. That is we express the non-negativity constraints for the symmetric system as matrix inequalities (Proposition 1 below). A natural extension of this Proposition is the generalization of the results to the asymmetric case. This is developed in Section 4.

3.1 Infinite Order Expansion

In this Section we will introduce a useful lemma. Having derived the 'univariate' representation we use it to obtain the SUE-P infinite-order expansion of each power transformed conditional variance (or mean in the case of the MEM model) in terms of convolutions of HEAVY/MEM/GARCH kernels and corresponding power transformed errors.

Firts, set

$$\beta(L) = 1 - \sum\nolimits_{i=1}^{N} \beta_i L^i = \prod\nolimits_{i=1}^{N} (1 - \phi_i L) = \det[\mathbf{I} - \mathbf{B}L], \tag{5}$$

which since we have assumed (see the Assumptions in the Appendix) that $\beta_N \neq 0$, it is a scalar polynomial of order N; ϕ_i are the roots of $\beta(z^{-1})$. Define the square matrix polynomial

$$\boldsymbol{\alpha}(L) = [a_{ij}(L)] = adj[\mathbf{I} - \mathbf{B}L]\mathbf{A}, \tag{6}$$

with $a_{ij}(L) = \sum\nolimits_{n=1}^{N} a_{ij}^{(n)} L^n$ (where the upper script with parenthesis denotes an index). Since we have assumed that $a_{ij}^{(N)} \neq 0$ for all $i, j = 1, \ldots, N$, the scalar polynomials $a_{ij}(L)$ are of the order N. Finally, define

$$\boldsymbol{\mu} = [\mu_i] = adj[\mathbf{I} - \mathbf{B}]\boldsymbol{\omega}. \tag{7}$$

To ease the following explanations some additional notation is needed. Let $\mathbf{\Psi}(L) = [\Psi_{ij}(L)]_{i,j=1,\ldots,N} = \boldsymbol{\alpha}(L)/\beta(L)$ with $\Psi_{ij}(L) = \Sigma_{k=1}^{\infty} \psi_{ij}^{(k)} L^k = \alpha_{ij}(L)/\beta(L)$ or equivalently $\mathbf{\Psi}(L) = \sum_{k=1}^{\infty} \mathbf{\Psi}_k L^k$, with $\mathbf{\Psi}_k = [\psi_{ij}^{(k)}]$.

Lemma 1 *Let Assumptions (A1) and (A2) be satisfied. Then eq. (**??**) can be rewritten as an SUE-P process of infinite order:*

$$\begin{aligned} \boldsymbol{\sigma}_t &= \frac{\boldsymbol{\mu}}{\beta(1)} + \mathbf{\Psi}(L)\boldsymbol{\varepsilon}_t \; or \\ \sigma_{it}^{\delta_i/2} &= \frac{\mu_i}{\beta(1)} + \sum_{j=1}^{N} \Psi_{ij}(L) \left|\varepsilon_{j,t}\right|^{\delta_j}. \end{aligned} \tag{8}$$

(see also Lemma 2 in Conrad and Karanasos, 2010). Here, each $\Psi_{ij}(L)$ can be thought of as an infinite-order kernel of a SUE-P- model of the order (N, N). Clearly, for the N-dimensional process in eq. (3) to be well-defined and the N (powered transformed) conditional variances to be positive almost surely for all t, all the constants μ_i must be positive and all the $\psi_{ij}^{(k)}$ coefficients in the 'univariate' representation of infinite order, that is eq. (8), must be non-negative: $\psi_{ij}^{(k)} \geq 0$, $i, j = 1, \ldots, N$ for $k = 1, 2, \ldots$

The non-negativity of the (powered transformed) variances is guaranteed if and only if all the kernels are non-negative, i.e., if the *infinite* number of coefficients in the infinite-order expansions of the N^2 kernels are non-negative. For this, one should express these coefficients as functions of the parameters of the original process. It is then can be shown that checking a *finite* number of inequality constraints on these parameters ensures the non-negativity of all HEAVY/MEM/GARCH kernels of the SUE-P model (see Conrad and Karanasos, 2010 who gave special attention only to the bivariate case of order $1,1$; and Proposition **??** in the Additional Appendix).

Alternative Infinite Representation

In this Section we make a general observation that will be applied tactically later on. That is, we shall make use of the following Lemma. The infinite-order expansion have been presented in Section 3.1. Here we present an alternative form for such an expansion.

Lemma 2 *Let Assumptions (A1) and (A2) be satisfied and all the roots ϕ_n, $n = 1, \dots N$, be distinct. Then eq. (4) can be rewritten in an alternative form as:*

$$\boldsymbol{\sigma}_t = \frac{\boldsymbol{\mu}}{\beta(1)} + \sum_{k=1}^{\infty} \mathbf{B}^{k-1}\mathbf{A}L^k\boldsymbol{\varepsilon}_t.$$

The above Lemma follows from Lemma (1) since $(\mathbf{I}-\mathbf{B}L)^{-1}\boldsymbol{\omega} = \frac{adj[\mathbf{I}-\mathbf{B}]\boldsymbol{\omega}}{\det[\mathbf{I}-\mathbf{B}]} = \frac{\boldsymbol{\mu}}{\beta(1)}$ and $\boldsymbol{\Psi}(L) = \frac{\boldsymbol{\alpha}(L)}{\beta(L)} = \frac{adj[\mathbf{I}-\mathbf{B}L]\mathbf{A}}{\beta(L)} = [\mathbf{I}-\mathbf{B}L]^{-1}\mathbf{A} = \sum_{k=1}^{\infty} \mathbf{B}^{k-1}\mathbf{A}L^{k-1}$ and hence $\boldsymbol{\Psi}_k = [\psi_{ij}^{(k)}]_{i,j=1,\dots,N} = \mathbf{B}^{k-1}\mathbf{A}$.

3.2 Tractable Expressions

In this Section we will show that the non-negativity conditions in Proposition (??) in the Additional Appendix can be expressed as simple inequalities involving $N \times N$ matrices. Our constraints in terms of inequalities of square matrices of order N are algorithmically solvable fast enough to be practically relevant. In other words, the results of this Section makes the problem of non-negativity conditions for N-dimensional SUE-P systems easily solvable and downright tractable.

Next we will present in the following Proposition our tractable non-negativity constraints. The constraints are expressed in terms of matrix inequalities. These inequalities can be easily computed fast enough to make them practical. But before we do that we will introduce some further notation. If $\mathbf{Y}^{(n)} = [y_{ij}^{(n)}]_{i,j=1,\dots,N}$, $n = 1, \dots, N$ (recall that the upper script with parenthesis denotes an index), then $\max[\mathbf{Y}^{(n)}] = \max(y_{ij}^{(n)})$ for $i, j = 1, \dots, N$. In other words, $\max[\mathbf{Y}^{(n)}]$ is the largest element of all the N^2 elements of $\mathbf{Y}^{(n)}$. In addition, $\max\limits_{1 \le n \le N}\{\mathbf{Y}^{(n)}\} = \max\{\mathbf{Y}^{(\mathbf{1})}, \dots, \mathbf{Y}^{(\mathbf{N})}\} = [\max\limits_{1 \le n \le N} \mathbf{y}_{ij}^{(n)}]_{i,j=1,\dots,N}$ is a matrix whose ijth element is the largest of the N $y_{ij}^{(n)}$ elements.

Proposition 1 *Consider the N-dimensional vector SUE-P model in eq. (4) and let Assumptions (A1)-(A2) be satisfied and all the roots ϕ_n, $n = 1, \dots N$, be distinct. Then, the necessary and sufficient conditions for $\sigma_{it}^{\delta_i/2} > 0$, $i = 1, \dots, N$ for all t given in Proposition (**??**) can be expressed as:*
(a) $\boldsymbol{\mu} = adj[\mathbf{I}-\mathbf{B}]\boldsymbol{\omega} > \mathbf{0}$

(b)
ϕ_1 *is real, and* $\phi_1 > 0$, *(C1)*
$adj[\mathbf{I}\phi_1 - \mathbf{B}]\mathbf{A} > \mathbf{0}$, *(C2*)*
$\boldsymbol{\Psi}^{(k)} = \mathbf{B}^{(k-1)}\mathbf{A} \overset{(k \le \kappa_{ij})}{\succeq} \mathbf{0}$, *for* $k = 1, \dots, \kappa$, *(C3*)*

(the symbol $\overset{(k \le \kappa_{ij})}{\succeq}$ means that if $k > \kappa_{ij}$ then the ijth scalar inequality of the matrix inequality $\boldsymbol{\Psi}^{(k)} \ge \mathbf{0}$ should be disregarded)[4] *where* $\kappa = \max[\mathbf{K}]$, $\mathbf{K} = [\kappa_{ij}]_{i,j=1,\dots,N}$, *and κ_{ij} is the smallest integer greater than or equal to* $\max\{0, \varphi_{ij}\}$ *with*

$$\boldsymbol{\Phi} = [\varphi_{ij}] = \{\log[\mathbf{H}^{(1)}] - \log[(N-1)\mathbf{H}]\}[\log(|\phi_2|) - \log(|\phi_1|)]^{-1},$$

[4] Let the ijth element of $\mathbf{B}^{\bar{\wedge} k_{ij}}$ be denoted by $\beta_{ij}^{[k_{ij}]}$, $i, j = 1, \dots, N$, and construct the $N \times N$ matrix $\mathbf{B}^{[k_{ij}]} = [\beta_{ij}^{[k_{ij}]}]_{j=1,\dots,N}$. Then condition (C3) in Proposition (??) in the Additional Appendix for $k_{ij} = 1, \dots, \kappa_{ij}$ amounts to: $\mathbf{B}^{[k_{ij}]}\mathbf{A} > \mathbf{0}$ for all $i, j = 1, \dots, N$. For computational convenience we write it instead as in condition (C3*): $\boldsymbol{\Psi}^{(k)} = \mathbf{B}^{\bar{\wedge}(k-1)}\mathbf{A} \overset{(k \le \kappa_{ij})}{\succeq} \mathbf{0}$.

where

$$\mathbf{H} = [\eta_{ij}] = \max_{2\leq n\leq N}\{\mathbf{H}^{(n)}\}=[\max_{2\leq n\leq N}\eta_{ij}^{(n)}]_{i,j=1,\ldots,N},$$

and

$$\mathbf{H}^{(n)} = [\eta_{ij}^{(n)}] = abs\left[\frac{adj[\mathbf{I}\phi_n - \mathbf{B}]\mathbf{A}}{\sum_{j=1}^{N} j\beta_j(\phi_n)^{N-j}}\right], \ \ 1 \leq n \leq N.$$

Proposition 1 follows directly from Proposition ?? (in the Additional Appendix) and Lemma 2. Interestingly, we only have to check: i) from condition (a) if all the N elements of the $N \times 1$ vector $adj[\mathbf{I}-\mathbf{B}]\boldsymbol{\omega}$ are positive, ii) from condition (C2*) if all the N^2elements of the $N\times N$ matrix $adj[\mathbf{I}\phi_1-\mathbf{B}]\mathbf{A}$ are positive, and iii) from condition (C3*) if all the N^2 elements of each of the κ $N\times N$ matrices: $\mathbf{B}^{(k-1)}\mathbf{A}$ are non-negative. In other words, our conditions are algorithmically solvable fast enough to be practical relevant. It is very easy for the practitioner to check if these matrix inequality constraints are satisfied.

Next we show that the matrix inequalities are easily represented in terms of scalar inequalities as well. As a last stage before we do that, however, we will introduce some additional notation. First, denote $\mathbf{B}^{\star} = [\beta_{ij}^{\star}] = \mathbf{I} - \mathbf{B}$, then condition (a) in Proposition 1 implies that:

$$\sum_{m=1}^{N}\beta_{im}^{\star\{a\}}\omega_m > 0, \text{ for all } i = 1,\ldots,N,$$

where $\beta_{im}^{\star\{a\}} = (-1)^{i+m}\det[(\mathbf{B}^{\star})_m^i]$, and the latter term is the determinant of the matrix obtained by deleting the mth row and the ith column from $\mathbf{B}^{\star}$. Similarly, denote $\mathbf{B}^{*} = [\beta_{ij}^{*}] = \mathbf{I}\phi_1 - \mathbf{B}$ then condition (C2*) in Proposition 1 is equivalent to:

$$\sum_{m=1}^{N}\beta_{im}^{*\{a\}}\alpha_{mj} > 0, \text{ for all } i,j = 1,\ldots,N \text{ (C2*)},$$

where $\beta_{im}^{*\{a\}} = (-1)^{i+m}\det[(\mathbf{B}^{*})_m^i]$. Next, let $\boldsymbol{\phi}=[\phi_i]$ be the vector of the N distinct roots. Then it exists a nonsingular matrix $\mathbf{D} = [d_{ij}]_{i,j=1,\ldots,N}$ (that is an $N \times N$ matrix with the N eigenvectors of $\mathbf{B}$) such that

$$\mathbf{B}^k = \mathbf{\Lambda} diag[\boldsymbol{\phi}^k]\mathbf{\Lambda}^{-1}.$$

Denote the ijth element of $\mathbf{\Lambda}^{-1}$ by λ_{ij}^{*}, that is $\mathbf{\Lambda}^{-1} = [\lambda_{ij}^{*}]_{i,j=1,\ldots,N}$. Then, for each k, condition (C3*) in Proposition 1, that the ijth element of $\mathbf{\Psi}^{(k)} = \mathbf{B}^{k-1}\mathbf{A}$ must be positive for all i,j, amounts to:

$$\psi_{ij}^{(k)} = \sum_{m=1}^{N}\sum_{l=1}^{N}\lambda_{il}\lambda_{lm}^{*}\phi_l^{k-1}\alpha_{mj} > 0, \text{ for all } i,j = 1,\ldots,N, \text{ (C3*)}.$$

The above inequality when $k = 1$ reduces to: $\alpha_{ij} > 0$, $i,j = 1,\ldots,N$, since $\sum_{l=1}^{N}\lambda_{il}\lambda_{lm}^{*} = 1$ if $i = m$ and zero otherwise.

3.3 Trivariate System and Numerical Examples

In this Section numerical examples are included to show the effectiveness of the proposed method. These may be helpful to the researcher who wishes to skip theoretical derivations and is mainly interested in the application of these constraints to a given N-dimensional system at hand. Next we will discuss a specific model in order to make our analysis more concise. That is, for illustrative purposes, we will consider the trivariate case.

Lemma 3 *Let Assumptions (A1) and (A2) be satisfied and $\phi_1 \neq \phi_2 \neq \phi_3$. The following conditions are necessary and sufficient for $\sigma_{it}^{\delta_i/2} > 0$, $i = 1,2,3$, for all t in the trivariate SUE-P model (with $\kappa = 3$):*
(a) For the three constants we require

$$\sum_{m=1}^{3} \beta_{im}^{\star\{a\}} \omega_m > 0, \ \textit{for all } i = 1,2,3 \textit{ where}$$
$$\beta_{im}^{\star\{a\}} = \begin{cases} (1-\beta_{ll})(1-\beta_{nn}) - \beta_{ln}\beta_{nl} & \textit{if} \quad i = m;\, l \neq n \neq i, \\ (-1)^{i+m+1}[\beta_{im}(1-\beta_{ll}) + \beta_{il}\beta_{lm}] & \textit{if} \qquad i \neq m \neq l \end{cases}$$
$$\phi_1 \textit{ is real, and } \phi_1 > 0, \qquad \textit{(C1'')}$$

(b)
$$\sum_{m=1}^{3} \beta_{im}^{*\{a\}} \alpha_{mj} > 0, \ \textit{for all } i,j = 1,2,3, \textit{ where}$$
$$\beta_{im}^{*\{a\}} = \begin{cases} (\phi_1-\beta_{ll})(\phi_1-\beta_{nn}) - \beta_{ln}\beta_{nl} & \textit{if} \quad i = m,\ l \neq n \neq i, \\ (-1)^{i+m+1}[\beta_{im}(\phi_1-\beta_{ll}) - \beta_{il}\beta_{lm}] & \textit{if} \qquad i \neq m \neq l, \end{cases} \qquad \textit{(C2'')}$$
$$a_{ij} \geq 0, \ \sum_{m=1}^{3} \beta_{im}\alpha_{mj} \geq 0, \textit{ and } \sum_{m=1}^{3}\sum_{l=1}^{3} \beta_{il}\beta_{lm}\alpha_{mj} > 0, \ \textit{for all } i,j = 1,2,3. \qquad \textit{(C3'')}$$

3.4 Numerical Examples

In what follows we graphically illustrate the necessary and sufficient parameter set for the trivariate SUE-AP system. This will provide a better understanding of the results presented in the previous subsection. We discuss four examples. We allow two off-diagonal elements of **B** to vary from -0.5 to 0.5. In the first example, we examine the situation where b_{13} and b_{31} vary. The purpose is to see if bidirectional negative spillovers are permitted. In the second example, we allow b_{21} and b_{31} (i.e., two coefficients in the first column of **B**) to vary. The purpose is to see if negative spillovers from one variable to the other two variables can be allowed. In the third example, we vary b_{21} and b_{23} (i.e., two coefficients in the second row of **B**). The purpose is to see if negative spillovers from the two variables to the third one can be allowed. In the fourth example, we examine if more than two off-diagonal elements of the **B** matrix can be negative. To do so, we restrict b_{21} to be negative and vary b_{13} and b_{31}.

The parameters chosen are mainly from empirical results in Table ? presented in the next Section (particularly the results from the FTSE index in dataset 2). The Data generation process are given by:

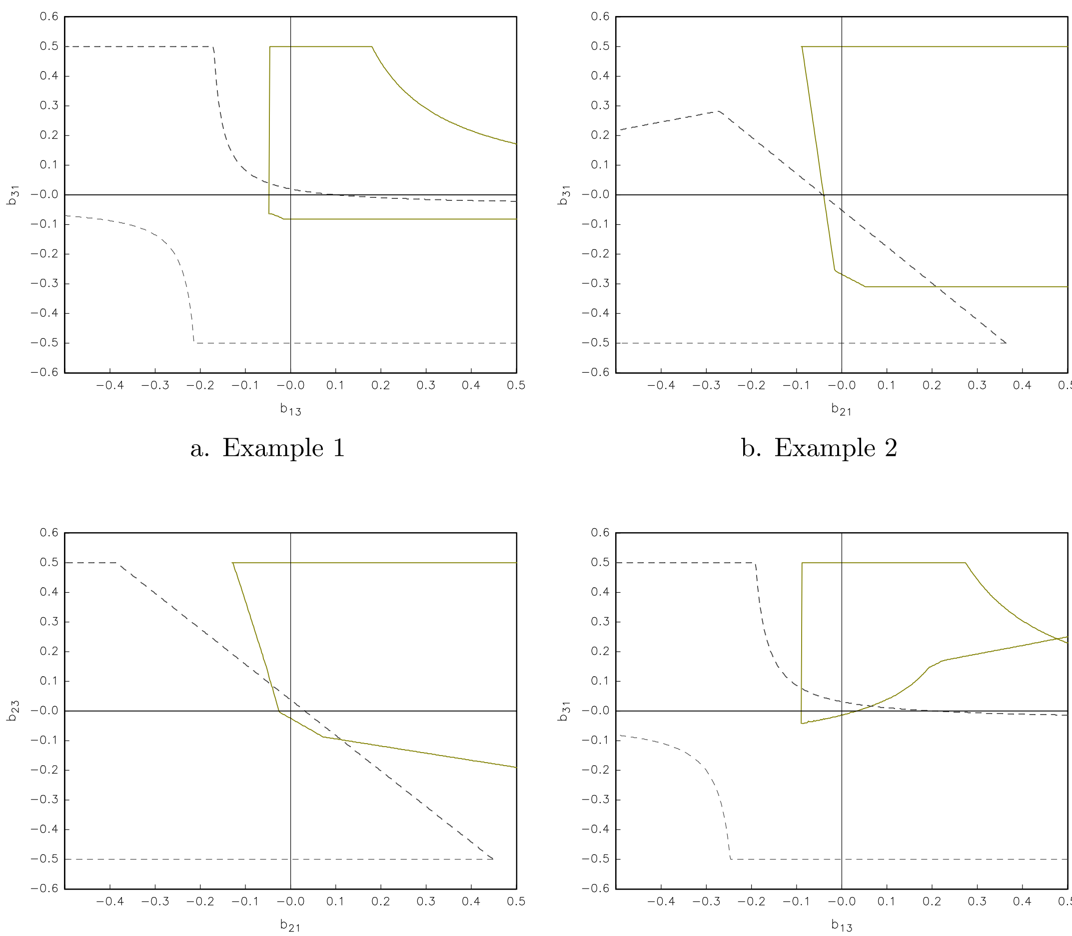
Table1A: Data Generation Process(DGP) for examples 1 and 2.

	DGP Ex.1	DGP Ex.2
$\boldsymbol{\omega}'$	$\begin{pmatrix} 0.149 & 0.074 & 0.124 \end{pmatrix}$	$\begin{pmatrix} 0.153 & 0.106 & 0.103 \end{pmatrix}$
A	$\begin{pmatrix} 0.064 & 0.021 & 0.158 \\ 0.008 & 0.005 & 0.108 \\ 0.028 & 0.043 & 0.198 \end{pmatrix}$	$\begin{pmatrix} 0.075 & 0.017 & 0.123 \\ 0.021 & 0.002 & 0.109 \\ 0.030 & 0.037 & 0.104 \end{pmatrix}$
B	$\begin{pmatrix} 0.790 & 0.032 & b_{13} \\ 0.001 & 0.808 & 0.006 \\ b_{31} & 0.137 & 0.616 \end{pmatrix}$	$\begin{pmatrix} 0.780 & 0.003 & 0.017 \\ b_{21} & 0.901 & 0.022 \\ b_{31} & 0.082 & 0.729 \end{pmatrix}$

Table1B: Data Generation Process(DGP) for examples 3 and 4.

	DGP Ex.3	DGP Ex.4
$\boldsymbol{\omega}'$	$\begin{pmatrix} 0.162 & 0.114 & 0.117 \end{pmatrix}$	$\begin{pmatrix} 0.214 & 0.184 & 0.164 \end{pmatrix}$
A	$\begin{pmatrix} 0.075 & 0.011 & 0.140 \\ 0.013 & 0.003 & 0.139 \\ 0.023 & 0.044 & 0.201 \end{pmatrix}$	$\begin{pmatrix} 0.078 & 0.012 & 0.171 \\ 0.012 & 0.005 & 0.100 \\ 0.048 & 0.029 & 0.228 \end{pmatrix}$
B	$\begin{pmatrix} 0.744 & 0.002 & 0.051 \\ b_{21} & 0.901 & b_{23} \\ 0.009 & 0.056 & 0.559 \end{pmatrix}$	$\begin{pmatrix} 0.743 & 0.031 & b_{13} \\ \mathbf{-0.028} & 0.851 & 0.053 \\ b_{31} & 0.111 & 0.548 \end{pmatrix}$

In the following figures, the bold solid lines show which combinations of the two freely parameters satisfy the necessary and sufficient conditions of Proposition 1 and of the second moments. We begin by discussing the implications of Example 1, which is presented in Figure 1a. First, all combinations of b_{13} and b_{31} that are bounded by the three bold solid lines satisfy the conditions of Proposition 1. Interestingly, both off-diagonal elements can be negative simultaneously. Second, the combinations of b_{13} and b_{31}, which are bounded by the dotted lines, satisfy the conditions for the existence of the second moments. Example 2 is visualized in figure 1b. The conditions of Proposition 1 allow for negative volatility spillovers from $\sigma_{1t}^{\delta_1/2}$ to $\sigma_{2t}^{\delta_1/2}$ and $\sigma_{3t}^{\delta_1/2}$. The negative parameter set that satisfies all conditions simultaneously is given by the area that is above the solid line and below the dotted line in the third quadrant. Figure 2c shows that, for the parameters in Example 3, the conditions of Proposition 1 allow for negative volatility spillovers from from $\sigma_{1t}^{\delta_1/2}$ and $\sigma_{3t}^{\delta_1/2}$ to $\sigma_{2t}^{\delta_1/2}$ only if b_{21} and b_{23} take rather small values. From example 4, it is interesting to observe that three off-diagonal elements in the **B** matrix can be negative, since the parameter set given by the area that is above the solid line in the third quadrant satisfies the conditions in Proposition 1.



a. Example 1

b. Example 2

c. Example 3

d. Example 4

FIGURE 1. Necessary and sufficient parameter sets for trivariate SUE-AP system from Examples 1 to 4. Bold solid lines represent the restrictions implied by Proposition 1. Dotted lines represent the restrictions implied by the existence of the unconditional second moment.

4 Extensions

TEXT

4.1 Assymetric Systems

TEXT

Asymmetric Power Model

In the asymmetric version of the N-dimensional system it will suffice to show that the matrix inequality constraints given in Proposition 1 are satisfied not only for $\mathbf{A}$ but for $\mathbf{\Gamma}$ as well. In other words, it is suffice to show that the constraints are satisfied in the two extreme cases: i) $\mathbf{\Gamma} = 0$ so there are no asymmetries and thus $\mathbf{A}_\tau = \mathbf{A}$, and ii) $\mathbf{A} = 0$ and $\mathbf{A}_\tau = \mathbf{\Gamma}$ for all τ in eq. (3), that is, only the negative errors affect the power transformed conditional variances and all the errors are negative at all times τ (see also the Additional Appendix). Therefore, the following Theorem holds.

Theorem 4 *Consider the N-dimensional vector SUE-AP model in eq. (3) and let Assumptions (A1) and (A2) be satisfied and all the roots ϕ_n, $n = 1, \ldots N$, be distinct. Then, necessary and sufficient conditions for $\sigma_{it}^{\delta_i/2} > 0$, $i = 1, \ldots, N$, for all t are given by:*
(a) $\boldsymbol{\mu} = adj[\mathbf{I} - \mathbf{B}]\boldsymbol{\omega} > \mathbf{0}$

(b)
ϕ_1 *is real, and* $\phi_1 > 0$, *(C1)*
$adj[\mathbf{I}\phi_1 - \mathbf{B}]\mathbf{A} > \mathbf{0}$ *and* $adj[\mathbf{I}\phi_1 - \mathbf{B}]\mathbf{\Gamma} > \mathbf{0}$, *(C2**)*
$\mathbf{\Psi}^{(k_a)} = \mathbf{B}^{(k_a-1)}\mathbf{A} \overset{(k_\alpha \leq \kappa_{ij,\alpha})}{\succeq} \mathbf{0}$ *and* $\mathbf{\Psi}^{(k_\gamma)} = \mathbf{B}^{(k_\gamma-1)}\mathbf{\Gamma} \overset{(k_\gamma \leq \kappa_{ij,\gamma})}{\succeq} \mathbf{0}$ *(C3**)*

(the symbol $\overset{(k_l \leq \kappa_{ij,l})}{\succeq}$ means that if $k_l > \kappa_{ij,l}$ then the ijth scalar inequality of the matrix inequality $\mathbf{\Psi}^{(k_l)} \geq \mathbf{0}$ should be disregarded), where $\kappa_l = \max\{\mathbf{K}_l\}$, $\mathbf{K}_l = [\kappa_{ij,l}]$, and $\kappa_{ij,l}$ is the smallest integer greater than or equal to $\max\{0, \varphi_{ij,l}\}$, $l = \alpha, \gamma$ with

$$\mathbf{\Phi}_l = [\varphi_{ij,l}] = \{\log[\mathbf{H}_l^{(1)}] - \log[(N-1)\mathbf{H}_l]\}[\log(|\phi_2|) - \log(|\phi_1|)]^{-1},$$

where

$$\mathbf{H}_l = [\eta_{ij,l}] = \max_{2 \leq n \leq N}\{\mathbf{H}_l^{(n)}\} = [\max_{2 \leq n \leq N} \eta_{ij,l}^{(n)}]_{i,j=1,\ldots,N},$$

and

$$\mathbf{H}_l^{(n)} = [\eta_{ij,l}^{(n)}]_{i,j=1,\ldots,N} = abs\left[\frac{adj[\mathbf{I}\phi_n - \mathbf{B}]\mathbf{A}^{[l]}}{\sum_{j=1}^{N} j\beta_j(\phi_n)^{N-(j-1)}}\right], \; 1 \leq n \leq N, \tag{9}$$

where $\mathbf{A}^{[l]} = \mathbf{A}$ if $l = \alpha$ and $\mathbf{A}^{[l]} = \mathbf{\Gamma}$ if $l = \gamma$ for all τ.

4.2 Models of Higher Order: $(1, q)$

In this Section we extent the order of the N-dimensional vector SUE-AP system in eq. (3) from $(1, 1)$ to $(1, q)$. That is we consider the multivariate process:

$$(\mathbf{I} - \mathbf{B}L)\boldsymbol{\sigma}_t = \boldsymbol{\omega} + \sum_{l=1}^{q} L^l \mathbf{A}_{lt}\boldsymbol{\varepsilon}_t, \tag{10}$$

where $\mathbf{A}_{lt} = \mathbf{A}_l + \mathbf{\Gamma}_{lt}$ with $\mathbf{A}_l = [\alpha_{ij}^{(l)}]$ and $\mathbf{\Gamma}_{lt} = \mathbf{\Gamma}_l diag\{\mathbf{s}_t\}$, $\mathbf{\Gamma}_l = [\gamma_{ij}^{(l)}]$. For the above process the following Proposition holds.

Proposition 2 *Consider the N-dimensional vector SUE-AP $(1, q)$ model in eq. (10) and let Assumptions (A1)-(A2) be satisfied and all the roots ϕ_n, $n = 1, \ldots N$, be distinct. Then, necessary and sufficient conditions for $\sigma_{it}^{\delta_i/2} > 0$, $i = 1, \ldots, N$ are as in Theorem 4 where we replace i) in condition (C2*) and eq. (9) $\mathbf{A}$ and $\mathbf{\Gamma}$ by $\widetilde{\mathbf{A}} = \sum_{j=1}^{q} \mathbf{A}_j \phi_i^{q-j}$ and $\widetilde{\mathbf{\Gamma}} = \sum_{j=1}^{q} \mathbf{\Gamma}_j \phi_i^{q-j}$, respectively, and ii) in Condition (C3**) $\mathbf{B}^{(k_a-1)}\mathbf{A}$ by $\sum_{s=1}^{\min(q,k_\alpha)} \mathbf{B}^{(k_\alpha - s)} \mathbf{A}_s$ and $\mathbf{B}^{(k_\gamma-1)}\mathbf{\Gamma}$ by $\sum_{s=1}^{\min(q,k_\alpha)} \mathbf{B}^{(k_\alpha - s)} \mathbf{\Gamma}_s$.*

An analogous result (not presented) holds for the N-dimensional SUE-AP model 2 of order $(1, q)$ (see the Appendix).

4.3 Examples

5 New Mixture Formulations

Once we have identified for which equation (power transformed conditional variance) the non-negativity constraints are not met, then they are easily removed by using a new mixture formulation.

5.1 Power-Exponential Model

First, we will introduce a new N-dimensional formulation, which we will term unrestricted extended asymmetric power-exponential (UE-APE). This formulation, with the Glosten et al. (1993) type of asymmetry, is given by

$$\sigma_{it}^{\delta_i/2} = \omega_i + \sum_{j=1}^{N} \exp[(\alpha_{ij} + \gamma_{ij}\xi_{j,t-1}) |e_{j,t-1}|^{\delta_j}] + \sum_{j=1}^{N} \beta_{ij}\sigma_{j,t-1}^{\delta_j/2}, \; i = 1, \ldots, N,$$

or in a matrix form

$$(\mathbf{I} - \mathbf{B}L)\boldsymbol{\sigma}_t^* = \boldsymbol{\omega} + \mathbf{\Lambda}\exp(\grave{\mathbf{E}}_t), \tag{11}$$

where $\mathbf{\Lambda} = [\lambda_{ij}]_{i,j=1,\ldots,N}$, $\grave{\mathbf{E}}_t = [\grave{e}_{ij,t}]_{i,j=1,\ldots,N}$ with $\grave{e}_{ij,t} = (\alpha_{ij} + \gamma_{ij}\xi_{jt}) |e_{jt}|^{\delta_j}$ and $\exp(\grave{\mathbf{E}}_t)$ denotes elementwise exponentiation. Since we take the exponential of each element of the stochastic matrix $\grave{\mathbf{E}}_t$ this model allows the matrices $\mathbf{A}$ and $\mathbf{\Gamma}$ to take any negative values. Therefore, the following Proposition holds.

Proposition 3 *Consider the N-dimensional UE-APE model in eq. (11) and let Assumptions (A1)-(A2) be satisfied and all the roots ϕ_n, $n = 1, \ldots N$, be distinct. Then, necessary and sufficient conditions for $\sigma_{it}^{\delta_i/2} > 0$, $i = 1, \ldots, N$ are as in Proposition 1 where we replace the matrix $\mathbf{A}$ by $\mathbf{\Lambda}$.*

TEXT

5.2 Asymmetric Mixture Model

Before we do that, however, we will introduce some additional notation. Let $\mathbf{I}_d$ be a diagonal matrix with ones in the first d diagonals and zeros elsewhere, and $\mathbf{I}_d^* = \mathbf{I} - \mathbf{I}_d$. Let also $\boldsymbol{\sigma}_{l,t} = [\ln(\sigma_{it})]$.

Define the $N \times 1$ vector $\boldsymbol{\sigma}_{M,t} = [\sigma_{Mi,t}] = \mathbf{I}_d\boldsymbol{\sigma}_t + \mathbf{I}_d^*\boldsymbol{\sigma}_{l,t}$, that is

$$\sigma_{Mi,t} = \begin{cases} \sigma_{it}^{\delta_i/2} & \text{if} \quad i = 1, \dots, d, \\ \ln(\sigma_{it}) & \text{if} \quad i = d+1, \dots, N, \end{cases} \tag{12}$$

with $d = 0, \dots, N$. In other words the first d conditional variances in this N-dimensional mixture formulation are modelled as power specifications (in the spirit of Ding et al., 1993) whereas the other $N - d$ are modelled as exponential specifications (in the spirit of Nelson, 1991).

Similarly, define the $N \times 1$ vector $\boldsymbol{\varepsilon}_{Mt} = [\varepsilon_{M,it}] = \mathbf{I}_d \boldsymbol{\varepsilon}_t + \mathbf{I}_d^* \mathbf{e}_t$:

$$\varepsilon_{Mi,t} = \begin{cases} |\varepsilon_{it}|^{\delta_i} & \text{if} \quad i = 1, \dots, d, \\ |e_{it}|^{\delta_i} & \text{if} \quad i = d+1, \dots, N. \end{cases} \tag{13}$$

Let also the matrix $\mathbf{J}_{1 \cdots d}$ indicate a binary $N \times N$ matrix with ones in its first d rows, and zeros elsewhere, where $d = 0, \dots, N$.

The N-dimensional vector SUE asymmetric mixture (AM) model consists of the following equations:

$$\begin{aligned} \sigma_{Mi,t} \quad = \quad & \omega_i + \sum_{j=1}^{d} (\alpha_{ij} + \gamma_{ij} s_{j,t-1}) |\varepsilon_{j,t-1}|^{\delta_j} + \sum_{j=d+1}^{N} (\alpha_{ij} + \gamma_{ij} s_{j,t-1}) |e_{j,t-1}|^{\delta_j} \\ & + \sum_{j=1}^{d} \beta_{ij} \sigma_{j,t-1}^{\delta_j/2} + \sum_{j=d+1}^{N} \beta_{ij} \ln(\sigma_{j,t-1}), \; i = 1, \dots, N. \end{aligned} \tag{14}$$

For this model we will use the acronym SUE-AM. The system in a matrix form can be written as

$$(\mathbf{I} - \mathbf{B}L) \boldsymbol{\sigma}_{M,t} = \boldsymbol{\omega} + L \mathbf{A}_t \boldsymbol{\varepsilon}_{M,t}, \tag{15}$$

where $\mathbf{B}$, $\boldsymbol{\omega}$ and $\mathbf{A}_t$ are as in eq. (3). When $d = N$, that is $\mathbf{I}_d = \mathbf{I}$ and $\mathbf{I}_d^* = \mathbf{0}$, then the AM model reduces to the AP model in eq. (3), whereas when $d = 0$, that is $\mathbf{I}_d^* = \mathbf{I}$ and $\mathbf{I}_d = \mathbf{0}$ then the AM becomes identical to the multivariate extension of the exponential specification of Nelson (1991).

Theorem 5 *Consider the N-dimensional vector SUE-AM model in eq. (14) and let Assumptions (A1)-(A2) be satisfied and all the roots ϕ_n, $n = 1, \dots N$, be distinct. Then, necessary and sufficient conditions for $\sigma_{it}^{\delta_i/2} > 0$, $i = 1, \dots, d$, for all t in Theorem 4 become*

(a) $\boldsymbol{\mu} = adj[\mathbf{I} - \mathbf{B}] \boldsymbol{\omega} \odot \mathbf{J}_{(\mathbf{d+1}) \cdots \mathbf{N}} \overset{d}{\succ} \mathbf{0}$

(b)

ϕ_1 *is real, and* $\phi_1 > 0$, *(C1)*

$adj[\mathbf{I}\phi_1 - \mathbf{B}] \mathbf{A} \odot \mathbf{J}_{(\mathbf{d+1}) \cdots \mathbf{N}} \overset{d}{\succ} \mathbf{0}$ *and* $adj[\mathbf{I}\phi_1 - \mathbf{B}] \boldsymbol{\Gamma} \odot \mathbf{J}_{(\mathbf{d+1}) \cdots \mathbf{N}} \overset{d}{\succ} \mathbf{0}$, *(C2***)*

$\boldsymbol{\Psi}^{(k_a)} \odot \mathbf{J}_{(\mathbf{d+1}) \cdots \mathbf{N}} = [\mathbf{B}^{\bar{\wedge}(k_a - 1)} \mathbf{A}] \odot \mathbf{J}_{(\mathbf{d+1}) \cdots \mathbf{N}} \overset{(k_\alpha \leq \kappa_{ij,\alpha})}{\succeq} \mathbf{0}$, *and*

$\boldsymbol{\Psi}^{(k_\gamma)} \odot \mathbf{J}_{(\mathbf{d+1}) \cdots \mathbf{N}} = [\mathbf{B}^{\bar{\wedge}(k_\gamma - 1)} \boldsymbol{\Gamma}] \odot \mathbf{J}_{(\mathbf{d+1}) \cdots \mathbf{N}} \overset{(k_\gamma \leq \kappa_{ij,\gamma})}{\succeq} \mathbf{0}$ *(C3***)*

(the symbol $\overset{d}{>}$ means that we check only the inequalities in the first d rows of the matrix and we disregard the other inequalities), where $\kappa_l = \max[\mathbf{K}_l \odot \mathbf{J}_{(\mathbf{d+1}) \cdots \mathbf{N}}]$ and $\mathbf{K}_l$ is as in Theorem 4.

The non-negativity constraints (matrix inequalities) in the above Theorem are similar to (actually exactly the same) those in Theorem 4 with the only difference that now for each matrix inequality we have only $d \times N$ scalar inequalities and not N^2, that is, we only have to check the scalar inequalities in the first d rows which are linked to the d power transformed conditional variances.

Table 1. Various HEAVYMEM/GARCH Models

Models↓	Formulations	Notation	Non-neagtivity Constraints
	Panel A, Semi-Unrestricted Extended Asymmetric Power (SUE-AP)		
SUE-AP	$(\mathbf{I}-\mathbf{B}L)\boldsymbol{\sigma}_t=\boldsymbol{\omega}+L\mathbf{A}_t\boldsymbol{\varepsilon}_t$ (Eq. (3))	$\mathbf{A}_t=\mathbf{A}+\boldsymbol{\Gamma}_t$, $\boldsymbol{\Gamma}_t=\boldsymbol{\Gamma}\mathbf{diag}\{\mathbf{s}_t\}$	Theorem (4)
If $\boldsymbol{\Gamma}=\mathbf{0}$, and $\delta_j=2\ \forall\ j$, then the above model reduces to the		SUE-HEAVY model: $(\mathbf{I}-\mathbf{B}L)\boldsymbol{\sigma}_t=\boldsymbol{\omega}+L\mathbf{A}\boldsymbol{\varepsilon}_t$ (Eq. (??))	
	Panel B. Mixture Formulations		
UE-APE	$(\mathbf{I}-\mathbf{B}L)\boldsymbol{\sigma}_t^*=\boldsymbol{\omega}+\boldsymbol{\Lambda}\exp[\grave{\mathbf{E}}_t]$ (Eq. (11))	$\grave{\mathbf{E}}_t=[\grave{e}_{ij,t}]$ with $\grave{e}_{ij,t}=(\alpha_{ij}+\gamma_{ij}s_{jt})\left\|e_{jt}\right\|^{\delta_j}$	Proposition(3)
SUE-AM	$(\mathbf{I}-\mathbf{B}L)\boldsymbol{\sigma}_{M,t}=\boldsymbol{\omega}+L\mathbf{A}_t\boldsymbol{\epsilon}_{M,t}$ (Eq. (14))	$\sigma_{Mi,t}=\begin{cases}\sigma_{it}^{\delta_i/2} & \text{if} \quad i=1,\ldots,d,\\ \ln(\sigma_{it}) & \text{if} \quad i=d+1,\ldots,N\end{cases}$ $\varepsilon_{Mi,t}=\begin{cases}\left\|\varepsilon_{it}\right\|^{\delta_i} & \text{if} \quad i=1,\ldots,d,\\ \left\|e_{it}\right\|^{\delta_i} & \text{if} \quad i=d+1,\ldots,N\end{cases}$	Theorem (5)

Combining restrictions on the parameter space with the sample data, using the restricted conditional maximum likelihood estimator...

The parameter estimates obtained from refitting the model with the extra restriction that

6 Moment Structure

Further, let $\overline{z}=\mathbb{E}(\mathbf{e}_t)=[\overline{e}_i]=[\mathbb{E}(|e_{it}|)^{\delta_i}]$ and $\overline{\mathbf{Z}}=\mathbb{E}(\mathbf{Z}_t)=diag\{\overline{\mathbf{z}}\}$. We will also use the following notation: $\overline{\mathbf{Z}}_{2\otimes}=\overline{\mathbf{Z}}\otimes\overline{\mathbf{Z}}$ and $\overline{\overline{\mathbf{Z}}}_{2\otimes}=\mathbb{E}(\mathbf{Z}_t\otimes\mathbf{Z}_t)=\mathbb{E}[(\mathbf{Z}_t)_{2\otimes}]$.

ARMA Representation

Assumption 1. a) $\overline{\mathbf{Z}}$, $\overline{\mathbf{Z}}_{2\otimes}<\mathbf{M}<\infty$, that is $\mathbb{E}\left|e_{it}\right|^{\delta_i}$ exists and it is bounded $\forall\ i$, and b) $\overline{\overline{\mathbf{Z}}}_{2\otimes}<\mathbf{M}<\infty$, that is $\mathbb{E}[\left|e_{it}\right|^{\delta_i}\left|e_{jt}\right|^{\delta_j}]$ exists and it is bounded $\forall\ i,\ j$; where $\mathbf{M}=[\mu_{ij}]$ and $\mu_{ij}\in\mathbb{R}^+$.

Corollary 6 *Under Assumption A.1a, the ARMA representation of the N-dimensional SUE-AP-process in eqs. () and (15) is given by*

$$[\mathbf{I}-\mathbf{C}_tL]\boldsymbol{\sigma}_{M,t}=\boldsymbol{\omega}_t+\mathbf{A}_t[\mathbf{I}_dL\mathbf{v}_{\varepsilon t}+\mathbf{I}_d^*L\mathbf{v}_{et}], \tag{16}$$

where $\mathbf{C}_t=\mathbf{B}+\mathbf{A}_t\mathbf{I}_d\overline{\mathbf{Z}}$, $\overline{\mathbf{Z}}$ has been defined above, and $\mathbf{B}$, $\mathbf{A}_t$ are given in eq. (3), and $\boldsymbol{\omega}_t=\omega+\mathbf{A}_t\mathbf{I}_d^*\overline{\mathbf{z}}$; $\mathbf{v}_{\varepsilon t}=(\boldsymbol{\varepsilon}_t-\overline{\mathbf{Z}}\boldsymbol{\sigma}_t)$ and $\mathbf{v}_{et}=(\mathbf{e}_t-\overline{\mathbf{z}})$ are two vectors of uncorrelated error terms. In addition, when $\mathbf{I}_d=\mathbf{I}$, and therefore $\mathbf{I}_d^*=\mathbf{0}$, that is when we have the AP formulation, then

$$[\mathbf{I}-\mathbf{C}_tL]\boldsymbol{\sigma}_t=\boldsymbol{\omega}+\mathbf{A}_tL\mathbf{v}_{\varepsilon t},$$

with $\mathbf{C}_t=\mathbf{B}+\mathbf{A}_t\overline{\mathbf{Z}}$, whereas when $\mathbf{I}_d^*=\mathbf{I}$, and therefore $\mathbf{I}_d=\mathbf{0}$, that is when we have the exponential formulation, then

$$[\mathbf{I}-\mathbf{B}L]\boldsymbol{\sigma}_{L,t}=\boldsymbol{\omega}_t+\mathbf{A}_tL\mathbf{v}_{et}.$$

General Solution

Next, define the partitioned (or block) two-diagonal matrix

$$\mathbf{C}_{t,k}=\begin{pmatrix}\mathbf{C}_t & -\mathbf{I} & & & & \\ & \mathbf{C}_{t-1} & \ddots & & & \\ & & \ddots & \ddots & & \\ & & & \ddots & \ddots & \\ & & & & \mathbf{C}_{t-k+2} & -\mathbf{I}\\ & & & & & \mathbf{C}_{t-k+1}\end{pmatrix},\ k\in\mathbb{Z}^+ \tag{17}$$

(here and in what follows empty spaces in partitioned matrix have to be replaced by the null matrix). The above matrix for $k \geq 1$, is a partitioned matrix of order kN that possesses 2 diagonals (the main diagonal and the superdiagonal) with nonzero 'matrix' entries. Next let $\mathbf{D}_{t,k}$ associated with the bivariate function

$$\mathbf{D}_{t,k} = \det(\mathbf{C}_{t,k}) = \prod_{r=0}^{k-1} \mathbf{C}_{t-r}, \tag{18}$$

coupled with the initial value $\mathbf{D}_{t,0} = \mathbf{I}$.

Proposition 4 *The general solution of eq. (16) with free constant (initial condition value) $\boldsymbol{\sigma}^{\diamond}_{t-k}$, in terms of the $\mathbf{D}$s exclusively, is given by*

$$\boldsymbol{\sigma}_{M,t} \triangleq \boldsymbol{\sigma}_{M,t;k} = \sum_{r=0}^{k-1} \mathbf{D}_{t,r}[\boldsymbol{\omega}_{t-r} + \mathbf{A}_{t-r}(\mathbf{I}_d L\mathbf{v}_{\epsilon,t-r} + \mathbf{I}^*_d L\mathbf{v}_{e,t-r})] + \mathbf{D}_{t,k}\boldsymbol{\sigma}_{M,t-k}. \tag{19}$$

In addition, if $\mathbf{I}_d = \mathbf{I}$, then

$$\boldsymbol{\sigma}_t \triangleq \boldsymbol{\sigma}_{t;k} = \sum_{r=0}^{k-1} \mathbf{D}_{t,r}(\boldsymbol{\omega} + \mathbf{A}_{t-r}L\mathbf{v}_{\epsilon,t-r}) + \mathbf{D}_{t,k}\boldsymbol{\sigma}_{t-k},$$

whereas if $\mathbf{I}_d = \mathbf{0}$, then

$$\boldsymbol{\sigma}_{L,t} \triangleq \boldsymbol{\sigma}_{L,t;k} = \sum_{r=0}^{k-1} \mathbf{B}^r(\boldsymbol{\omega}_{t-r} + \mathbf{A}_{t-r}L\mathbf{v}_{e,t-r}) + \mathbf{B}^k\boldsymbol{\sigma}_{L,t-k}.$$

In the above Theorem $\boldsymbol{\sigma}_{M,t;k}$ is decomposed in two parts: the homogeneous part consists of the free constant $\boldsymbol{\sigma}_{M,t-k}$; the particular part contains the drift ($\boldsymbol{\omega}$s) and the $\mathbf{v}$s from time $t-k$ to time $t-1$. Notice that the 'matrix coefficients' or weights of eq. (19), that is the $\mathbf{D}$s are expressed as two-diagonal partitioned determinants. Moreover, for '$k = 0$' (for $i > j$ we will use the convention $\sum_{r=i}^{j}(\cdot) = 0$), since $\mathbf{D}_{t,0} = \mathbf{I}$ (see eq.(18)) eq. (19) becomes an 'identity': $\boldsymbol{\sigma}_{M,t;k} = \boldsymbol{\sigma}_{M,t}$. Similarly, when $k = 1$ eq. (19), since $\mathbf{D}_{t,1} = \mathbf{C}_t$, reduces to 'eq. (16)'.

Optimal Predictors

Taking the conditional expectation of eq. (19) with respect to the σ field $\mathcal{F}_{t-k-1}$ yields the following Proposition.

Proposition 5 *The k-step-ahead optimal (in L_2 sense) linear predictor of $\boldsymbol{\sigma}_{M,t}$,$\mathbb{E}(\boldsymbol{\sigma}_{M,t}\,|\mathcal{F}_{t-k-1})$, is readily seen to be*

$$\mathbb{E}(\boldsymbol{\sigma}_{M,t}\,|\mathcal{F}_{t-k-1}) = (\mathbf{I} - \mathbf{C})^{-1}(\mathbf{I} - \mathbf{C}^k)\overline{\boldsymbol{\omega}} + \mathbf{C}^k\boldsymbol{\sigma}_{M,t-k}, \tag{20}$$

where $\mathbf{C} = \mathbb{E}[\mathbf{C}_t] = \mathbf{B} + \overline{\mathbf{A}}$, with $\overline{\mathbf{A}} = (\mathbf{A} + \boldsymbol{\Gamma}\frac{1}{2})\mathbf{I}_d\overline{\mathbf{Z}}$ and $\overline{\boldsymbol{\omega}} = \boldsymbol{\omega} + \overline{\mathbf{A}}\mathbf{I}^*_{\mathbf{d}}\overline{\mathbf{z}}$. In addition, the first-order moment vector $\overline{\boldsymbol{\sigma}}_M = \mathbb{E}(\boldsymbol{\sigma}_{M,t})$ exists if

$$\lim_{k\to\infty}(\mathbf{C}^k) = 0 \text{ or } \lambda(\mathbf{C}) < 1, \tag{21}$$

where $\lambda(\mathbf{C})$ refers to the modulus of the largest eigenvalue of $\mathbf{C}$. Under (21) $\overline{\boldsymbol{\sigma}}_M$ is given by

$$\overline{\boldsymbol{\sigma}}_M = \lim_{k\to\infty}\mathbb{E}(\boldsymbol{\sigma}_{M,t}\,|\mathcal{F}_{t-k-1}) = (\mathbf{I} - \mathbf{C})^{-1}\overline{\boldsymbol{\omega}}. \tag{22}$$

Further, $\mathbf{C}^k$ can be expressed as

$$\mathbf{C}^k = \widetilde{\mathbf{C}} diag\{\boldsymbol{\phi}^{\wedge k}\}\widehat{\mathbf{C}},$$

(see, for example, Hamilton,) where $\widetilde{\mathbf{C}} = [\widetilde{c}_{ij}]$ is the matrix with the N eigenvectors of $\mathbf{C}$, $\boldsymbol{\phi} = [\phi_i]$ is the vector of the N eigenvalues ($^\wedge$ denotes elementwise exponentiation) and $\widehat{\mathbf{C}} = [\widehat{c}_{ij}]$ is the inverse of $\widetilde{\mathbf{C}}$. Thus we can write

$$\mathbb{E}(\sigma_{M,it}\,|\mathcal{F}_{t-k-1}) = \overline{\omega}_i^{(k)} + \sum_{m=1}^{N}\sum_{l=1}^{N}\widetilde{c}_{il}\widehat{c}_{lm}\phi_l^k\sigma_{Mm,t-k},$$

where $(\mathbf{I}-\mathbf{C})^{-1}(\mathbf{I}-\mathbf{C}^k)\overline{\boldsymbol{\omega}}=\overline{\boldsymbol{\omega}}_k = [\overline{\omega}_i^{(k)}]$.

In the next Corollary we will examine the two special cases, that of the AP model and the one of the exponential multivariate formulation.

Corollary 7 *If $\mathbf{I}_d = \mathbf{I}$, that is when we have the AP model then*

$$\mathbb{E}(\boldsymbol{\sigma}_t\,|\mathcal{F}_{t-k-1}) = (\mathbf{I}-\mathbf{C})^{-1}(\mathbf{I}-\mathbf{C}^k)\boldsymbol{\omega} + \mathbf{C}^k\boldsymbol{\sigma}_{t-k},$$

and Under (21)

$$\overline{\boldsymbol{\sigma}} = \lim_{k\to\infty}\mathbb{E}(\boldsymbol{\sigma}_t\,|\mathcal{F}_{t-k-1}) = (\mathbf{I}-\mathbf{C})^{-1}\boldsymbol{\omega}.$$

On the other hand if $\mathbf{I}_d = \mathbf{0}$, that is when we have the exponential formulation then

$$\mathbb{E}(\boldsymbol{\sigma}_{L,t}\,|\mathcal{F}_{t-k-1}) = (\mathbf{I}-\mathbf{B})^{-1}(\mathbf{I}-\mathbf{B}^k)\overline{\boldsymbol{\omega}}+\mathbf{B}^k\boldsymbol{\sigma}_{L,t-k},$$

and if $\lim_{k\to\infty}(\mathbf{B}^k) = 0$ or $\lambda(\mathbf{B}) < 1$, then

$$\overline{\boldsymbol{\sigma}}_L = \lim_{k\to\infty}\mathbb{E}(\boldsymbol{\sigma}_{L,t}\,|\mathcal{F}_{t-k-1}) = (\mathbf{I}-\mathbf{B})^{-1}\overline{\boldsymbol{\omega}}.$$

Next, we define $\boldsymbol{\varepsilon}_{L,t} = [\ln(\varepsilon_{it}^2)] = \mathbf{e}_{L,t}+\boldsymbol{\sigma}_{L,t}$ where $\boldsymbol{\sigma}_{L,t}$ has been defined in eq. (12) and $\mathbf{e}_{L,t} = [\ln(e_{it}^2)]$ with $\overline{\mathbf{e}}_L = \mathbb{E}[\ln(e_{it}^2)]$. Further, $\boldsymbol{\varepsilon}_{ML,t} = \mathbf{I}_d\boldsymbol{\varepsilon}_t + \mathbf{I}_d^*\boldsymbol{\varepsilon}_{L,t}$, where $\mathbf{I}_d$ and $\mathbf{I}_d^*$ have been defined in eq. (12). We can express $\boldsymbol{\varepsilon}_{ML,t}$ as

$$\boldsymbol{\varepsilon}_{ML,t} = \mathbf{I}_d(\mathbf{v}_{\varepsilon t} + \overline{\mathbf{Z}}\boldsymbol{\sigma}_t) + \mathbf{I}_d^*(\mathbf{e}_{L,t} + \boldsymbol{\sigma}_{L,t}), \tag{23}$$

where $\mathbf{v}_{\varepsilon t}$ has been defined in eq. (16). The following Corollary gives the

Corollary 8 *The k-step-ahead optimal (in L_2 sense) linear predictor of $\boldsymbol{\varepsilon}_{ML,t}$ is given by*

$$\begin{aligned}\mathbb{E}(\boldsymbol{\varepsilon}_{ML,t}\,|\mathcal{F}_{t-k-1}) &= \mathbf{I}_d\overline{\mathbf{Z}}\mathbb{E}(\boldsymbol{\sigma}_t\,|\mathcal{F}_{t-k-1}) + \mathbf{I}_d^*(\overline{\mathbf{e}}_L + \mathbb{E}(\boldsymbol{\sigma}_{L,t}\,|\mathcal{F}_{t-k-1}) \;\; \textit{or} \\ \mathbb{E}(\boldsymbol{\varepsilon}_{ML,t}\,|\mathcal{F}_{t-k-1}) &= (\mathbf{I}_d\overline{\mathbf{Z}} + \mathbf{I}_d^*)\mathbb{E}(\boldsymbol{\sigma}_{M,t}\,|\mathcal{F}_{t-k-1}) + \mathbf{I}_d^*\overline{\mathbf{e}}_L,\end{aligned}$$

Further, if $\mathbf{I}_d^* = \mathbf{0}$:

$$\mathbb{E}(\boldsymbol{\varepsilon}_t\,|\mathcal{F}_{t-k-1}) = \overline{\mathbf{Z}}\mathbb{E}(\boldsymbol{\sigma}_t\,|\mathcal{F}_{t-k-1}),$$

whereas if $\mathbf{I}_d = \mathbf{0}$ then

$$\mathbb{E}(\boldsymbol{\varepsilon}_{L,t}\,|\mathcal{F}_{t-k-1}) = \mathbb{E}(\boldsymbol{\sigma}_{L,t}\,|\mathcal{F}_{t-k-1}) + \overline{\mathbf{e}}_L.$$

In addition, under condition (21)

$$\overline{\boldsymbol{\varepsilon}}_{ML} = \lim_{k\to\infty}\mathbb{E}(\boldsymbol{\varepsilon}_{ML,t}\,|\mathcal{F}_{t-k-1}) = (\mathbf{I}_d\overline{\mathbf{Z}} + \mathbf{I}_d^*)(\mathbf{I}-\mathbf{C})^{-1}\overline{\boldsymbol{\omega}} + \mathbf{I}_d^*\overline{\mathbf{e}}_L.$$

Finally, if $\mathbf{I}_d^* = \mathbf{0}$ then

$$\overline{\boldsymbol{\varepsilon}} = \overline{\mathbf{Z}}(\mathbf{I}-\mathbf{C})^{-1}\boldsymbol{\omega},$$

whereas when $\mathbf{I}_d = \mathbf{0}$:

$$\overline{\boldsymbol{\varepsilon}}_L = (\mathbf{I}-\mathbf{B})^{-1}\overline{\boldsymbol{\omega}} + \overline{\mathbf{e}}_L.$$

Second Moments

Next we will introduce some further notation. Let $\mathbf{\Gamma}(k)$ denote the multidimensional covariance function of $\{\boldsymbol{\sigma}_{M,t}\}$, that is

$$\begin{aligned}\mathbf{\Gamma}(l) &= [\gamma_{M,ij}(l)] = \mathbb{E}[(\boldsymbol{\sigma}_{M,t-l} - \overline{\boldsymbol{\sigma}}_M)(\boldsymbol{\sigma}_{M,t} - \overline{\boldsymbol{\sigma}}_M)'], \\ &= \mathbf{\Sigma}(l) - \overline{\boldsymbol{\sigma}}_M\ \overline{\boldsymbol{\sigma}}'_M,\ \ l \in \mathbb{Z},\end{aligned} \tag{24}$$

where $\mathbf{\Sigma}(l) = [\sigma_{M,ij}(l)] = [\mathbb{E}(\sigma_{Mi,t-l}\sigma_{Mj,t})] = \mathbb{E}(\boldsymbol{\sigma}_{M,t-l}\boldsymbol{\sigma}_{M,t})$. In addition,

$$\mathbf{s}(l) = vec(\mathbf{\Sigma}(l)),\ \ \boldsymbol{\gamma}(l) = vec(\mathbf{\Gamma}(l)), \tag{25}$$

Exact form solutions for the $\mathbf{\Gamma}_M$s and $\mathbf{\Sigma}_M$s and conditions for their existence will be presented below. Let also

$$\mathbf{D} = diag\{\sqrt{\gamma_{M,11}(0)}, \dots, \sqrt{\gamma_{M,NN}(0)}\}, \tag{26}$$

where $\gamma_{M,ii}(0)$ is the ith diagonal element of $\mathbf{\Gamma}_M(0)$. To fix notation, write the lth-order autocorrelation matrix of $\{\boldsymbol{\sigma}_{M,t}\}$ as

$$\mathbf{R}(l) = \mathbf{D}^{-1}\mathbf{\Gamma}(l)\mathbf{D}^{-1} \tag{27}$$

for $l \geq 1$.

Next, we define $\boldsymbol{\varepsilon}_{L,t} = [\ln(\varepsilon_{it}^2)]$, $\boldsymbol{\varepsilon}_{ML,t} = \mathbf{I}_d\boldsymbol{\varepsilon}_t + \mathbf{I}_d^*\boldsymbol{\varepsilon}_{L,t}$ and

$$\mathbf{\Gamma}_\epsilon(l) = \mathbf{\Sigma}_{\epsilon_{ML}}(l) - \overline{\boldsymbol{\epsilon}}_{ML}\ \overline{\boldsymbol{\epsilon}}'_{ML} \text{ where } \mathbf{\Sigma}_{\epsilon_{ML}}(l) = \mathbb{E}(\boldsymbol{\epsilon}_{ML,t-l}\boldsymbol{\epsilon}'_{ML,t}), \tag{28}$$

and $\mathbf{s}_{\epsilon_{ML}}(l) = vec(\mathbf{\Sigma}_{\varepsilon_{ML}}(l))$. Moreover, the lth-order autocorrelation matrix of $\{\boldsymbol{\epsilon}_{ML,t}\}$ is defined as

$$\mathbf{R}_{\epsilon_{ML}}(l) = \mathbf{D}_{\epsilon_{ML}}^{-1}\mathbf{\Gamma}_{\epsilon_{ML}}(l)\mathbf{D}_{\epsilon_{ML}}^{-1}, \tag{29}$$

with

$$\mathbf{D}_{\epsilon_{ML}} = diag\{\sqrt{\gamma_{11,\varepsilon_{ML}}(0)}, \dots, \sqrt{\gamma_{NN,\varepsilon_{ML}}(0)}\},$$

where $\gamma_{ii,\epsilon_{ML}}(0)$ is the ith diagonal element of $\mathbf{\Gamma}_{\epsilon_{ML}}(0)$. Similarly, let

$$\mathbf{\Gamma}_e(l) = \mathbf{\Sigma}_e(l) - \overline{\mathbf{e}}\ \overline{\mathbf{e}}',$$

where $\mathbf{\Sigma}_e(l) = \mathbb{E}(e_te'_t)$.

Next, we will introduce some further notation. Let $\mathbf{C}_{\otimes I} = \mathbf{C}\otimes\mathbf{I}$, $\mathbf{C}_{I\otimes} = \mathbf{I}\otimes\mathbf{C}$, $\overline{\overline{\mathbf{A}}}_{2\otimes} = \mathbb{E}(\mathbf{A}_t\mathbf{I}_d\otimes\mathbf{A}_t\mathbf{I}_d)$ and $\overline{\overline{\mathbf{A}}}^*_{2\otimes} = \mathbb{E}(\mathbf{A}_t\mathbf{I}_d^*\otimes\mathbf{A}_t\mathbf{I}_d^*)$. Also $\mathbf{C}_\otimes^*$ is the following $2N^2 \times 2N^2$ matrix:

$$\mathbf{C}_\otimes^* = \begin{bmatrix} \overline{\overline{\mathbf{A}}}_{2\otimes}[\overline{\overline{\mathbf{Z}}}_{2\otimes} - \overline{\mathbf{Z}}_{2\otimes}] & \mathbf{C}_{I\otimes} \\ \mathbf{C}_{\otimes I} & \mathbf{I} \end{bmatrix}, \tag{30}$$

(recall that $\overline{\mathbf{Z}}_{2\otimes}$, $\overline{\overline{\mathbf{Z}}}_{2\otimes}$ have been defined in eq. (**??**)).

Theorem 9 *Consider the N-dimensional vector SUE-AM-process in eqs. (**??**) and (**??**). Assume that condition (21) holds. Then the multidimensional unconditional variance matrix $\mathbf{\Gamma}(0)$ of $\{\boldsymbol{\sigma}_{Mt}\}$ exists if*

$$\lambda(\mathbf{C}_\otimes^*) < 1. \tag{31}$$

Under (31),

$$\boldsymbol{\gamma}(0) = [\mathbf{I} - \overline{\overline{\mathbf{A}}}_{2\otimes}(\overline{\overline{\mathbf{Z}}}_{2\otimes} - \overline{\mathbf{Z}}_{2\otimes}) - \mathbf{C}_{2\otimes}]^{-1}[\overline{\overline{\mathbf{A}}}_{2\otimes}(\overline{\overline{\mathbf{Z}}}_{2\otimes} - \overline{\mathbf{Z}}_{2\otimes})vec(\overline{\boldsymbol{\sigma}\boldsymbol{\sigma}'}) + \overline{\overline{\mathbf{A}}}^*_{2\otimes}\gamma_e(0)]. \tag{32}$$

Further, the vec form of the autocovariance function for lag $l \geq 1$, $\mathbf{s}(l)$, is given by

$$\boldsymbol{\gamma}(l) = (\mathbf{C}^l)_{\otimes I}\boldsymbol{\gamma}(0). \tag{33}$$

When $\mathbf{I}_d = \mathbf{I}$, then $\boldsymbol{\gamma}(0)$ becomes

$$\boldsymbol{\gamma}(0) = [\mathbf{I} - \overline{\overline{\mathbf{A}}}_{2\otimes}(\overline{\overline{\mathbf{Z}}}_{2\otimes} - \overline{\mathbf{Z}}_{2\otimes}) - \mathbf{C}_{2\otimes}]^{-1}\overline{\overline{\mathbf{A}}}_{2\otimes}(\overline{\overline{\mathbf{Z}}}_{2\otimes} - \overline{\mathbf{Z}}_{2\otimes})vec(\overline{\boldsymbol{\sigma}\boldsymbol{\sigma}'})$$

(see He and Teräsvirta, 2004, and Karanasos et al. 2016a,b), whereas when $\mathbf{I}_d = \mathbf{0}$, $\boldsymbol{\gamma}(0)$ reduces to

$$\boldsymbol{\gamma}(0) = [\mathbf{I} - \mathbf{B}_{2\otimes}]^{-1}[\overline{\overline{\mathbf{A}}}^*_{2\otimes}\boldsymbol{\gamma}_e(0)],$$

with $\mathbf{B}_{2\otimes} = \mathbf{B} \otimes \mathbf{B}$ and $\overline{\overline{\mathbf{A}}}^*_{2\otimes} = \mathbb{E}(\mathbf{A}_t \otimes \mathbf{A}_t)$

Theorem 10 *The equivalent moments for $\boldsymbol{\gamma}_\varepsilon(l)$, $l = 0, 1, 2, \ldots$ are*

$$\begin{aligned} \boldsymbol{\gamma}_\varepsilon(l) &= \overline{\mathbf{Z}}_{2\otimes}\boldsymbol{\gamma}(l) + (\overline{\mathbf{Z}}\mathbf{C}^l\overline{\mathbf{A}})_{I\otimes}[\overline{\overline{\mathbf{Z}}}_{2\otimes} - \overline{\mathbf{Z}}_{2\otimes}]\mathbf{s}(0), \; l \geq 1, & (34) \\ \boldsymbol{\gamma}_\varepsilon(0) &= \overline{\overline{\mathbf{Z}}}_{2\otimes}\boldsymbol{\gamma}(0) + [\overline{\overline{\mathbf{Z}}}_{2\otimes} - \overline{\mathbf{Z}}_{2\otimes}]vec(\overline{\boldsymbol{\sigma}\boldsymbol{\sigma}'}) & (35) \end{aligned}$$

*(recall that $\overline{\mathbf{Z}}_{2\otimes}$, $\overline{\overline{\mathbf{Z}}}_{2\otimes}$ have been defined in eq. (**??**)).*

7 Empirical Results

Assuming that the conditional distribution is the log-normal we estimate trivariate and four variate SUE MEM systems (see also Cipollini, et al. 2013):

$$\boldsymbol{\varepsilon}^*_t = \mathbf{E}^*_t\boldsymbol{\sigma}^*_t,$$

where $\boldsymbol{\varepsilon}^*_t$, $\mathbf{E}^*_t$ and $\boldsymbol{\sigma}^*_t$ have been defined in eq. (1). We now assume that the stochastic vector $\mathbf{e}^*_t = [e_{it}]$ with $\mathbf{E}^*_t = diag\{\mathbf{e}^*_t\}$ is independent and identically distributed (i.i.d) with unit vector $\mathbf{1}$ as an expectation and positive definite correlation matrix $\mathbf{R} = [\rho_{ij}]_{i,j=1,\ldots,N}$ with $\rho_{ii} = 1$, and covariance matrix $\mathbf{Q} = [q_{ij}] = diag\{\mathbf{q}^*\}\mathbf{R}diag\{\mathbf{q}^*\}$ with $\mathbf{q}^* = [\sqrt{q_{ii}}]$. The elements of $\boldsymbol{\varepsilon}^*_t$ could be different measures of volatility (i.e. realized volatility, absolute return, high-low range volatility) for an individual asset, or the volatility proxy (i.e. high-low range volatility) for several financial markets, or the intraday trading duration, volume and volatility.

Following Xu (2013) we use the multivariate log-normal for the innovation vector $\mathbf{e}^*_t$, which is a random vector defined in $[\mathbf{0}, +\infty)$, that is $\mathbf{e}^*_t \sim \ln N(\mathbf{1}, \mathbf{Q})$. The log likelihood function is given by

$$l(\theta) = \sum_{t=1}^{T} \ln f(x_t|\theta)$$

where

$$\ln f(\boldsymbol{\varepsilon}^*_t|\boldsymbol{\theta}) = -\frac{N}{2}\ln(2\pi) - \frac{1}{2}\ln|\mathbf{Q}| - \sum_{j=1}^{N}\ln\varepsilon_{it} - \frac{1}{2}(\ln\boldsymbol{\varepsilon}^*_t - \ln\mathbf{e}^*_t - \mathbf{1})'\mathbf{Q}^{-1}(\ln\boldsymbol{\varepsilon}^*_t - \ln\mathbf{e}^*_t - \mathbf{1}). \tag{36}$$

The model can be estimated consistently by quasi maximum likelihood (QMLE). An alternative estimation method was proposed by Cipollini et al. (2013). They bypass the specification of the error's conditional distribution and made use of only the first two conditional moments of the errors by using an efficient generalized methods of moments (GMM) estimation method. By a simulation study Xu(2013) has shown that both the QML and GMM estimation techniques are consistent and that the efficient loss of the QML compare with GMM due to misspecification of the error distribution is trivial.

In what follows we estimate three (two trivariate and one four-variate) MEM models, based on data and model availability. The first example is a trivariate system of trading duration, stock volume and volatility. An equation-by-equation specification of this model is estimated by Manganelli (2005) and we use the same dataset and estimate a trivariate SUE-MEM$(1,1)$ model[5].The second example is a trivariate SUE-MEM$(1,1)$ system of three volatility indicators (realized volatility, absolute returns and high-low range volatility), which is proposed by Engle and Gallo(2006) and also estimated by Cipollini et al. (2013). [6] The third example is by Cipollini et al. (2010), who estimated a four-variate MEM model by using daily high-low range data in four EU stock markets (UK, France, German, Swiss). We estimate the model by using the latest data (from 01/01/2003 to 31/12/2014). All the models are estimated by using the QMLE stragegy initially proposed by Xu (2013). The QMLE estimation results are reported in the Tables below[7].

Duration, Volume and Volatility

Table 1: Trivariate SUE-MEM(1,1) of trading duration, stock volume and volatility.

	AVT			COX			CP			DLP			GAP
A	**0.141**	-	-	**0.132**	-	-	**0.044**	-	-	**0.090**	-	-	**0.101**
	(14.6)			(21.49)			(10.79)			(16.06)			(13.60)
	0.004	0.045	-	-	**0.004**	**0.006**	-	**0.029**	**0.003**	-	**0.050**	-	-
	(0.55)	(1.29)			(9.13)	(3.72)		(2.04)	(2.03)		(4.12)		
	-	0.063	**0.288**	-	**0.041**	**0.314**	**0.151**	**0.424**	**0.768**	-	**0.057**	**0.267**	-
		(1.03)	(16.4)		(7.29)	(50.33)	(29.60)	(21.71)	(107.8)		(5.90)	(34.75)	
B	**0.873**	-	-	**0.873**	0.001	-	**0.950**	0.001	-	**0.912**	**0.004**	0.000	**0.921**
	(90.2)			(148.3)	(0.18)		(176.3)	(0.69)		(169.1)	(3.14)	(0.72)	(149.6)
	-	**0.775**	-	**0.000**	**0.987**	**-0.004**	-	**0.873**	-	-	**0.858**	-	-
		(3.12)		(3.04)	(587.8)	(5.64)		(11.97)			(20.8)		
	0.017	-0.223	**0.826**	0.002	0.003	**0.686**	**0.033**	**-0.040**	**0.232**	**0.024**	**-0.093**	**0.826**	**0.529**
	(2.21)	(0.75)	(81.9)	(0.45)	(0.04)	(110.1)	(2.83)	(3.59)	(32.64)	(7.82)	(3.66)	(162.4)	(17.48)

Notes: We use the same data set with Manganelli (2005). Bollerslev-Wooldridge robust t-statistics in parenthesis. Variables significant at the 5 percent confidence level formatted inbold.

[5] See subsection 4.1 in Manganelli (2005) for a concise description of how the data are prepared for use in his and our paper.

[6] For the description of the data set see Cipollini et al. (2013).

[7] We restrict the A matrix to be nonnegative in estimation.