View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Research Papers in Economics

Massimiliano Caporin
Paolo Paruolo

Spatial effects in Multivariate ARCH

2005/1

UNIVERSITÀ DELL'INSUBRIA
FACOLTÀ DI ECONOMIA

http://eco.uninsubria.it

I punti di vista espressi nei quaderni della Facoltà di Economia riflettono unicamente le opinioni degli autori, e non rispecchiano necessariamente quelli della Facoltà di Economia dell'Università dell'Insubria.

The views expressed in the working papers reflect the opinions of the authors only, and not necessarily the ones of the Economics Faculty of the University of Insubria.

© Copyright Massimiliano Caporin and Paolo Paruolo
Printed in Italy in May 2005
Università degli Studi dell'Insubria
Via Ravasi 2, 21100 Varese, Italy

All rights reserved. No part of this paper may be reproduced in any form without permission from the Authors.

Spatial effects in multivariate ARCH[*]

Massimiliano Caporin
Università di Padova

Paolo Paruolo
Università dell'Insubria

May 31, 2005

Abstract

This paper proposes a new approach for the specification of multivariate GARCH models for data sets with a potentially large cross-section dimension. The approach exploits the spatial dependence structure associated with asset characteristics, like industrial sectors and capitalization size. We use the acronym SEARCH for this model, short for Spatial Effects in ARCH.

This parametrization extends current feasible specifications for large scale GARCH models, while keeping the numbers of parameters linear with respect to the number of assets. An application to daily returns on 20 stocks from the NYSE for the period January 1994 to June 2001 shows the benefits of the present specification.

Keywords: Spatial models, GARCH, Common features, Volatility, Large scale models.
JEL code: C32, C51, C52.

[*]Both authors gratefully acknowledge partial financial support from Italian COFIN 2002 and COFIN 2004 grants. The second author also acknowledges partial financial support from FAR 2004 and 2005 grants. This paper was presented at the JAE conference 'Changing structures in International and Financial Markets and the effects on financial decision making', Venice (Italy) June 2-3, 2005. A previous version was also presented at the first 'Italian conference on econometrics and applied economics', Venice (Italy) January 24-25, 2005.

Contents

1 Introduction

One of the most serious challenges faced by multivariate GARCH models is parameter parsimony for data sets with a large cross-section dimension; see e.g. the review paper by Bauwens et al. (2003). In the most general GARCH specifications, the number of parameters – i.e. the model dimension – is in fact quartic in the number of assets, see e.g. Engle and Kroner (1995). This features has limited the applicability of unrestricted GARCH models to systems of limited dimensions, i.e. with up to a dozen assets.

Bollerslev (1990) proposed the constant conditional correlations model, CCC, hence initiating a class of constrained models, whose dimension is quadratic in the number of assets when employing the general specification of McAleer and Ling (2002). This class of models called CC in the following, increased the maximum feasible system dimension in applications.[1] Being quadratic in the cross–section dimensions, however, it may be estimated for system with limited dimensions.

More recently Engle (2002) has proposed a dynamic version of the CC class, the DCC. The DCC has the same number of parameters of the CCC plus a handful (2 in the basic version) of extra parameters that govern the dynamics of correlations. The DCC model dimension is thus of the same order of magnitude as the CC class, i.e. quadratic in the number of asset.

In a parallel strand of literature, Alexander (2001) and Van Der Weide (2002), have proposed orthogonal GARCH models, OG in the following, where univariate GARCH models are applied to an invertible linear transformation of deviations from the conditional mean. When the orthogonal linear transformation is estimated, also the OG model dimension is quadratic in the number of assets, despite a diagonal specification for the GARCH dynamics.

Both the CC and the OG classes contain restricted versions that are equivalent to univariate models for each single asset; this reduces the model dimensions to be linear in the cross-section dimension. However, under this restriction no attempt is made to model the interaction between units (asset returns), which is of primary importance in volatility models. The need for multivariate models is well documented also on empirical grounds, see e.g. Bauwens et al. (2003).

Other specifications have been proposed in the literature; Factor GARCH have been suggested by Engle et al. (1990), Lin (1992), Bollerslev and Engle (1993) and Vronton et al. (2003). These models are more parsimonious than the ones which do not impose a factor structure; some of the factor specifications are indeed linear in the cross-section dimension. However the factor loadings and the latent factors may be not easy to interpret without further characterization.

In this paper we propose a different approach, based on notions of proximity derived from a sectoral classification of assets. The present class of models is a spatially-restricted multivariate ARCH specification that borrows both from the CC, the OG and the factor GARCH model classes. Asset returns are defined as first neighbors when they share a number of characteristics, e.g. when the correspond-

[1] Chan, Hoti and McAleer (2003) generalized these models allowing for dynamic correlations. This generalization increases the number of parameters, whose number however remains quadratic in the number of assets.

ing firms belong to the same industrial sector and/or capitalization class. Other factor, both qualitative and quantitative, can be deduced from asset fundamentals or derived from the analysis of balance-sheet data. This defines a spatial structure that can be exploited to define (parsimonious) GARCH specifications, which we call SEARCH, short for Spatial Effects in ARCH.

Spatial econometrics has a long history, see e.g. Anselin (1988). Recent references include Giacomini and Granger (2004), who study aggregation of time series, Pesaran et al. (2004) who use spatial concepts to specify a world macro-model, Baltagi et al. (2004), which contains applications of spatial statistics to panel data models. However, applications of ideas from spatial statistics to multivariate GARCH modelling are unknown to the authors; the present paper hence adds a novel approach to design GARCH specifications.

The asymptotic properties of this model class can be deduced from the results in Ling and McAleer (2003) and Comte and Lieberman (2003), as shown in Section 4 below.

We present an empirical application to a set of 20 stocks of the New York Stock Exchange. We use daily data from January 1994 to June 2001. This sample choice excludes the technology market bubble burst. The results show the relevance of capitalization size and industrial sector in defining a proximity structure. The results for SEARCH favorably compare with corresponding results for the CC model.

The rest of the paper is organized as follows. Section 2 presents the SEARCH class, while Section 3 discusses its relations to the CC, DCC, OG classes. Section 4 analyzes the likelihood function and discusses asymptotics. Section 5 report the application while Section 6 concludes. Proofs are placed in 2 final appendices.

In the paper we use the following operators and definitions: $a := b$ or $b =: a$ indicates that a is defined by b; vec denotes the column stacking operator; vech indicates the column stacking of the lower triangular portion of a matrix, including the main diagonal; $\mathrm{vecd}(A)$ indicates a column vector containing the diagonal elements of a matrix A; $\mathrm{diag}(A_1, ..., A_n)$ indicates a matrix with A_1, ..., A_n as blocks on the main diagonal; when the diagonal blocks A_i are scalars, we collect $a_1, ..., a_n$ into a vector $a := (a_1 : ... : a_n)'$ and write $\mathrm{diag}(a) := \mathrm{diag}(a_1, ..., a_n)$. $\mathrm{dg}(A)$ is a matrix with off–diagonal elements equal to 0 and diagonal elements equal to the ones on the diagonal of A; hence $\mathrm{diag}\left(\mathrm{vecd}\left(A\right)\right) = \mathrm{dg}\left(A\right)$; ι_n indicates a $n \times 1$ vector of ones and I is the identity matrix; $1()$ is the indicator function and δ_{ij} is Kronecker's delta, which equals 1 for $i = j$ and 0 otherwise; $\odot$ indicates Hadamard's element-wise product, while $\otimes$ indicates Kronecker's product, $A \otimes B = (a_{ij}B)$; calligraphic letters $\mathcal{D}$, $\mathcal{F}$, $\mathcal{H}$, $\mathcal{K}$, $\mathcal{M}$, $\mathcal{N}$ are used to indicate various selection and duplication matrices, while calligraphic $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$ are used to indicate sets; in particular, $\mathcal{K}_{m,n}$ is a commutation matrix, i.e. a matrix with elements 0 and 1 that satisfies $\mathcal{K}_{mn}\mathrm{vec}(A) = \mathrm{vec}(A')$ for any $m \times n$ matrix A.

2 Search models for conditional variances

Consider an n_x dimensional vector time series $\{x_t\}_{t\in\mathbb{N}}$ and the associated filtration $\mathcal{I}_t := \sigma\left(x_{t-q}, q \geq 0\right)$. Let also $x_t := (y_t' : z_t')'$ be partitioned into a n subvector of variables of interest y_t and a n_z subvector of other information variables. We assume that x_t has finite second moments conditional on $\mathcal{I}_{t-1}$. Indicate by $\mathbb{E}_{t-1}(\cdot) := \mathbb{E}(\cdot|\mathcal{I}_{t-1})$ the conditional expectation operator, and let $\mu_t := \mathbb{E}(y_t|\mathcal{I}_{t-1})$ be the conditional mean of y_t.

Consider a parametric model for the conditional mean $\mu_t := \mathbb{E}_{t-1}(y_t)$, taken for simplicity to be linear. Specifically let $w_t := (y'_{t-1} : ... : y'_{t-p_\mu}, z'_t)'$, a $n_w \times 1$ vector and

$$y_t =: \mathbb{E}_{t-1}(y_t) + \varepsilon_t = \mu w_t + \varepsilon_t, \tag{1}$$

where $\varepsilon_t := y_t - \mu_t$ is the $n \times 1$ vector of deviations from the conditional mean. μ is a matrix of parameters, where $\theta_\mu := \text{vec}(\mu')$ is the $v_\mu \times 1$ vector of parameters.

We indicate by Σ_t the conditional variance covariance matrix of ε_t, $\Sigma_t := \mathbb{V}_{t-1}(\varepsilon_t) = \mathbb{E}_{t-1}(\varepsilon_t\varepsilon'_t)$. Assume that the cross section dependence of ε_t due to asset proximity can be summarized in a linear relation of the form

$$\varepsilon_t = S\varepsilon_t + \eta_t, \tag{2}$$

where η_t is a $n \times 1$ vector of random variables with $\mathbb{E}_{t-1}(\eta_t) = 0$ and $\mathbb{V}_{t-1}(\eta_t) = \mathbb{E}_{t-1}(\eta_t\eta'_t) =: \Gamma_t$. Eq. (2) is similar to standard specifications in the structural VAR literature, see e.g. Amisano and Giannini (1997). The S matrix, which describes the contemporaneous inter-dependence of the elements in ε_t, is assumed to be of spatial structure (to be defined in Subsection 2.1 below). Eq. (2) defines a spatial autoregression (SAR) process with S of spatial structure, see Cressie (1993) eq. (6.3.8) and reference therein.

A direct consequence of the assumption (2) is that

$$\Sigma_t := \mathbb{V}_{t-1}(\varepsilon_t) = (I - S)^{-1}\Gamma_t(I - S')^{-1}.$$

provided $I - S$ is invertible, which we assume in the following. Similarly to the CC class, we assume that the errors η_t in (2) have constant correlations, i.e.

$$\Gamma_t = D_t R D_t \tag{3}$$

where $D_t = \text{diag}(\sigma_{1,t}, \sigma_{2,t}, ...\sigma_{n,t})$ is a diagonal matrix with (possibly) time varying standard deviations and R is a time invariant correlation matrix of the standardized innovations $\psi_t := D_t^{-1}(I - S)\varepsilon_t = D_t^{-1}\eta_t$, $\mathbb{V}_{t-1}(\psi_t) = \mathbb{E}_{t-1}(\psi_t\psi'_t) = R$.

We next consider a diagonal GARCH specification for D_t^2. Let

$$h_t := \text{vecd}(D_t^2) = \begin{pmatrix} \sigma_{1,t}^2 \\ \vdots \\ \sigma_{n,t}^2 \end{pmatrix}, \qquad e_t := \text{vecd}(\eta_t\eta'_t) = \begin{pmatrix} \eta_{1,t}^2 \\ \vdots \\ \eta_{n,t}^2 \end{pmatrix} \tag{4}$$

where by the definition above $h_t = \mathbb{E}_{t-1}(e_t)$, so that $e_t - h_t$ is an innovation with respect to $\mathcal{I}_{t-1}$.

We assume the following GARCH(p_E, p_A) dynamics for h_t driven by e_t:

$$h_t = c + E(L)h_{t-1} + A(L)e_{t-1}, \tag{5}$$

$E(L) := \sum_{l=1}^{p_E} E_l L^{l-1}$ and $A(L) := \sum_{l=1}^{p_A} A_l L^{l-1}$, E_l, A_l are $n \times n$ matrices, c is a n-dimensional vector. The fact that $e_t - h_t$ is an innovation with respect to $\mathcal{I}_{t-1}$

allows to calculate multi-step ahead predictions on the conditional variances using recursions.

The present conditional heteroscedasticity model (2)-(5) is closely related to the CC and OG specifications. In fact choosing $S = 0$ and E_l, A_l diagonal delivers the CC model. Setting instead $I - S$ to be an orthogonal matrix and $R = I_n$ gives the OG specification. Hence (2)-(5) nests both the CC and OG specification. However, without further restrictions, model (2)-(5) has dimension that is quadratic in the number of assets. Further restrictions are hence needed in order to render estimation of the model feasible on large cross sections.

The SEARCH class of models is characterized as the submodel of (2)-(5) with the property that (some of) the matrices S, E_l, A_l, R (SEAR) which define the conditional heteroscedasticity (CH) features of the model, have *spatial structure* in the sense defined in the next subsection.

In the following we indicate with θ a $v \times 1$ vector of parameters. Parameters within θ are partitioned into $\theta = (\theta'_\mu, \theta'_S, \theta'_{EA}, \theta'_R)'$ where θ_S, θ_{EA}, θ_R are the subvectors of parameters in the S, EA and R specification. The dimension of the subvectors θ_S, θ_{EA}, θ_R is indicated as v_S, v_{EA}, v_R respectively, with $v = v_\mu + v_S + v_{EA} + v_R$. The θ_{EA} parameters are further partitioned into c, θ_E, θ_A in an obvious notation.

2.1 Spatial structure

In this subsection we define matrices of spatial structure. This applies to some or all the S, E_l, A_l, R matrices, which are indicated in this subsection by the generic symbol C. We say that C has spatial structure if it can be written as a linear combination of some spatial weight matrices W_q, $q = 1, ..., m$, i.e. if

$$C = c_0 I_n + \sum_{q=1}^{m} c_q W_q^\circ \tag{6}$$

for $m \geq 1$ and c_0, c_1,...,c_m real scalars and W_1°, ..., W_m° pre-defined spatial weight matrices.

A matrix $W^\circ = (w_{ij}^\circ)$ is called a spatial weight matrix, or simply a spatial matrix, if it is has real entries contained in the closed interval $[0, 1]$, and diagonal elements equal to 0, i.e. $0 \leq w_{ij}^\circ \leq 1$, $\text{vecd}(W^\circ) = 0$. If the row-sums equal 1, $W^\circ \iota_n = \iota_n$, then the spatial matrix is called 'normalized'.

A special case of a spatial matrix is the one containing first order neighbors $W^{(1)} = (w_{ij}^{(1)})$, where the entries $w_{ij}^{(1)}$ in row i are non-zero the j-th unit is a first order neighbor of the i-th unit. A similar structure can be used to define spatial lag matrices of generic order q, $W^{(q)}$.

In the following, we use spatial matrices $W^\circ = (w_{ij}^\circ)$ where non-zero entries w_{ij} indicate that the i-th and the j-th units belong to the same classification group. The classification is based on one or more characteristics, called 'factors', borrowing terminology from the statistical analysis of variance. Example of factors are industrial sector and capitalization size. Further factors are defined by constructing interactions among primary factors. Each factor defines a particular spatial matrix. An example of this is given in Subsection 2.4 below.

Other spatial proximity structures may be defined on the basis of one or more variables. Consider a single ordinal variable, and let the units be ordered in increasing or decreasing order. One may define as first order neighbors of unit i in the list, (either or both) the units preceding and following it. Each of these proximity structures is associated with specific spatial matrices W°. An example of this is given in Subsection 2.5 below.

Any combinations of the above constructions may be used to define a spatial structure for each of the matrices S, E_l, A_l, R. In the following we use a categorical classification for industrial sector and capitalization size for the specification of S and an ordering based on capitalization within each class to characterized the spatial structure of E_l, A_l and R.

2.2 Generalized spatial structure

In order to allow for heterogeneity across units in the spatial specification, we define θ_{c_q} be a $n \times 1$ vector $\theta_{c_q} := (\theta_{c_q,1} : ... : \theta_{c_q,n})'$, when n is the dimension of W_q°. We say that C has a generalized spatial structure if

$$c_q = \mathrm{diag}(\theta_{c_q,1} : ... : \theta_{c_q,n}) =: \mathrm{diag}(\theta_{c_q}), \tag{7}$$

$q = 0, ..., m$ in (6). When one or more rows of W_q° have all 0 entries, we define the corresponding entries in θ_{c_q} to be equal to zero, i.e. $J'W_q^\circ = 0$ implies $J'\theta_{c_q} = 0$. This specification is also called the diagonal specification, in constrast with the scalar specification when c_q are scalars.

In (6) the elements $\theta_{c_q,i}$ represent (possibly different) loadings of each asset on the same factor. For simplicity m is here assumed to be the same for all rows i, but this format can easily accommodate different number of terms in the sum by setting to zero appropriate elements $c_{q,j}$. When $c_{q,1} = ... = c_{q,n}$ the diagonal specification reduces to the scalar one.

2.3 Factors and levels

In (6) a subset of terms in $\sum_{q=1}^m c_q W_q^\circ$ may represent a single 'factor', where each spatial matrix W° is associated with a level of the factor. In this subsection we introduce notation for different parametrizations. An application of these terms is given in the next subsection 2.4.

Assume that (6) represents g factors, $\sum_{q=1}^m c_q W_q^\circ = \sum_{h=1}^g \sum_{k=1}^{m_h} c_{h,k} W_{h,k}^\circ$, where $\sum_{k=1}^{m_h} c_{h,k} W_{h,k}^\circ$ represents the h-th factor, $h = 1, ..., g$. A classification of the several possible parameterizations of a single factor is the following.

- **Heterogeneous specification** (HET): all the $c_{h,k}$ coefficient are diagonal; in this specification the loadings on the levels of the factor are asset-specific.

- **Mixed specification** (MIX): at least one $c_{h,k}$ coefficient is scalar and at least one $c_{h,k}$ coefficient is diagonal; in this case some level of the factor has asset-specific loadings, whereas some other level has homogeneous loadings.

Capitalization level $\rightarrow$ Industrial sector $\downarrow$	level 1	level 2	...	level ℓ
sector 1	$\mathcal{C}_{11}$	$\mathcal{C}_{12}$	...	$\mathcal{C}_{1\ell}$
...	...	...	...	...
sector k	$\mathcal{C}_{k1}$	$\mathcal{C}_{k2}$	...	$\mathcal{C}_{k\ell}$

Table 1: Two-way classification based on industrial sector and capitalization.

- **Homogeneous specification** (HOM): all the $c_{h,k}$ coefficient are scalar; in this case the levels of the factor have the same loadings on all assets, and these loadings are not necessarily equal.
- **Restricted homogeneous specification** (RHOM) all the $c_{h,k}$ coefficient are scalar and equal; in this case all levels of the factor have the same loading on all assets.

2.4 S specification

In this subsection we discuss possible spatial specifications for the S matrix. S is restricted to be of spatial nature with all diagonal elements equal to zero, in order to avoid simultaneous effects from ε_{it} to itself. We assume that

$$S = \sum_{i=1}^{m_S} s_i W_i \tag{8}$$

for an appropriate definition of the W_i spatial matrices. The elements s_i must guarantee that $I - S$ is invertible,[2] but are not otherwise constrained. Note that in general there is no need to restrict S to be symmetric or positive definite.

We next consider the specification of the spatial matrices W_i in (8). We here give one of the possible definitions of the spatial matrices W_i, that reflects a two–way classification of the assets into k industrial sectors and ℓ capitalization classes. Sector and size are hence 'factors'; we also consider interaction between the two as another 'factor'.[3]

Consider an integer label q for each asset, where $q \in \mathcal{C}_{\cdot\cdot} := \{1, 2, ..., n\}$. A possible classification is specified in Table 1, where

$$\mathcal{C}_{ij} := \{q \in \mathcal{C}_{\cdot\cdot} : \text{stock } q \text{ belongs to sector } i \text{ and capitalization level } j\} \tag{9}$$

indicates the set of labels of the stocks that belong to industrial sector i and capitalization class j.

Define also the aggregated label sets for each sector $\mathcal{C}_{i\cdot} := \cup_{j=1}^{\ell} \mathcal{C}_{ij}$, for each capitalization class $\mathcal{C}_{\cdot j} := \cup_{i=1}^{k} \mathcal{C}_{ij}$ and for the whole market $\mathcal{C}_{\cdot\cdot} = \cup_{i=1}^{k} \cup_{j=1}^{\ell} \mathcal{C}_{ij} = \cup_{i=1}^{k} \mathcal{C}_{i\cdot} = \cup_{j=1}^{\ell} \mathcal{C}_{\cdot j}$. One possible choice of spatial matrices is the following.

[2] Hence when $m_S = 1$ in the scalar case, this implies that s_1 should differ from the reciprocal of the eigenvalues of W_1.

[3] For simplicity we assume that the capitalization levels are the same for all industrial sectors, even though this simplification may be relaxed. We also assume that the classification is fixed and known throughout the sample and forecast periods.

1. **Market-wide spatial matrix**: $W_{..} := \iota_n \iota_n' - I_n$.

2. **Industrial sector** i **spatial matrix**: $W_{i\cdot} := (w_{i\cdot,ql})$ where $w_{i\cdot,ql} := 1(q \in \mathcal{C}_{i\cdot})\,(1(l \in \mathcal{C}_{i\cdot}) - \delta_{ql})$ for $i = 1, ..., k$.

3. **Capitalization class** j **spatial matrix**: $W_{\cdot j} := (w_{\cdot j,ql})$ where $w_{\cdot j,ql} := 1(q \in \mathcal{C}_{\cdot j})\,(1(l \in \mathcal{C}_{\cdot j}) - \delta_{ql})$ for $j = 1, ..., \ell$.

4. **Interaction spatial matrices**: $W_{ij} := (w_{ij,ql})$ where $w_{ij,ql} = 1(q \in \mathcal{C}_{ij})\,(1(l \in \mathcal{C}_{ij}) - \delta_{ql})$ for $i = 1, ..., k$, $j = 1, ..., \ell$.

Several specifications can be obtained by making use of the above spatial matrices. We here list two polar cases, labelled the heterogenous and homogeneous cases, and a class of intermediate models.

2.4.1 HET

This is a saturated model where all classification factors and interactions thereof may have a different effect. We assume

$$S = s_{..}W_{..} + \sum_{i=1}^{k} s_{i\cdot}W_{i\cdot} + \sum_{j=1}^{\ell} s_{\cdot j}W_{\cdot j} + \sum_{i=1}^{k}\sum_{j=1}^{\ell} s_{ij}W_{ij}. \tag{10}$$

This specification allows for heterogeneity across different classes $\mathcal{C}_{ij}$, $\mathcal{C}_{\cdot j}$, $\mathcal{C}_{i\cdot}$. The parameters to be estimated are $s_{..}$, $s_{i\cdot}$ ($i = 1$, ..., k), $s_{\cdot j}$, ($j = 1$, ..., ℓ), s_{ij} ($i = 1$, ..., k; $j = 1$, ..., ℓ). For the scalar specification the number of parameters is $v_S := k + \ell + k\ell + 1$, while for the diagonal specification $v_S := n(k + \ell + k\ell + 1)$.

2.4.2 RHOM

This is submodel of the previous model where all terms in each of the three sums in (10) have the same coefficient, i.e.

$$s_{i\cdot} = s_{0\cdot} \text{ for all } i, \tag{11}$$
$$s_{\cdot j} = s_{\cdot 0} \text{ for all } j, \tag{12}$$
$$s_{ij} = s_{00} \text{ for all } i \text{ and } j. \tag{13}$$

These restrictions imply homogeneous relations for different levels of the factors and of the interactions. Corresponding to the restrictions in (11), (12), (13), define the following aggregate spatial matrices for all industrial sectors $W_{0\cdot}$, capitalization levels $W_{\cdot 0}$ and interactions W_{00} :

$$W_{0\cdot} := \sum_{i=1}^{k} W_{i\cdot} \qquad W_{\cdot 0} := \sum_{j=1}^{\ell} W_{\cdot j} \qquad W_{00} := \sum_{i=1}^{k}\sum_{j=1}^{\ell} W_{ij}. \tag{14}$$

With these definition, the homogeneous model is defined as

$$S = s_{..}W_{..} + s_{0\cdot}W_{0\cdot} + s_{\cdot 0}W_{\cdot 0} + s_{00}W_{00}, \tag{15}$$

where the parameters to be estimated are just $v_S = 4$ for the scalar specification, and consist of $s_{..}$, $s_{0\cdot}$, $s_{\cdot 0}$, s_{00} in (15). For the diagonal specification their number is $v_S = 4n$.

2.4.3 Intermediate models and examples

Intermediate models are obtained by imposing a subset of the restrictions (11), (12), (13); this results in models with an intermediate number of parameters, i.e. between 4 and $k + \ell + k\ell + 1$ for the scalar specification.

Example 1 (HOM) *Take $k = 2$, $\ell = 1$ and two stocks for each class $\mathcal{C}_{ij}$, for a total of 4 assets. The assets are ordered following Table 1. (In case $\ell > 1$ we proceed row-wise). This gives*

$$W_{\cdot\cdot} = \begin{pmatrix} & 1 & 1 & 1 \\ 1 & & 1 & 1 \\ 1 & 1 & & 1 \\ 1 & 1 & 1 & \end{pmatrix} \qquad W_{1\cdot} = \begin{pmatrix} & 1 & & \\ 1 & & & \\ & & & \\ & & & \end{pmatrix} \qquad W_{2\cdot} = \begin{pmatrix} & & & \\ & & & \\ & & & 1 \\ & & 1 & \end{pmatrix}.$$

and the specification $S = s_{\cdot\cdot}W_{\cdot\cdot} + s_{1\cdot}W_{1\cdot} + s_{2\cdot}W_{2\cdot}$ results in

$$S = \begin{pmatrix} & a_1 & s_{\cdot\cdot} & s_{\cdot\cdot} \\ a_1 & & s_{\cdot\cdot} & s_{\cdot\cdot} \\ s_{\cdot\cdot} & s_{\cdot\cdot} & & a_2 \\ s_{\cdot\cdot} & s_{\cdot\cdot} & a_2 & \end{pmatrix},$$

where $a_i := s_{\cdot\cdot} + s_{i\cdot}$, $i = 1, 2$. Here and in the following we omit 0 entries unless needed for clarity.

Example 2 (RHOM) *Take $k = 2$, $\ell = 2$ and two stocks for each class $\mathcal{C}_{ij}$, for a total of 8 assets. The assets are ordered following Table 1 proceeding row-wise. This gives*

$$W_{\cdot\cdot} = \begin{pmatrix} & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & \end{pmatrix} \qquad W_{1\cdot}{+}W_{2\cdot} = \begin{pmatrix} & 1 & 1 & 1 & & & & \\ 1 & & 1 & 1 & & & & \\ 1 & 1 & & 1 & & & & \\ 1 & 1 & 1 & & & & & \\ & & & & & 1 & 1 & 1 \\ & & & & 1 & & 1 & 1 \\ & & & & 1 & 1 & & 1 \\ & & & & 1 & 1 & 1 & \end{pmatrix},$$

$$W_{\cdot 1}{+}W_{\cdot 2} = \begin{pmatrix} & 1 & & & 1 & 1 & & \\ 1 & & & & 1 & 1 & & \\ & & & 1 & & & 1 & 1 \\ & & 1 & & & & 1 & 1 \\ 1 & 1 & & & & 1 & & \\ 1 & 1 & & & 1 & & & \\ & & 1 & 1 & & & & 1 \\ & & 1 & 1 & & & 1 & \end{pmatrix} \qquad \sum_{i=1}^{2}\sum_{j=1}^{2} W_{ij} = \begin{pmatrix} & 1 & & & & & & \\ 1 & & & & & & & \\ & & & 1 & & & & \\ & & 1 & & & & & \\ & & & & & 1 & & \\ & & & & 1 & & & \\ & & & & & & & 1 \\ & & & & & & 1 & \end{pmatrix}.$$

and the specification (15) *results in the following structure for* S*:*

$$S=\begin{pmatrix} & s & a & a & b & b & s_{..} & s_{..}\\ s & & a & a & b & b & s_{..} & s_{..}\\ a & a & & s & s_{..} & s_{..} & b & b\\ a & a & s & & s_{..} & s_{..} & b & b\\ b & b & s_{..} & s_{..} & & s & a & a\\ b & b & s_{..} & s_{..} & s & & a & a\\ s_{..} & s_{..} & b & b & a & a & & s\\ s_{..} & s_{..} & b & b & a & a & s & \end{pmatrix},$$

where $s := s_{..} + s_{0\cdot} + s_{\cdot 0} + s_{00}$*,* $a := s_{..} + s_{0\cdot}$*,* $b := s_{..} + s_{\cdot 0}$*.*

In the following we usually normalize the spatial matrices W to have row sums equal to 1. In the above examples this normalization would just imply a rescaling of the parameters, given that the row sums of the various W matrices are all equal. This would not be the case if the number of assets in each class is different.

2.5 ARCH dynamics

We next consider the specification of the matrix polynomials $E(L) := \sum_{l=1}^{p_E} E_l L^{l-1}$ and $A(L) := \sum_{l=1}^{p_A} A_l L^{l-1}$ in (5), i.e. $h_t = c + E(L)h_{t-1} + A(L)e_{t-1}$. The main interaction *between* classes $\mathcal{C}_{ij}, \mathcal{C}_{\cdot j}, \mathcal{C}_{i\cdot}$, see (9), is addressed via the S-specification. The ARCH dynamics in (5) is hence assumed to reflect possible dependencies *within* classes $\mathcal{C}_{ij}$.

Many criteria may be applied to classify assets within each class $\mathcal{C}_{ij}$. Some possibilities are given by earnings before income and taxes, dividend/price, dividend/earning ratios etc. In absence of further information, the single ordering criterion within each class $\mathcal{C}_{ij}$ may be given by capitalization level.

For simplicity, we here assume that a single criterion is used to order assets within each class $\mathcal{C}_{ij}$. Without loss of generality, assume that assets within the blocks for class $\mathcal{C}_{ij}$ are ordered as in Examples 1 and 2, i.e. the first n_{11} assets belong to the class $\mathcal{C}_{11}$, the next n_{12} belong to the class $\mathcal{C}_{12}$, and so forth, proceeding row-wise with respect to Table 1, where n_{ij} indicates the number of assets in class $\mathcal{C}_{ij}$, $n_{ij} := \#(\mathcal{C}_{ij})$, and $\#(\cdot)$ indicates the cardinality of the argument set. Obviously $n = \sum_{i=1}^{k}\sum_{j=1}^{\ell} n_{ij}$.

Consider the partition of e_t into subvectors $e_t := (e'_{11,t},\ e'_{12,t},\ \ldots,\ e'_{1\ell,t},\ e'_{21,t},\ \ldots e'_{k\ell,t})'$, where $e_{ij,t}$ is the subvector of e_t corresponding to the class $\mathcal{C}_{ij}$. Partition also η_t, c and h_t conformably. We assume that the $n \times n$ matrices E_l, A_l matrices in $E(L) := \sum_{l=1}^{p} E_l L^{l-1}$ and $A(L) := \sum_{l=1}^{q} A_l L^{l-1}$ to be block diagonal, where blocks are conformable with the partition of e_t. We indicate the blocks of E_l and A_l corresponding to $e_{ij,t-1}$ as $E_{ij,l}$ and $A_{ij,l}$ respectively; similarly we indicate the corresponding blocks of $E(L)$, $A(L)$ as $E_{ij}(L)$, $A_{ij}(L)$.

We assume that assets within the subvector $e_{ij,t}$ are ordered according to a single ordering, which creates a spatial proximity within each block. We consider two main EA-specifications; the first one is of spatial nature, the second one allows for a factor structure. The latter case will be shown to be a special case of the former. In both cases the intercept c_{ij} is a linear function of some underlying vector of parameters $\theta_{c_{ij}}$,

$$c_{ij} = \mathcal{F}_{ij}\theta_{c_{ij}}. \tag{16}$$

We note in passing that the standard requirements for second order stationarity of GARCH processes apply, i.e. the roots of $|I - E(L) - A(L)| = 0$ must be outside the unit circle. By the block-diagonal assumption of $E(L)$ and $A(L)$, this corresponds to the requirement that the roots of $\left|I_{n_{ij}} - E_{ij}(L) - A_{ij}(L)\right| = 0$ are outside the unit circle.

2.5.1 Spatial EA dynamics

We assume that $E_{ij,l}$ and $A_{ij,l}$ have a spatial nature, i.e. they are of the following form:

$$E_{ij,l} := \sum_{q=0}^{m_{E_{ij,l}}} \beta_{ij,lq} W^*_{ij,q} \qquad A_{ij,l} := \sum_{q=0}^{m_{A_{ij,l}}} \alpha_{ij,lq} W^*_{ij,q}, \tag{17}$$

where $\beta_{ij,lq}$ and $\alpha_{ij,lq}$, are scalars or diagonal matrices. The matrices $W^*_{ij,q}$ for $q > 0$ are $n_{ij} \times n_{ij}$ spatial weight matrices that reflect proximity according to the intra-class ordering criterion. For the spatial EA dynamics we assume $W^*_{ij,0} := I_{n_{ij}}$ and $\mathcal{F}_{ij} = I_{n_{ij}}$ in (16).

For a single ordering criterion, two possible choices for the spatial matrices $W^*_{ij,q}$ are the following

$$W^*_{ij,q} := U_{ij}^{q\prime} \tag{18}$$

$$\text{or} \qquad W^*_{ij,q} := U_{ij}^{q} + U_{ij}^{q\prime} \qquad \text{where} \qquad \underset{n_{ij}\times n_{ij}}{U_{ij}} := \begin{pmatrix} 0 & 0 \\ I_{n_{ij}-1} & 0 \end{pmatrix}. \tag{19}$$

In the case of a scalar specification, this implies that the $E_{ij,l}$ and $A_{ij,l}$ matrices are Toeplitz matrices.

These specifications have the following interpretation: the spatial structure within each block relates each stock within $e_{ij,t}$ with the preceding one in the list; this is the case for specification (18), which implies an lower triangular system.

Alternatively, consider specification (19); each stock within $e_{ij,t}$ is related to the one preceding and the one following it in the list, which implies a symmetric Toeplitz matrix for the scalar specification. The lower triangular system can be obtained from (18) simply reversing the ordering within the block, and hence it is not treated separately.

The leading specification of the GARCH dynamics is the GARCH(1,1) system, which we consider in order to count parameters. Let $m_{EA_{ij}} := \max(m_{E_{ij,l}}, m_{A_{ij,l}}) \leq n_{ij}$ for specifications (18) and (19). The parameters are c_{ij}, $\beta_{ij,1q}$, $\alpha_{ij,1q}$ for $q = 0$, ..., $m_{EA_{ij}}$, possibly with zero elements. This gives $v_{EA} = \sum_{i=1}^{k} \sum_{j=1}^{\ell} \left(n_{ij} + 2 + m_{E_{l,ij}} + m_{A_{l,ij}}\right)$ for the EA-scalar specification. If $m_{EA_{ij}} = m_{EA}$ for all i, j, this number becomes $v_{EA} = n + 2\,(1 + m_{EA})\,k\ell$. The case with no spatial effect corresponds to $m_{EA} = 0$, where the number of parameters is simply $v_{EA} = n + 2k\ell$.

For the EA-diagonal specification one has to distinguish between specifications (18) and (19). In case of (18), the first row of U is a zero vector, because the first asset has no preceding asset in the classification. This implies that the first elements in $\beta_{ij,1q}$ is constrained to 0 and the remaining $n_{ij}-1$ elements are free. The same applies to $\alpha_{1q,ij}$.

When $m_{E_{ij,l}} = m_{A_{ij,l}} = m_{EA}$ this gives a total number of parameters for the EA-diagonal specification equal to $v_{EA} = \sum_{i=1}^{k}\sum_{j=1}^{\ell}\left(n_{ij}+2\left(n_{ij}-1\right)m_{EA}\right)\right) = n+2\left(n-k\ell\right)m_{EA}$. For the EA-diagonal specification (19) one has instead $v_{EA} = \sum_{i=1}^{k}\sum_{j=1}^{\ell} n_{ij}\left(1+2m_{EA}\right) = n\left(1+2m_{EA}\right)$. Intermediate cases are characterized by an intermediate number of parameters.

The conditions for positive definiteness of the conditional variance matrix require the EA dynamics to deliver always positive definite conditional variances h_t. A sufficient condition for this is that the c, α and β parameters to be positive.

2.5.2 Factor EA dynamics

In this subsection we consider a factor GARCH model for the dynamics within each class $\mathcal{C}_{ij}$, with a time-varying component driven a lower-dimensional GARCH process. This specification implies common features in variance in the sense of Engle and Kozicki (1993).

In Appendix A we show that a GARCH(1,1) specification that embodies this factor structure is given by

$$h_{ij,t} = \xi_{ij,\perp} q_{ij} + \overset{\circ}{\gamma}_{ij,\perp} h_{ij2,t}, \tag{20}$$

$$h_{ij2,t} = \omega_{ij} + \alpha_{ij}\xi_{ij}' e_{ij,t-1} + \beta_{ij} h_{ij2,t-1} \tag{21}$$

where q_{ij} is $n_{ij}-r_{ij}\times 1$ and ω_{ij} is $r_{ij}\times 1$, $\overset{\circ}{\gamma}_{ij,\perp} := \gamma_{ij,\perp}(\xi_{ij}'\gamma_{ij,\perp})^{-1}$ and $h_{ij2,t} := \xi_{ij}' h_{ij,t}$, both of dimension $r_{ij}\times 1$, where $r_{ij} < n_{ij}$.

Often one wishes to specify a one-factor model for each class $\mathcal{C}_{ij}$, of the GARCH type, $r_{ij}=1$, possibly pre-specifying also γ_{ij} and ξ_{ij} as follows

$$\xi_{ij} = \gamma_{ij,\perp} = \iota_{n_{ij}}. \tag{22}$$

This leaves $n_{ij}-1$ free elements to estimate in q_{ij} in (20) and 3 parameters in (21), given by ω_{ij}, α_{ij}, β_{ij}. Overall this EA-scalar specification has number of parameters in the ARCH dynamics equal to $v_{EA} = \sum_{i=1}^{k}\sum_{j=1}^{\ell}\left(n_{ij}+2\right) = n+2k\ell$. Note that the EA-diagonal specification is not available in this case, because in this case the $E(L)$ and $A(L)$ polynomial do not present a spatial nature.

Under constraint (22) $\xi_{ij} = \gamma_{ij,\perp} = \iota_{n_{ij}}$, one can choose $\xi_{ij,\perp}$ as $\xi_{ij,\perp} = \left(I_{n_{ij}} - U\right) Q_{ij} =: b_{ij}$, where U is defined in (19) and Q_{ij} is the $n_{ij}\times(n_{ij}-1)$ matrix containing the first $n_{ij}-1$ columns of $I_{n_{ij}}$. Specification (20) (21) becomes

$$\begin{aligned} h_{ij,t} &= b_{ij} q_{ij} + n_{ij}^{-1}\iota_{n_{ij}} h_{ij2,t}, \\ h_{ij2,t} &= \omega_{ij} + \alpha_{ij}\iota_{n_{ij}}' e_{ij,t-1} + \beta_{ij} h_{ij2,t-1}. \end{aligned}$$

Note that by construction $h_{ij2,t} := \iota_{n_{ij}}' h_{ij,t}$ because $\iota_{n_{ij}}' b_{ij} = 0$. Let also $J_{ij} := \frac{1}{n_{ij}}\iota_{n_{ij}}\iota_{n_{ij}}' = \iota_{n_{ij}}\left(\iota_{n_{ij}}'\iota_{n_{ij}}\right)^{-1}\iota_{n_{ij}}'$. Substituting the second equation into the first one and setting $h_{ij2,t-1} := \iota_{n_{ij}}' h_{ij,t-1}$ one finds

$$h_{ij,t} = \mathcal{F}_{ij} c_{ij}^{*} + \left(\beta_{ij} J_{ij}\right) h_{ij,t-1} + \left(\alpha_{ij} J_{ij}\right) e_{ij,t-1}, \tag{23}$$

where $\theta_{c_{ij}} := \left(q'_{ij} : \omega_{ij}\right)'$ is $n_{ij} \times 1$ and $\mathcal{F}_{ij} := \left(b_{ij} : n_{ij}^{-1}\iota_{n_{ij}}\right)$ This shows that the factor EA specification is a different submodel of (5) than (17) with

$$E_{ij,1} = \beta_{ij} J_{ij}, \qquad A_{ij,1} = \alpha_{ij} J_{ij}, \tag{24}$$

and $\mathcal{F}_{ij} := \left(b_{ij} : n_{ij}^{-1}\iota_{n_{ij}}\right)$ in (16).

Formally we can define $W^*_{ij,0} := J_{ij}$ and consider (24) a special case of the scalar specification (17) with $m_{E_{ij,1}} = 0$, $m_{A_{ij,1}} = 0$, $W^*_{ij,0} := J_{ij}$ and $\mathcal{F}_{ij} := \left(b_{ij} : n_{ij}^{-1}\iota_{n_{ij}}\right)$. In the following we treat the factor EA specification as a special case of the EA spatial dynamics (17) for this particular choice of $m_{E_{ij,1}}$, $m_{A_{ij,1}}$, $W^*_{ij,0}$, $\mathcal{F}_{ij}$.

2.6 R specification

Consider next the matrix R, a positive definite, symmetric matrix with ones on the main diagonal. Again assuming the same ordering of the blocks as in the previous subsection, we assume that R is block diagonal with diagonal blocks R_{ij}, where the subscripts ij refer to the class $\mathcal{C}_{ij}$.

As in the previous subsection recall that assets are ordered within the subvector $e_{ij,t}$ according to a single ordering critierion. Within each diagonal block R_{ij}, we consider the following spatial specification:

$$R_{ij} = I_{n_{ij}} + \sum_{q=1}^{m_{R_{ij}}} \rho_{ij,q} W^*_{ij,q} \qquad \text{where} \qquad W^*_{ij,q} := U^q_{ij} + U^{q\prime}_{ij}, \tag{25}$$

and U_{ij} is defined in (19), $m_{R_{ij}} \leq n_{ij} - 1$. Note that the spatial nature of the R specification matches the one for the EA dynamics. In fact both reflect the same intra-class classification.

For example for $n_{ij} = 4$ one has

$$R_{ij} = \begin{pmatrix} 1 & \rho_{ij,1} & \rho_{ij,2} & \rho_{ij,3} \\ \rho_{ij,1} & 1 & \rho_{ij,1} & \rho_{ij,2} \\ \rho_{ij,2} & \rho_{ij,1} & 1 & \rho_{ij,1} \\ \rho_{ij,3} & \rho_{ij,2} & \rho_{ij,1} & 1 \end{pmatrix}.$$

The number of parameters in the R-scalar specification (25) is $n_{ij} - 1$ for each class $\mathcal{C}_{ij}$, for a total number of parameters equal to $v_R = \sum_{i,j}(n_{ij} - 1) = n - k\ell$. Note that the R-diagonal specification does not make sense, given the symmetry of each block R_{ij}. We hence consider only the scalar R specification.

A special case of (25) is given by the spatial autoregression of order one, where $\rho_{ij,q} = \rho^q_{ij}$ for some scalar ρ_{ij}. In this case the number of parameters reduces to 1 for each class $\mathcal{C}_{ij}$, for a total number of parameters equal to $v_R = k\ell$.

2.7 Extensions

A number of possible extensions and variations can be considered on the basic scheme proposed in the previous sub-sections. We here list a few. These extensions go beyond the scope of the present paper, and they are not considered further, even if they well deserve attention in future research on the issue.

1. One may consider a conditional spatial autoregressive (CAR) scheme in place of S for the spatial autoregression defined in (2), see also Cressie (1993) eq. (6.3.13) and (6.3.8) respectively for definitions. The CAR scheme is more restrictive than the SAR one, but represents an important alternative approach.

2. The GARCH specification for the dynamics of the conditional variances can be easily extended to allow for asymmetries in the news impact curve, non-linear transformations of the input and output of the GARCH dynamics. The latter transformations, however, come at the expense of preventing the use of recursion in the multi-step ahead prediction of variances.

3. More complicated spatial weight matrices W may be defined from multivariate metric distances computed on the basis of quantitative indicators. For example one may define $W = (w_{ij})$ with $w_{ij} = 1/d_{ij}$, $i \neq j$, $w_{ii} = 0$, where d_{ij} is a metric distance between units i and j. Different distances d_{ij} can be defined on the analysis of balance–sheet data.

4. The two-way classification used in the S-specification may be directly extended to more complicated multi-way classifications. More (or less) than 2 factors may be used in the factorial design in the definition of the matrices S, E_l, A_l and R. One may expect that different spatial classifications and orderings may give rise to non-nested SEARCH models. This raises the question of testing non nested specifications in the present context.

5. The present specification constrains R and S to be time-invariant. This is the simplest option, which could be relaxed in many ways. One may e.g. allow the parameters θ_S in S to be time-varying, possibly with GARCH dynamics. A time-varying spatial structure of S could model periods of higher and lower spatial dependence.

3 Relation with other specifications

In this section we discuss the relation of the SEARCH model with other GARCH specifications, including the CC and the OG specifications. The general model we consider, which we call the generalized variance model GV, consists of the following general specification of the first and second conditional moments

$$\begin{aligned}
y_t &=: \mathbb{E}_{t-1}(y_t) + \varepsilon_t = \mu w_t + \varepsilon_t, \\
\Sigma_t &:= \mathbb{E}_{t-1}(\varepsilon_t \varepsilon_t') = V D_t R D_t V', \\
h_t &:= \operatorname{vecd}\left(D_t^2\right) = c + E(L) h_{t-1} + A(L) e_{t-1}, \\
e_t &:= \operatorname{vecd}\left(V^{-1} \varepsilon_t \varepsilon_t' V^{-1\prime}\right),
\end{aligned} \tag{26}$$

and D_t is a diagonal matrix. We assume that the process is second order stationary, so that $\Sigma := \mathbb{E}(\varepsilon_t \varepsilon_t')$ is also well defined. We first show in Subsection 3.1 how several different specifications are nested within (26). In subsection 3.2 we compare model dimensions.

3.1 CC, OG, factor GARCH and SEARCH

The general specification (26) nests various submodels. Specifically:

1. When taking $R = I$ and $E(L)$, $A(L)$ diagonal and $V = P\Lambda^{1/2}$, with P and Λ respectively equal to the matrix of eigenvectors and eigenvalues of Σ, one obtains the standard OG specification. The parameters to be estimated consist of the diagonal elements in $E(L)$, $A(L)$ and the $n(n+1)/2$ free elements in V.

2. When taking $R = I$ and $E(L)$, $A(L)$ diagonal, V equal to $P\Lambda^{1/2}U_0$, with P and Λ respectively equal to the matrix of eigenvectors and eigenvalues of Σ, and U_0 as an orthogonal matrix, one obtains the generalized OG specification of van Der Weide (2002). The parameters are the same as in the OG with the addition of $n(n-1)/2$ parameters in U_0.

3. When taking $R = I$ and $E(L)$, $A(L)$ diagonal, and V lower triangular with diagonal elements equal to 1, one obtains the full factor GARCH model of Vrontos et al (2003). The number of parameters in the GARCH dynamics is $n\,(1 + p_A + p_E)$ plus the $n(n-1)/2$ unrestricted elements in V.

4. When taking $V = I$ and $E(L)$, $A(L)$ diagonal, R unrestricted, one obtains the standard CC specification. The number of parameters is $n(1+p_A+p_E)$ in the ARCH dynamics with the addition of $n(n-1)/2$ parameters in R.

5. Consider the k-factor GARCH model of Engle, see Bollerslev and Engle (1993) eq. (12),
$$\Sigma_t = \Psi + C\Lambda_t C'$$
where Ψ is an $n \times n$ matrix (not necessarily diagonal), C is $n \times r$ and of full column rank, $\Lambda_t = \operatorname{diag}(h_{1t})$,
$$h_{1t} = E_1(L)\operatorname{vecd}\left(\overline{C}'\Sigma_t\overline{C}\right) + A_1(L)\operatorname{vecd}\left(\overline{C}'\varepsilon_t\varepsilon_t'\overline{C}\right). \qquad (27)$$
Here $\overline{C} := C(C'C)^{-1}$ for any matrix C. The matrix polynomials $E_1(L) := \sum_{i=1}^{p_E} E_{1i}L^i$ and $A_1(L) := \sum_{i=1}^{p_A} A_{1i}L^i$ are assumed to have diagonal structure. It is simple to verify, using orthogonal projections, that the model may be written as
$$\Sigma_t = V\left(\Phi + \operatorname{diag}(\Lambda_t, 0)\right)V'$$
with $V = (C : C_\perp)$. When Φ is diagonal, this model is in the format (26) with $R = I$, $V = (C : C_\perp)$ and $E(L) = \operatorname{diag}(E_1(L), 0)$, $A(L) = \operatorname{diag}(A_1(L), 0)$. Note in fact that $\eta_t = V^{-1}\varepsilon_t = \left(\overline{C} : \overline{C}_\perp\right)'\varepsilon_t$ in $e_t := \operatorname{vecd}(\eta_t\eta_t')$, so that this conforms with the definition of the innovations in (27). With Φ diagonal, the parameters to be estimated are Φ, C and the elements on the diagonal of $E_1(L)$ and $A_1(L)$, for a total of $n(r+1) + r(1 + p_A + p_E)$ parameters.

6. SEARCH specifications are obtained as the special cases where $V := (I - S)^{-1}$ and S, $E(L)$, $A(L)$, R are constrained to have a spatial structure. The parametrization of the SEARCH has been described at length in Section 2.

More general multivariate GARCH specifications are also reported here for comparison. Consider the vech GARCH formulation, see Engle and Kroner (1995); this is given by

$$\mathrm{vec}\left(\Sigma_t\right) = \mathrm{vec}\left(C\right) + \sum_{i=1}^{p_E} E_i^* \mathrm{vec}\left(\Sigma_{t-i}\right) + \sum_{i=1}^{p_A} A_i^* \mathrm{vec}\left(\varepsilon_{t-i}\varepsilon_{t-i}'\right). \tag{28}$$

Let $\mathcal{D}$ be the $n^2 \times u$ duplication matrix, $u := n(n+1)/2$, that satisfies $\mathrm{vec}\left(\Sigma_t\right) = \mathcal{D}\,\mathrm{vech}\left(\Sigma_t\right)$. Given the symmetry of Σ_t, the parameters in the GARCH vech specification are constrained to satisfy $\mathrm{vec}\left(C\right) = \mathcal{D}\,\mathrm{vech}\left(C\right)$ (symmetry of C), $E_i^* = \mathcal{D}E_i^\circ\overline{\mathcal{D}}'$, $A_i^* = \mathcal{D}A_i^\circ\overline{\mathcal{D}}'$ where E_i° and A_i° are $u \times u$.

In order to ensure the positive definiteness of Σ_t, the BEKK specification moreover requires

$$E_i^* = \sum_{j=1}^{m_{E_i^*}} \left(E_{ij} \otimes E_{ij}\right) \qquad \text{and} \qquad A_i^* = \sum_{j=1}^{m_{A_i^*}} \left(A_{ij} \otimes A_{ij}\right) \tag{29}$$

where E_{ij} and A_{ij} are $n \times n$ real matrices. In the following we take for simplicity $m_{E_i^*} = m_{A_j^*} =: m_{EA}$ for all i and j.

We next consider model dimensions.

3.2 Model dimensions

The models surveyed in the previous subsection have very different dimensions. The model dimensions $v$are decomposed into the number of parameters in the conditional mean specification v_1 and in the number of parameters in the conditional variance v_2; $v = v_1 + v_2$.

The different specifications are compared for: an equal number of parameters for the conditional mean specification v_1; an equal number of lags in the GARCH specification, indicated as p_E and p_A; for an equal number of terms in the SEARCH specification of E_{ij} and A_{ij}, indicated as $m_{E_{ij}} = m_{A_{ij}} = m_{EA}$ for all i and j. This number m_{EA} is assumed to coincide also with $m_{E_i^*} = m_{A_i^*}$ in the BEKK specification (29).

The model dimensions v_2 for the conditional variances are compared in Table 2. The following remarks are in order.

1. The SEARCH specifications are linear in n, the number of assets.
2. The only other specification that is also linear in n is the factor GARCH, assuming that the number of factors r is independent of n.
3. The dimension of the OG class is quadratic in n, due to the estimation of the $n \times n$ matrix of eigenvectors V of Σ. This also holds for the generalized OG, which has more parameters.
4. The unrestricted CC estimates the parameters in the correlation matrix R unrestrictedly, and this makes the model quadratic in n.

5. One could also consider a constrained version of the BEKK specification, restricting the parameters E_{ij}, A_{ij} in (29) to be of spatial nature, in the sense of Section 2. This would reduce the number of parameters in E_i^* and A_i^* to be linear in n. However the model dimension would still be quadratic in n when the constant matrix C is estimated unrestrictedly. When also C is restricted to be of spatial nature, the model dimension becomes linear in n. However the requirement that C is of spatial nature may be too restrictive.

4 Likelihood based inference

In this section we consider (quasi-) maximum likelihood, QML, estimation of the SEARCH model with the following GARCH(1,1) specification

$$\begin{aligned}
&y_t =: \mathbb{E}_{t-1}(y_t) + \varepsilon_t = \mu w_t + \varepsilon_t, \\
&\Sigma_t := \mathbb{E}_{t-1}(\varepsilon_t\varepsilon_t') = (I-S)^{-1} D_t R D_t (I-S')^{-1}, \\
&h_t = c + E h_{t-1} + A e_{t-1}. \\
&S = \sum_{i=1}^{m_S} s_i W_i \qquad\qquad E = \operatorname{diag}(E_{11}, E_{12}, ..., E_{k\ell}) \\
&A = \operatorname{diag}(A_{11}, A_{12}, ..., A_{k\ell}) \qquad R = \operatorname{diag}(R_{11}, R_{12}, ..., R_{k\ell})
\end{aligned}$$

and the specification of the blocks E_{ij}, A_{ij} is defined in (17) or (20) and (21) and the one for R_{ij} in (25).[4]

The parameter vector θ is partitioned as $\theta = (\theta_\mu' : \theta_S' : \theta_{EA}' : \theta_R')'$ where the subvectors contain parameters in the μ, S, EA and R specifications, see Section 2. The θ_{EA} subvector is partitioned as $\theta_{EA} = (\theta_c' : \theta_E' : \theta_A')'$, where for the spatial-$EA$ specification in Subsection 2.5.1 θ_c contains c_{ij}, whereas for the factor EA specification, c contains c_{ij}^*, see Subsection 2.5.2 and Appendix B.

When ε_t is assumed conditionally Gaussian, the log-likelihood function is $\ln L(\theta) = \sum_{t=1}^T \ln f_t(\theta)$, where $\ln f_t(\theta) := \ln f(y_t|\mathcal{I}_{t-1}, \theta)$ is given by

$$\ln f_t(\theta) = -\frac{1}{2}\left(\ln\det(\Sigma_t) + \operatorname{tr}\left(\Sigma_t^{-1}\varepsilon_t\varepsilon_t'\right)\right) = -\frac{1}{2}(f_{1t} + f_{2t}), \text{ say}$$

with

$$f_{1t} := \ln\det(\Sigma_t) = -2\ln\det(I-S) + \ln\det(R) + \ln\det\left(D_t^2\right)$$

and

$$\begin{aligned}
f_{2t} &:= \operatorname{tr}\left(\Sigma_t^{-1}\varepsilon_t\varepsilon_t'\right) = \operatorname{tr}\left(\Gamma_t^{-1}(I-S)\varepsilon_t\varepsilon_t'(I-S')\right) = \\
&= \operatorname{tr}\left(R^{-1}D_t^{-1}\eta_t\eta_t' D_t^{-1}\right) = \operatorname{tr}\left(R^{-1}\psi_t\psi_t'\right).
\end{aligned}$$

Derivatives of $\ln f_t(\theta)$ with respect to the parameters in $\theta = (\theta_\mu' : \theta_S' : c' : \theta_E' : \theta_A' : \theta_R')'$ are given in Appendix B.

In the following subsections we discuss identification in Subsection 4.1. In Subsection 4.2 we present possible ways to maximize the Gaussian likelihood, while in Subsection 4.3 we discuss LM tests for SEARCH.

[4]Extensions to ARCH dynamics of higher orders are straightforward and hence omitted.

Model: v_2 (num. of par. in $\mathbb{V}_{t-1}(\varepsilon_t)$)	Order q in $v_2 = O(n^q)$	Leading term in n
GARCH vech, eq. (28): $\frac{1}{4}(p_E + p_A)\, n^2 (n+1)^2 + \frac{1}{2} n (n+1)$	4	$\frac{1}{4}(p_E + p_A)\, n^4 + O(n^2)$
BEKK, eq. (29): $m_{EA}(p_E + p_A)\, n^2 + \frac{1}{2} n (n+1)$	2	$\left(m_{EA}(p_E + p_A) + \frac{1}{2}\right) n^2 + O(n)$
CC: $(1 + p_E + p_A)\, n + \frac{1}{2} n (n+1)$	2	$\frac{1}{2} n^2 + O(n)$
OG: $(p_E + p_A)\, n + \frac{1}{2} n (n+1)$	2	$\frac{1}{2} n^2 + O(n)$
GOG: $(p_E + p_A)\, n + n^2$	2	$n^2 + O(n)$
r-factors GARCH, eq. (27): $(p_E + p_A + n)\, r + n$	1	$(r+1)\, n + O(1)$
SEARCH		
diag, hetero S, spatial EA: $n\left((k+1)(\ell+1) + 2 + 2m_{EA}\right) - k\ell$	1	$(k\ell + \ell + k + 3 + 2m_{EA})\, n + O(1)$
diag, hetero S, factor EA: $n(k+1)(\ell+1) + 2n + k\ell$	1	$(k\ell + \ell + k + 3)\, n + O(1)$
diag, homo S, spa EA: $n(6 + 2m_{EA}) - k\ell$	1	$(6 + 2m_{EA})\, n + O(1)$
diag, homo S, factor EA: $6n + k\ell$	1	$6n + O(1)$
scalar, hetero S, spatial EA: $(k+1)(\ell+1) + 2n+$ $+(1 + 2m_{EA})\, k\ell$	1	$2n + O(1)$
scalar, hetero S, factor EA: $(k+1)(\ell+1) + 2n + k\ell$	1	$2n + O(1)$
scalar, homo S, spatial EA: $k + \ell + 1 + 4 + n(2 + 2m_{EA})$	1	$2(1 + m_{EA})\, n + O(1)$
scalar, homo S, factor EA: $(k+1)(\ell+1) + 4 + 2n + k\ell$	1	$2n + O(1)$

Table 2: Model dimensions as a function of n; the lower panel contains SEARCH specifications.

4.1 Identification

Identification in volatility models has been addressed in Sentana and Fiorentini (2001) and more recently in Lucchetti (2004). We note that, when the ARCH dynamics is shut down, $E_l = 0$, $A_l = 0$, for the SEARCH specification in S, c, and R to be identified, one needs to have $n_{ij} > 2$. In the following we assume that models are identified.

4.2 Maximization

The Gaussian likelihood is well defined and continuously differentiable wrt the parameters $\theta = (\theta'_\mu : \theta'_S : \theta'_c : \theta'_E : \theta'_A : \theta'_R)'$. This allows to perform QML estimation using gradient methods, like BGFS, or even of Newton-Raphson type. These methods require to perform the maximization of the likelihood jointly wrt to all the parameters.

For a large cross-section dimension n, this feature may be seen as a cost in terms of speed of computations, and simpler methods may be sought. In the next subsection we describe a alternating algorithm that may be employed to maximize the likelihood which requires considerably fewer computations in each iteration (possibly at the expense of more iterations).

The switching algorithm is based on the alternating maximization of the likelihood with respect to two disjoint and complementary subset of parameters. In particular we consider the following steps:

1. For fixed μ, S consider each single block $h_{ij,t}$ in turn and maximize wrt the parameters in c_{ij}, E_{ij}, A_{ij}, R_{ij}.

2. For fixed c, E, A, R, maximize wrt μ, S.

We now illustrate computations for each step. Consider step 1. In f_{1t} the only part that depends on parameters is $\ln\det\left(D_t^2\right) = \sum_{i=1}^{k}\sum_{j=1}^{\ell} f_{1ij,t}$ where $f_{1ij,t} = \sum_{q=1}^{n_{ij}} \ln h_{ijq,t}$ and $h_{ijq,t}$ is the q-th element in $h_{ij,t}$. Note also that $\eta_t = (I - S)\,\varepsilon_t$ are known constants in f_{2t}, which, thanks to the block diagonal structure of R and the diagonal structure of D_t, decomposes into

$$f_{2t} = \operatorname{tr}\left(R^{-1}D_t^{-1}\eta_t\eta_t' D_t^{-1}\right) = \sum_{i=1}^{k}\sum_{j=1}^{\ell} f_{2ij,t}$$

where $f_{2ij,t} = \eta'_{ij,t}D_{ij,t}^{-1}R_{ij}^{-1}D_{ij,t}^{-1}\eta_{ij,t}$. Further note that $\ln f_t(\theta)$ depends on $\theta_{c_{ij}}$, E_{ij}, A_{ij}, R_{ij} only through $f_{1ij,t}$ and $f_{2ij,t}$, so that one may write

$$\ln f_t(\theta) = c + \sum_{i=1}^{k}\sum_{j=1}^{\ell} f_{ij,t} \qquad f_{ij,t} := -\frac{1}{2}\left(f_{1ij,t} + f_{2ij,t}\right)$$

and maximization of $\sum_{t=1}^{T} \ln f_t(\theta)$ wrt c_{ij}, E_{ij}, A_{ij}, R_{ij} is obtained by considering the separate subsystems consisting of $\varepsilon_{ij,t}$ and $h_{ij,t}$. This step can take advantage of existing software for the estimation of GARCH processes.

Next we consider the second step, which maximize wrt μ, S for fixed c, E, A, R. In this step the maximization involves all the cross section. However, the number of parameters is here limited to the parameters of the conditional mean and S. In particular the number v_S of parameters is very small, speeding up calculations.

As with any alternating algorithm, this switching algorithm ensures that the likelihood is increased in each step.

4.3 LM tests

The asymptotic theory for the quasi maximum likelihood (QML) estimator follows directly from the results in Ling and McAleer (2003), see also Comte and Lieberman (2003). Under second order stationarity requirements and under the assumption of the existence of moments up to order 8, the QML estimator is $\sqrt{T}$ asymptotically Gaussian with covariance matrix $A_0^{-1}B_0A_0^{-1}$ where $A_0 := \mathbb{E}\left(\frac{\partial^2 \ln f_t(\theta_0)}{\partial\theta\partial\theta'}\right)$, $B_0 := \mathbb{E}\left(\frac{\partial \ln f_t(\theta_0)}{\partial\theta}\frac{\partial \ln f_t(\theta_0)}{\partial\theta'}\right)$ under regularity conditions.

These asymptotics provide the basis for the development of LM tests for SEARCH effects. We consider the univariate GARCH specification as a starting point. This corresponds to setting $S = 0$, $E_1 = \text{diag}\,(\xi_1, ..., \xi_n)$, $A_1 = \text{diag}\,(\alpha_1, ..., \alpha_n)$, $R = I$. Let $\widetilde{\theta}$ be the corresponding restricted QML estimator of θ.

We consider the LM statistics

$$LM = \iota' X(X'X)^{-1}X'\iota \qquad \text{where } X := \left(\frac{\partial \ln f_1\left(\widetilde{\theta}\right)}{\partial\theta} : ... : \frac{\partial \ln f_T\left(\widetilde{\theta}\right)}{\partial\theta}\right)', \tag{30}$$

and ι is a vector of ones and $\widetilde{\theta}$ is the restricted ML estimator, where the expressions for G are given in Appendix B. This is the usual outer-product of the gradient (OPG) form of LM tests, which can be performed using artificial regression.

Under the regularity condition for the QML estimator, LM converges to a χ^2 with degrees of freedom equal to the number of constraints.

A different LM test can be applied for nested SEARCH specifications. As an example, one can derive LM tests for (some or all) the homogeneity restrictions (11), (12), (13) based on the estimation of the homogeneous S specification (15). Indicate in fact with $\widetilde{\theta}$ the ML estimator of a SEARCH specification (15). One can test the hypotheses (11), (12), (13) jointly by employing the test statistics (30). Alternatively each of these restrictions may be tested using the appropriate subset of the score vector. These LM test can be robustified by using the technique in Bollerslev and Wooldridge (1992).

5 Empirical application

This section present an empirical application of the proposed models. We selected 20 assets from the New York Stock Exchange; the stocks are listed in Table 3. The data were downloaded from Yahoo!Finance. The market capitalization refers to the beginning of September 2004, while the downloaded series cover the range January 1994 to June 2001.

Ticker	Asset	Sub-sector	Cap.	$\mathcal{C}_{ij}$
DD	E.I. Du Pont De Nemours & Co.[a]	CPR	42600	$\mathcal{C}_{11}$
DOW	Dow Chemical Co.[a]	CPR	40200	$\mathcal{C}_{11}$
KMB	Kimberly-Clark Corp[a]	PPP	33700	$\mathcal{C}_{11}$
AA	Alcoa Inc.[a]	MM	28600	$\mathcal{C}_{11}$
NEM	Newmont Mining Corp. [a]	GS	19400	$\mathcal{C}_{11}$
VAL	The Valspar Corp.[b]	CM	2500	$\mathcal{C}_{12}$
SON	Sonoco Products Co.[b]	PPP	2500	$\mathcal{C}_{12}$
SMG	The Scotts Co.[b]	CM	2100	$\mathcal{C}_{12}$
BOW	Bowater Inc.[b]	PPP	2000	$\mathcal{C}_{12}$
LZ	Lubrizol Corp.[b]	CM	1900	$\mathcal{C}_{12}$
HIN	Honeywell International Inc[a]	AD	31300	$\mathcal{C}_{21}$
ITW	Illinois Tool Works Inc.[a]	MCG	28100	$\mathcal{C}_{21}$
LMT	Lockheed Martin Corp.[a]	AD	24300	$\mathcal{C}_{21}$
NOC	Northrop Grumman Corp.[a]	AD	18400	$\mathcal{C}_{21}$
DE	Deere and Co-[a]	CAM	15700	$\mathcal{C}_{21}$
TOL	Tell Brothers Inc.[b]	CS	3400	$\mathcal{C}_{22}$
SPW	SPX Corp.[b]	MCG	2700	$\mathcal{C}_{22}$
ATK	Alliant Techsystems Inc.[b]	AD	2200	$\mathcal{C}_{22}$
HOV	Hovnanian Enterprises Inc.[b]	CS	2200	$\mathcal{C}_{22}$
JEC	Jacobs Engineering Group Inc.[b]	CS	2200	$\mathcal{C}_{22}$

Table 3: Assets included in the analysis; top 2 panels: Basic Materials, bottom 2 panels: Capital Goods. a: Index SP500, b: Index SP400. Sub-sectors: AD: Aerospace & Defense; CAM: Constr. & Agric. Machinery; CM: Chemical Manufacturing; CPR: Chemicals-Plastics & Rubber; CS: Construction & Services; GS: Gold & Silver; MCG: Misc. Capital Goods; MM: Metal Mining; PPP: Paper & PaperProducts.

Missing data across assets were replaced with a zero return in order to get an homogeneous sample. The data were also corrected for stock splits and dividends. The analysis was then performed on the log-returns $r_t = \ln p_t - \ln p_{t-1}$.

The estimation period covered January 1994 to December 1999 while the subsample from January 2000 to June 2001 was used for out-of-sample analysis. The two sample periods include 1513 and 377 dates, respectively.

The conditional mean specification included the following information variables z_t: the lagged log-returns of the Standard & Poor's 500 Index, the lagged first difference of the 3 Months Treasury Bills and of the 10 Years Notes, the lagged log-returns of Oil Prices (Texas), a set of dummy variables for the day of the week effect and the January effect.

For computational simplicity, we fixed μ at the estimated values from the ML (OLS) estimation on the conditional mean with conditional homoscedasticity. We calculated deviations from the conditional mean, and fitted a simple CC model on them, with an unrestricted R matrix.

We next specified a SEARCH model with homogeneous S specification (15), with spatial matrices defined as in (14). The classification was based on industrial sector and capitalization level, as described in Subsection 2.4. The four classes $\mathcal{C}_{ij}$ thus corresponded to the following associations. $\mathcal{C}_{11}$: Basic materials - Large capitalization, $\mathcal{C}_{12}$: Basic materials - Medium capitalization , $\mathcal{C}_{21}$: Capital goods - Large capitalization, $\mathcal{C}_{22}$: Capital goods - Medium capitalization. Assets within each class $\mathcal{C}_{ij}$ were ordered on the basis of capitalization level.

We next calculated the LM test against a SEARCH specification with homogeneous and scalar S, and scalar spatial EA specification, leaving R unrestricted. The LM test has been computed in the robustified version suggested by Bollerslev and Wooldridge (1992). This gave a LM test statistic of 517.761, which, when compared with a $\chi^2(44)$, gives a p-value inferior to 10^{-5}, thus suggesting the presence of significant spatial effects.

We next estimated a SEARCH model with scalar S-specification and scalar spatial EA specification, constraining R to the specification (25). The homogeneous-scalar-S and scalar-EA model includes 56 coefficients (4 in S, 20 in c, 8 in E and in A and 16 in R).

The estimated coefficients were obtained by QML using the BFGS algorithm and the analytical gradients included in the Appendix B. The coefficients standard errors were computed with the sandwich estimator. The estimates and asymptotic standard errors are reported in Tables 4. It can be seen that spatial effects are present and significant. Graphs of the estimated conditional variances and covariances for the SEARCH and CCC specifications are reported in Fig. 1 to Fig. 4.

We also considered the diagonal specifications for S and EA. A robust LM test for the given specification against a homogeneous diagonal S specification with scalar-EA specifications was considered; the test implies 56 constraints. The LM statistic was equal to 180.628 with an p-value less than 10^{-5} when compared with a $\chi^2(56)$ distribution. This evidence points to the generalization of S to a more a HET specification.

We also considered an LM test of the specification in Tables 4 and **??** against a homogeneous scalar S specification, with a diagonal EA specification. The test

S	$s_{..}$	$s_{.0}$	$s_{0.}$	s_{00}				
	0.02	0.01	0.01	-0.03				
	(462.53)	(160.72)	(129.66)	(-394.39)				
c	$c_{11,1}$	$c_{11,2}$	$c_{11,3}$	$c_{11,4}$	$c_{11,5}$			
	0.20	0.10	0.20	0.24	0.46			
	(30.83)	(91.75)	(74.66)	(79.50)	(30.73)			
	$c_{12,1}$	$c_{12,2}$	$c_{12,3}$	$c_{12,4}$	$c_{12,5}$			
	0.44	0.18	0.44	0.53	0.69			
	(62.71)	(50.22)	(59.59)	(58.85)	(61.80)			
	$c_{21,1}$	$c_{21,2}$	$c_{21,3}$	$c_{21,4}$	$c_{21,5}$			
	0.52	0.35	0.21	0.16	0.89			
	(93.60)	(79.98)	(61.42)	(63.87)	(96.82)			
	$c_{22,1}$	$c_{22,2}$	$c_{22,3}$	$c_{22,4}$	$c_{22,5}$			
	0.80	0.71	0.34	1.31	0.45			
	(104.11)	(100.69)	(85.95)	(109.07)	(99.44)			
E	$\beta_{11,0}$	$\beta_{12,0}$	$\beta_{21,0}$	$\beta_{22,0}$	$\beta_{11,1}$	$\beta_{12,1}$	$\beta_{21,1}$	$\beta_{22,1}$
	0.81	0.75	0.53	0.69	0.00	0.00	0.10	0.00
	(150.35)	(230.55)	(119.49)	(329.16)	(0.00)	(0.00)	(62.78)	(0.00)
A	$\alpha_{11,0}$	$\alpha_{12,0}$	$\alpha_{21,0}$	$\alpha_{22,0}$	$\alpha_{11,1}$	$\alpha_{12,1}$	$\alpha_{21,1}$	$\alpha_{22,1}$
	0.11	0.10	0.16	0.15	0.01	0.01	0.00	0.00
	(47.11)	(107.46)	(142.42)	(184.94)	(15.21)	(48.43)	(37.13)	(23.30)
R	$\rho_{11,1}$	$\rho_{11,2}$	$\rho_{11,3}$	$\rho_{11,4}$	$\rho_{12,1}$	$\rho_{12,2}$	$\rho_{12,3}$	$\rho_{12,4}$
	0.15	0.10	0.09	-0.02	0.03	0.03	0.07	0.09
	(404.01)	(236.51)	(185.24)	(-18.20)	(74.04)	(71.19)	(151.64)	(123.76)
	$\rho_{21,1}$	$\rho_{21,2}$	$\rho_{21,3}$	$\rho_{21,4}$	$\rho_{22,1}$	$\rho_{22,2}$	$\rho_{22,3}$	$\rho_{22,4}$
	0.16	0.11	0.16	0.16	0.04	0.03	0.11	0.08
	(402.57)	(260.61)	(297.79)	(212.13)	(134.16)	(59.06)	(243.67)	(115.30)

Table 4: SEARCH estimates; t-ratios in parenthesis, $c_{ij,h}$ are the estimates of the GARCH intercept for asset h in class $\mathcal{C}_{ij}$.

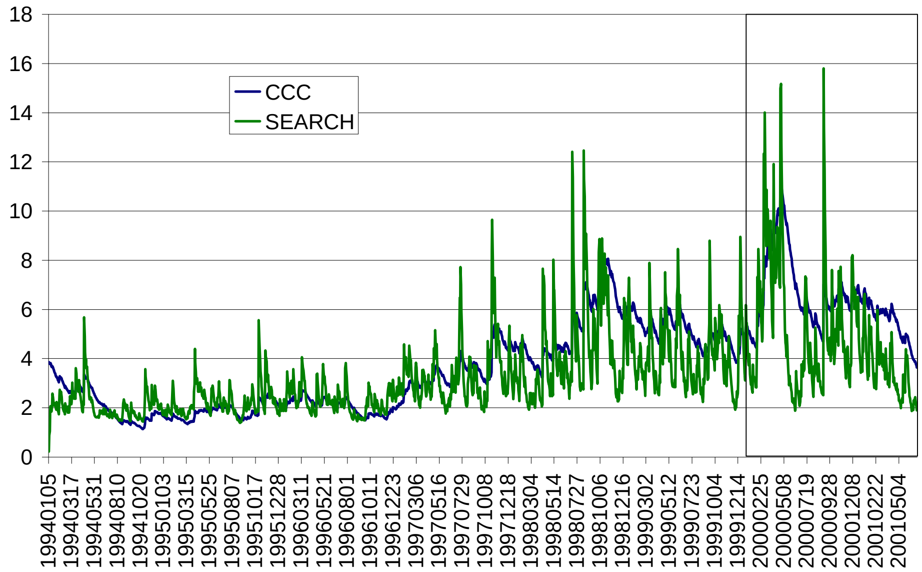

Figure 1: Estimates and forecasts of element $(1, 1)$ of Σ_t for the CCC and the SEARCH models.

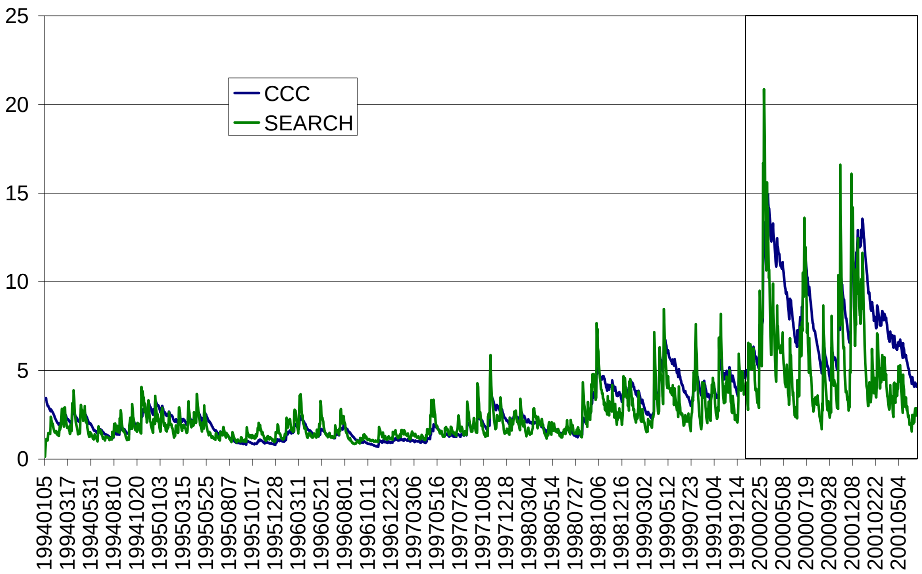

Figure 2: Estimates and forecasts of element $(2, 2)$ of Σ_t for the CCC and the SEARCH models.

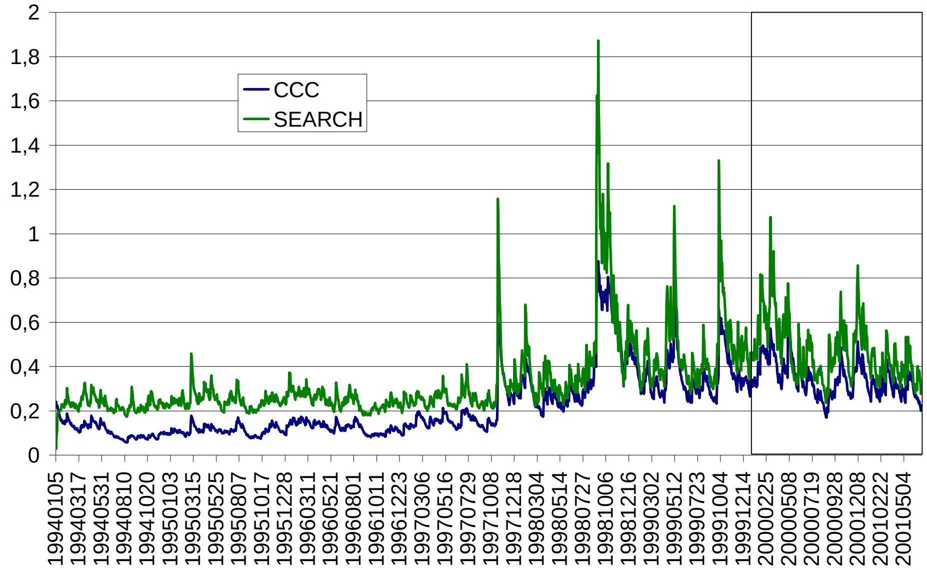

Figure 3: Estimates and forecasts of element $(1, 5)$ of Σ_t for the CCC and the SEARCH models.

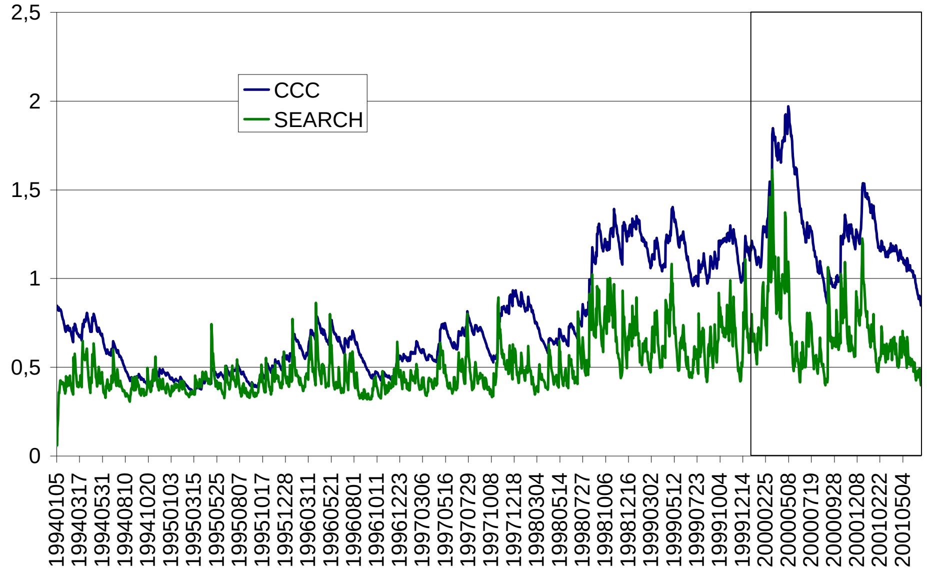

Figure 4: Estimates and forecasts of element $(1, 10)$ of Σ_t for the CCC and the SEARCH models.

Model	log-Lik	number of parameters in Σ_t
SEARCH	-31879.321	$\underset{S}{4}\ \underset{c}{20}\ \underset{E}{8}\ \underset{A}{8}\ \underset{R}{16} = 56$
CCC	-31258.465	$\underset{GARCH}{60}\ \underset{R}{190} = 250$
DCC	-31257.704	$\underset{GARCH}{60}\ \underset{R}{190}\ \underset{DCC}{2} = 252$
OGARCH	-31974.045	$\underset{GARCH}{40}\ \underset{FL}{210} = 250$

Table 5: Log-likelihood values

implies 64 equality restrictions; the test statistics results was equal to 179.565 with a p-value less than 10^{-5}. Also this test points to a HET specification.

A LM test against an heterogeneous S specification was computed; this gave a $\chi^2(5)$ test statistics of LM=112.766, with a p-value less than 10^{-5}. A similar test against an heterogeneous S specification with diagonal EA specification was computed; this gave a $\chi^2(69)$ test statistics of LM=270.031, with a p-value less equal to 10^{-5}.

These specification tests call for HET generalizations of the spatial specification reported in Tables 4 and **??** to diagonal and heterogeneous specifications. These specifications were not estimated due to time constraints and will be added in a future version of this paper.

In order to compare the SEARCH model to alternative multivariate GARCH specifications we estimated also a Full Factor OGARCH model and a DCC model, in addition to the CCC. The DCC provided estimated values very close results to the CCC model which was preferred by a LM test. Table 5 reports the maximized log-likelihoods of the various models; note that in general models are non-nested. One can note that the SEARCH model, with a maximize log-likelihood comparable with the OG and inferior to the models in the CC class, has much fewer parameters.

The literature on the diagnostic checking of multivariate GARCH models includes statistics based on the standardised residuals, as in Ling and Li (1997), Tse and Tsui (1999), Wong and Li (2002), Duchesne and Lancette (2003), and Duchesne and Roy (2004). Different approaches include a residaual based analysis, Tse (2002) and a multivariate Portmanteau statistics, Hosking (1980). Apart the last one, the various approache suffer form loss of power or inconsistency against alternative models.

In order to compare the estimated models, SEARCH, CC and OGARCH we use simple Ljung-Box tests on the univariate standardised residuals and the multivariate Portmanteau of Hosking (1980). These statistics are not reported for brevity but we focus instead on the worst cases and on the rejections of the LB null hypothesis of absence of autocorrelations in squared standardized residuals.

For each model we computed the LB p-values for each single asset for lag lenght from 1 to 20. We then averaged the p-values over the 20 lags and selected the asset which presented the lowest average p-value; this was asset NEM for all models. Results are in Fig. 5. One observes that the OG provides the worst results while

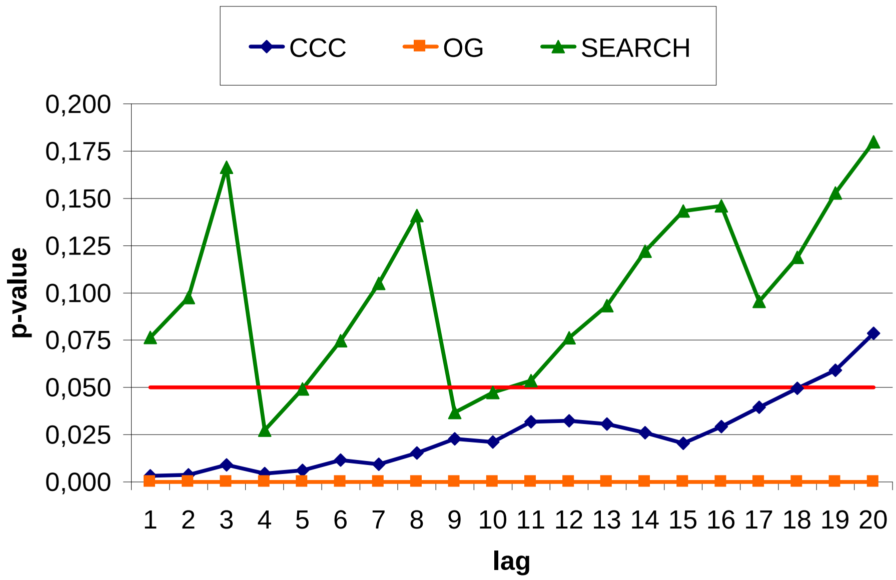

Figure 5: p-values of the univariate LB autocorrelation tests performed on squared standardized residuals of asset NEM for the CCC, OGARCH and SEARCH models. The asset NEM was selected as the ones with minimal p-values across lags 1 to 20 for each of the 3 models.

SEARCH dominates the CCC.

The same exercise was perfomed selecting the asset with minimal p-value across lags 1 to 20. Results are presented in Fig. 6. OG still provides the worst results while SEARCH and CCC have similar performances, with SEARCH doing marginally better. The same results may be summarized by calculating the number of rejections of the univariate LB tests for each model; these are reported in Fig. 7. One observes that the OG provides the worst results with SEARCH doing marginally better than CCC.

Furthermore, we computed the multivariate Portmanteau test which rejected the null hypothesis for all models. The "omnibus" multivariate normality test of Doornik and Hansen (1994) was also computed on standardized residuals. Also this test rejected the null of normality for all models. This indicates that the none of the estimated models is able to capture the multivariate volatility dependence completely.

The models may also be compared in terms of value-at-risk measures, in order to asses their forecast ability. In this case the attention is on the out-of-sample estimated variances. We consider an equally weighted portfolio and computed its returns and its variances with the various fitted models. Then, we used a back-testing approach computing exceptions, the tests of Kupiec (1995) and Christoffersen (1998), and the ones based on the specification of loss functions.

Consider the former. Table 6 reports the number of exceptions and the tests of Unconditional Coverage (UC), Independence (I) and Conditional Coverage (CC). The exceptions are above the VaR confidence level in all cases except for the OG

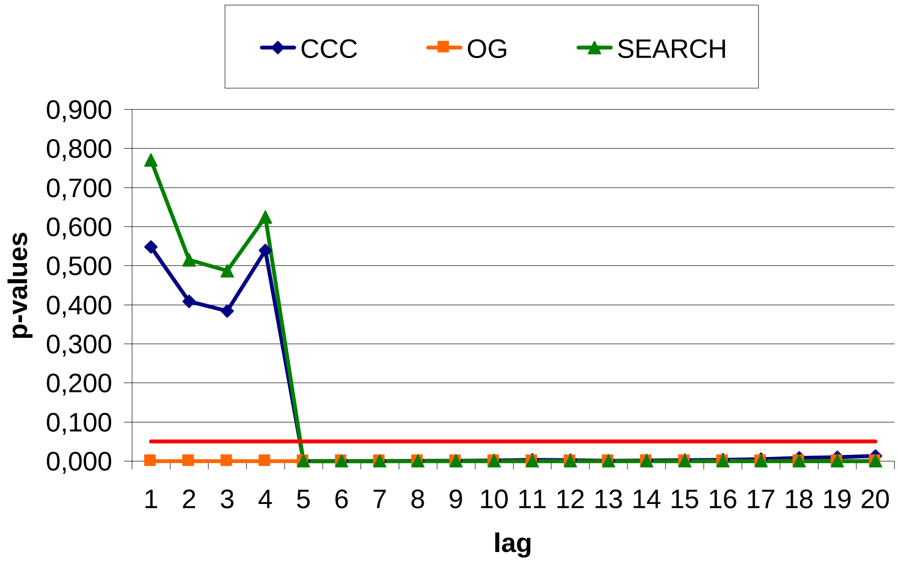

Figure 6: p-values of the univariate LB autocorrelation tests performed on standardized residuals for the CCC, OG and SEARCH models. The selected asset was asset HOV for the CCC and the SEARCH, and asset NEM for the OG. The assets were selected as the ones with minimal p-value across lags 1 to 20.

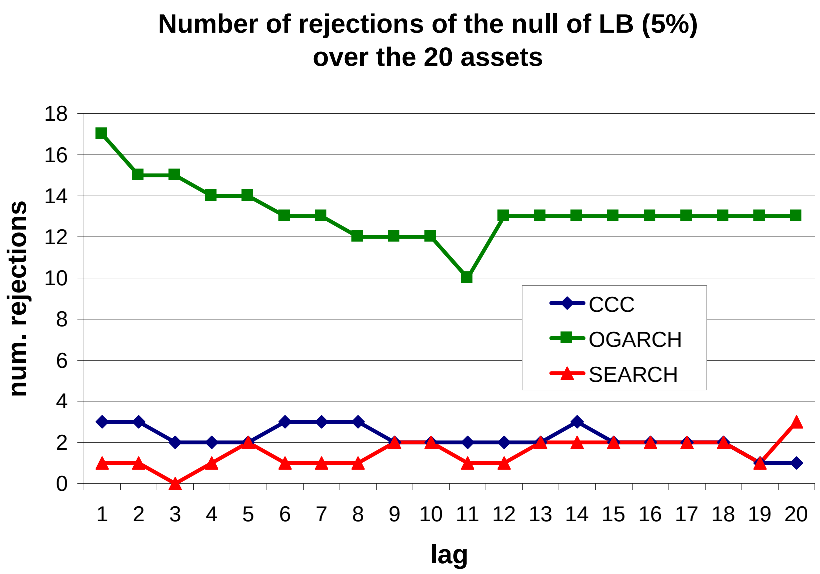

Figure 7: Number of rejections (out of 20) of the univariate LB autocorrelation tests performed on standardized residuals for the CCC, OGARCH and SEARCH models, by lag.

	CCC	DCC	OG	SEARCH
Exceptions (VaR significance level)				
Num. (1%)	5	5	0	8
Fraction	0.013	0.013	0.000	0.021
Num. (5%)	28	27	3	34
Fraction	0.074	0.071	0.008	0.090
Tests (VaR significance level)				
UC (1%)	0.368	0.368	7.578	3.626
UC (5%)	4.095	3.290	*21.364*	*10.460*
I (1%)	3.559	3.559	0.000	1.858
I (5%)	1.536	1.868	5.576	2.579
CC (1%)	3.926	3.926	7.578	5.484
CC (5%)	5.630	5.158	*26.940*	*13.039*

Table 6: Conditional and unconditional coverage tests; significant values in italics.

model, which provides the highest forecasted variances, and, as a consequence the widest VaR bands. Rejection of the null hypothesis is reported in italics. UC, I and CC tests are asymptotically distributed as a χ^2 with one degree of freedom for the UC and I test, while with 2 degrees of freedom for the CC test. There is some evidence of dependence among exceptions for the SEARCH and OG model for the 5% VaR.

We next considered the evaluation based on loss function. The following loss functions were calculated, where v_t indicates the value-at-risk and u_t the univariate prediction error:

$$LE_{j,t} := \begin{cases} LF_{j,t} & u_t < v_t \\ 0 & u_t \geq v_t \end{cases},$$

$$LF_{0,t} := 1 + (u_t - v_t)^2, \qquad LF_{1,t} := \left|1 - \left|\frac{u_t}{v_t}\right|\right|,$$

$$LF_{2,t} := \frac{(|u_t| - v_t)^2}{|v_t|}, \qquad LF_{3,t} := |u_t - v_t|.$$

$LE_{0,t}$ was suggested by Lopez (1999), while Caporin (2003) proposed $LF_{i,t}$, $LE_{i,t}$, $i = 1, 2, 3$. Note that the $LF_{i,t}$ functions weight all observations, while $LE_{i,t}$ assign positive loss only to exceptions to the value-at-risk measure. Table 7 reports the summed losses over the forecasting period. The minimal value for the various losses is obtained either by the DCC for by the SEARCH.

The loss functions signal a preference for the DCC if computed only on the exceptions, i.e. using $LE_{i,t}$, if one excludes the OG given its very wide value-at-risk bands. A different picture emerges, however, when using the loss functions $LF_{i,t}$, as designed in Caporin (2003). The purpose of the $LF_{i,t}$ functions is to mesure how far the value-at-risk is from the realised return series, monitoring the opportunity cost of the value-at-risk. For these loss functions, there is a clear preference for the SEARCH, implying that the SEARCH-based value-at-risk is closer to the returns series. This ensures a smaller value-at-risk band without excessively increasing the

Loss Function and VaR level	CCC	DCC	OG	SEARCH
LE_0 1%	7.314	*7.004*	0.000	11.307
LE_0 5%	40.126	*38.117*	4.088	48.960
LE_1 1%	1.163	*1.033*	0.000	1.581
LF_1 1%	234.179	236.495	290.668	*227.796*
LE_1 5%	7.811	*7.060*	0.732	10.116
LF_1 5%	206.026	207.155	258.786	*202.771*
LE_2 1%	0.901	*0.760*	0.000	1.383
LF_2 1%	426.454	441.743	1099.251	*386.386*
LE_2 5%	6.859	*6.150*	0.474	9.037
LF_2 5%	253.060	260.201	652.482	*234.025*
LE_3 1%	2.894	*2.635*	0.000	3.672
LF_3 1%	960.741	980.700	1735.604	*909.088*
LE_3 5%	13.605	*12.570*	1.729	16.522
LF_3 5%	707.050	719.457	1235.260	*675.261*

Table 7: Summed loss functions over forecasting period; minimal values are reported in italics.

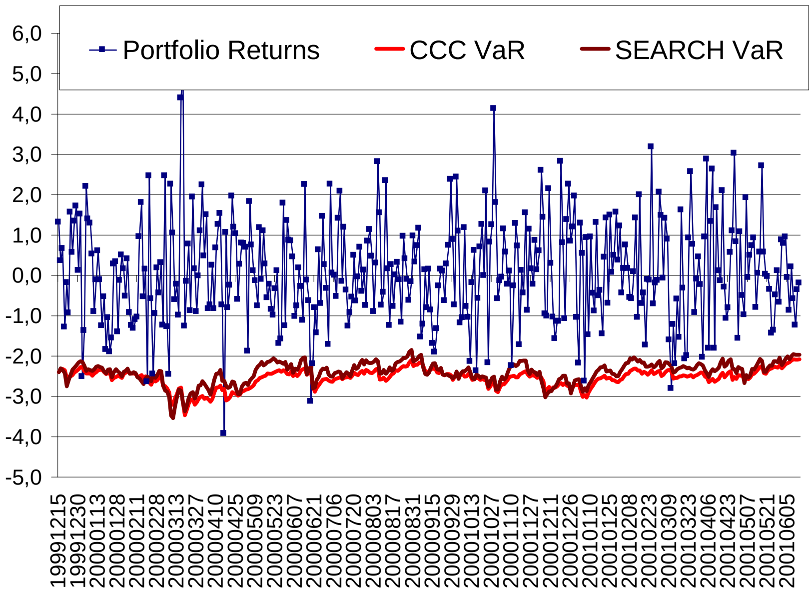

Figure 8: Returns on an equally weighted portfolio and the 1% value-at-risk for the CCC and SEARCH models

value-at-risk exceptions. Fig. 8 reports the portfolio returns and the CCC and SEARCH 1% VaR bands.

These results could be further extended allowing for changes in the portfolio composition in order to extend the model comparison to the portfolio weights evolution.

6 Conclusions

In this paper we have presented a spatial multivariate ARCH specification, called SEARCH, which employs spatially restricted parameter matrices to obtain a model dimension that is linear in the number of assets n.

An application to daily returns on 20 stocks from the NYSE for the period January 1994 to June 2001 shows the benefits of the present specification.

References

Anselin, L. 1988. Spatial Econometrics: Methods and Models, Dorddrecht: Kluwer Academic Publishers

Alexander C. 2001. Orthogonal GARCH. In Mastering Risk. Financial Times–Prentice Hall: London; 2: 21–38.

Amisano G., C. Giannini, 1997. Topics in structural VAR econometrics, 2nd revised and enlarged edition, Berlin: Springer Verlag.

Baltagi B. H., S. H. Song, B. C. Jung, W. Koh, 2004, Testing for Serial Correlation, Spatial Autocorrelation and Random Effects Using Panel Data, Working paper.

Bauwens L., S. Laurent and J.K.V. Rombouts, 2003, Multivariate GARCH models: a survey, CORE discussion paper 2003/31.

Bollerslev T. 1990. Modelling the coherence in short-run nominal exchange rates: A multivariate generalized ARCH approach. Review of Economics and Statistics 72: 498–505.

Bollerslev T., J. Wooldridge 1992. Quasi-Maximum Likelihood Estimation and Inference in Dynamic Models with Time-Varying Covariances. Econometric Reviews 11, 143-172.

Bollerslev T, Engle RF. 1993. Common persistence in conditional variances. Econometrica 61: 167– 186.

Caporin M., 2003, Evaluating Value-at-Risk measures in presence of long memory conditional volatility, mimeo, University of Venice 'Ca' Foscari'.

Christoffersen P.F., 1998, Evaluating interval forecasts, International Economic Review, 39, 841-862.

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

Cressie, N.A.C. 1993 Statistics for spatial data, revised edition. John Wiley & Sons. New York.

Doornik J. A. and H. Hansen 1994. An Omnibus Test for Univariate and Multivariate Normality. Paper available on J. Doornik's homepage.

Duchesne, P. and S. Lancette, 2003, On testing for multivariate ARCH effects in vector time series models, The Canadian Journal of Statistics, 31-(3), 275-292.

Duchesne, P. and R. Roy, 2004, On consistent testing for serial correlation of unknown form in vector time series models, Journal of Multivariate Analysis, 89, 148-180.

Engle R. 2002. Dynamic conditional correlation - a simple class of multivariate GARCH models. Journal of Business and Economic Statistics 20: 339–350.

Engle R. F. and S. Kozicki, 1993, Testing for Common Features, Journal of Business and Economic Statistics 11, 369 - 380.

Engle R, Kroner K. 1995. Multivariate simultaneous generalized ARCH. Econometric Theory 11: 122–150.

Engle RF, Ng VK, Rothschild M. 1990. Asset pricing with a factor-ARCH covariance structure: Empirical estimates for treasury bills. Journal of Econometrics 45: 213–238.

Giacomini R., C. W.J. Granger 2004. Aggregation of space-time processes. Journal of Econometrics 118: 7–26

Hansen P. R. and A. Lunde 2004. Consistent Ranking of Volatility Models. Working paper, Brown University.

Hosking, J.R.M., 1980, The multivariate portmanteau statistic, Journal of the American Statistical Association, 75-(371), 602-608.

Kupiec P., 1995, "Techniques for verifying the accuracy of value-at-risk measurement models", Journal of Derivatives, 3, 73-84.

Ling, S. and McAleer, M. 2003. Asymptotic theory for a new vector ARMA-GARCH model. Econometric Theory 19: 280-310.

Ling, S. and W.K. Li, 1997, Diagnostic checking on nonlinear multivariate time series with multivariate ARCH errors, Journal of TIme Series Analysis, 18-(5), 447-464.

Lopez A. J., 1999, Methods for evaluating value-at-risk estimates, Federal Reserve Bank of San Francisco Economic Review, 2.

Lucchetti, R. 2004. Identification of covariance structures. Working paper, University of Ancona, Italy.

Magnus J. 1988. Linear structures, Oxford University Press.

Magnus J. and H. Neudecker, 1999. Matrix differential calculus with applications in statistics and econometrics. New York: Wiley.

Pesaran M. H., T. Schuermann and Weiner S. 2004. Modelling Regional Interdependencies Using A Global Error-Correcting Macroeconometric Model. Journal of Business Economics and Statistics, 22, 129-162.

Sentana E. and G. Fiorentini 2001. Identification, estimation and testing of conditionally heteroskedastic factor models, Journal of Econometrics 102, 143-164.

Tse, Y.K., 2002, Residual based diagnostics for conditional heteroscedasticity models, Econometrics Journal, 5, 358-373.

Tse, Y.K. and A.K.C. Tsui, 1999, A note on diagnosing multivariate conditional heteroscedasticity models, Journal of Time Series Analysis, 20-(6), 679-691.

Van Der Weide R. 2002. GO-GARCH: a multivariate generalised orthogonal GARCH model, Journal of Applied Econometrics, 17, 549-564

Vrontos I.D., Dellaportas P, Politis D.N. 2003. A full-factor multivariate GARCH model. Econometrics Journal 6: 311–333.

Wong, H., and W.K. Li, 2002, Detecting and diagnostic checking multivariate conditional heteroscedastic time series models, Annals of the Institute of Statistical Mathematics, 54-(1), 45-59.

A Appendix A: Factor Arch representation

In this appendix we motivate the restriction of a multivariate GARCH specification implied by requiring the ARCH component to be common, see e.g. Engle and Bollerslev (1993) and reference therein for the definition of common features in volatility. We here take the feature to be time-varying conditional variances.

Consider the GARCH process

$$h_t^* = c^* + A^*(L)e_{t-1}^* + E^*(L)h_{t-1}^* \tag{31}$$

where h_t^* and e_t^* are $n^* \times 1$. Here all variables and matrices are starred, in order to distinguish them from the ones in the main text. We first give basic definition.

The GARCH process (31) is called *non-trivial* if $A^*(L)$ is not identically zero. We say that the GARCH process is driven by a subset $\xi^{*\prime}e_{t-1}^*$ of e_{t-1}^* if

$$A^*(L) = A^\dagger(L)\xi^{*\prime}$$

for a $n^* \times r^*$ polynomial matrix $A^\dagger(L)$ with $r^* < n^*$. We say that h_t^* has GARCH common features, GARCH–CF, if it is non-trivial and if there exists a non-zero matrix $n^* \times k^*$ matrix γ^* such that $\gamma^{*\prime}h_t^*$ is constant, for an appropriate choice of h_0^*.

Proposition 3 *The following are sufficient conditions for a non-trivial GARCH process* (31) *to be driven by a subset $\xi^{*\prime}e_{t-1}^*$ of e_{t-1}^*, which itself follows a GARCH process, and to present common features:*

$$(\gamma^* : \xi^*) \text{ is a full rank square } n^* \times n^* \text{ matrix} \tag{32}$$

$$(\gamma^* : \xi^*)' \, E^*(L) = \begin{pmatrix} E_1(L)\gamma^{*\prime} \\ E_2(L)\xi^{*\prime} \end{pmatrix} \tag{33}$$

$$A^*(L) = A^\dagger(L)\xi^{*\prime} \tag{34}$$

$$\gamma^{*\prime}A^\dagger(L) = 0 \tag{35}$$

with $E_1(L)$ a $k^* \times k^*$ and $E_2(L)$ a $r^* \times r^*$ matrix polynomials, where $E_1(L)$ is stable.

Note that (33) can be described as the conditions that γ^* and ξ^* are matrices of left eigenvectors of $E^*(L)$.

A leading example is the GARCH(1,1) process

$$h_t^* = c^* + A^*e_{t-1}^* + E^*h_{t-1}^* \tag{36}$$

for which the above conditions (33) (34) (35) translate into the following conditions:

$$(\gamma^* : \xi^*)' \, E^* = \begin{pmatrix} E_1\gamma^{*\prime} \\ E_2\xi^{*\prime} \end{pmatrix}$$

$$A^* = A^\dagger\xi^{*\prime}$$

$$\gamma^{*\prime}A^\dagger = 0$$

which can be imposed in a restricted parametrization of the GARCH(1,1) process as specified by the following proposition.

Proposition 4 *Under the conditions* (32) (33), (34), (35), *the GARCH*$(1,1)$ *process* (36) *admits the following factorial representation*

$$h_t^* = \xi_\perp^* q^* + \mathring{\gamma}_\perp h_{2,t}^*, \tag{37}$$

$$h_{2,t}^* = c_2^* + a e_{2,t-1}^* + E_2^* h_{2,t-1}^* \tag{38}$$

where $c_2^* := \xi^{*\prime} c^*$, $\alpha := \xi^{*\prime} A^\dagger$, $q^* := (\gamma^{*\prime}\xi_\perp^*)^{-1}(I - E_1^*)^{-1}\gamma^{*\prime}c^*$, $\mathring{\gamma}_\perp := \gamma_\perp^*(\xi^{*\prime}\gamma_\perp^*)^{-1}$, $h_{2,t}^* := \xi^{*\prime} h_{t-1}^*$, $e_{2,t}^* := \xi^{*\prime} e_{t-1}^*$. *The GARCH*$(1,1)$ *process for* $h_{2,t}^*$, $e_{2,t}^*$ *is stable (i.e. covariance stationary) iff the eigenvalues of* $\alpha + E_2^*$ *are less than* 1 *in modulus.*

The decomposition in (37) separates the time-invariant part $\xi_\perp^* q^*$ of the conditional variance and the time-varying part given in (38). We also observe that the unconditional variance, under the given condition for stability of the process for $h_{2,t}^*$, $e_{2,t}^*$, is

$$\mathbb{E}(h_t^*) = \xi_\perp^* q^* + \mathring{\gamma}_\perp (I - \alpha - E_2^*)^{-1} \xi^{*\prime} c^*,$$

with a contribution from both components.

Proof. Consider the projection identity

$$I_{n^*} = \xi_\perp^*(\gamma^{*\prime}\xi_\perp^*)^{-1}\gamma^{*\prime} + \gamma_\perp^*(\xi^{*\prime}\gamma_\perp^*)^{-1}\xi^{*\prime}$$

which holds thanks to (32), and pre-multiply h_t^* by its r.h.s.. Let $h_{1t}^* := \gamma^{*\prime} h_t^*$ and $h_{2t}^* := \xi^{*\prime} h_t^*$ so that the following decomposition applies:

$$h_t^* = \left(\xi_\perp^*(\gamma^{*\prime}\xi_\perp^*)^{-1} : \gamma_\perp^*(\xi^{*\prime}\gamma_\perp^*)^{-1}\right) \begin{pmatrix} h_{1t}^* \\ h_{2t}^* \end{pmatrix}.$$

We wish to show that h_{1t}^* can be made constant by an appropriate choice of h_0^* and that h_{2t}^* is driven by a GARCH process for $\xi^{*\prime} e_t$. We first consider h_{1t}^*.

$$h_{1t}^* = \gamma^{*\prime}\left(c^* + E^* h_{t-1}^*\right) = \gamma^{*\prime} c^* + E_1^* \gamma^{*\prime} h_{t-1}^* =: c_1^* + E_1^* h_{1t-1}^*$$

where $c_1^* := \gamma^{*\prime} c^*$ and the eigenvalues of E_1^* are all less than one in modulus by the requirement of stability of $E_1^*(L) = E_1^*$. Solving backwards one obtains

$$h_{1t}^* = \left(\sum_{i=0}^{t-1} E_1^{*i}\right) c_1^* + E_1^{*t} h_{10}^*$$

By choosing $h_{10}^* = \left(\sum_{i=0}^{\infty} E_1^{*i}\right) c_1^* = (I - E_1^*)^{-1} c_1^*$ one obtains

$$h_{1t}^* = \left(\sum_{i=0}^{t-1} E_1^{*i}\right) c_1^* + E_1^{*t}\left(\sum_{i=0}^{\infty} E_1^{*i}\right) c_1^* = \left(\sum_{i=0}^{\infty} E_1^{*i}\right) c_1^* = h_{10}^* = (I - E_1^*)^{-1} c_1^*,$$

which is constant. Next consider h_{2t}^*

$$\begin{aligned} h_{2t}^* &= \xi^{*\prime}\left(c^* + A^* e_{t-1}^* + E^* h_{t-1}^*\right) = \xi^{*\prime} c^* + \xi^{*\prime} A^\dagger \xi^{*\prime} e_{t-1}^* + E_2^* \xi^{*\prime} h_{t-1}^* = \\ &=: c_2^* + \alpha \xi^{*\prime} e_{t-1}^* + E_2^* h_{2t-1}^* \end{aligned}$$

where $\alpha := \xi^{*\prime} A^{\dagger}$. This defines a GARCH(1,1) for $\xi^{*\prime} e^*_{t-1}$, where $\mathbb{E}_{t-1}\left(\xi^{*\prime} e^*_t\right) = h^*_{2t} = \xi^{*\prime}\mathbb{E}_{t-1}\left(e^*_t\right)$, where $h^*_t := \mathbb{E}_{t-1}\left(e^*_t\right)$. Hence

$$h^*_t = \xi^*_{\perp} q^* + \overset{\circ}{\gamma}_{\perp} h^*_{2,t},$$
$$h^*_{2,t} = c^*_2 + \alpha e^*_{2,t-1} + E^*_2 h^*_{2,t-1}$$

where $q^* := \left(\gamma^{*\prime}\xi^*_{\perp}\right)^{-1}\left(I - E^*_1\right)^{-1}\gamma^{*\prime} c^*$, $\overset{\circ}{\gamma}_{\perp} := \gamma^*_{\perp}(\xi^{*\prime}\gamma^*_{\perp})^{-1}$, $e^*_{2,t} := \xi^{*\prime} e^*_{t-1}$, $h^*_{2,t} := \xi^{*\prime} h^*_{t-1}$. The conditions for covariance stationarity of $e^*_{2,t} := \xi^{*\prime} e^*_{t-1}$, $h^*_{2,t} := \xi^{*\prime} h^*_{t-1}$ are the usual ones. ■

We conclude this Appendix by showing how one can aggregate (23) at the system level. Let $q := (q'_{11} : \ldots : q'_{\ell k})'$, $h_{2,t} := (h'_{112,t} : \ldots : h'_{\ell k 2,t})'$, $\omega := (\omega_{11} : \ldots : \omega_{\ell k})'$, $\alpha := (\alpha_{11} : \ldots : \alpha_{\ell k})'$, $\beta := (\beta_{11} : \ldots : \beta_{\ell k})'$, $b := \operatorname{diag}(b_{ij})$, $N := \operatorname{diag}\left(\iota_{n_{ij}}\right)$, $\overline{N} := \operatorname{diag}\left(\frac{1}{n_{ij}}\iota_{n_{ij}}\right)$ $\alpha^* := \operatorname{diag}(\alpha)$, $\beta^* := \operatorname{diag}(\beta)$. Note also that as $h_{2,t} = N' h_t$. Collecting various blocks i, j, one finds[5]

$$h_t = bq + \overline{N} h_{2,t},$$
$$h_{2,t} = \omega + \alpha^* N' e_{t-1} + \beta^* h_{2,t-1}.$$

Substituting the second equation into the first one gets

$$h_t = \left(b : \overline{N}\right)\begin{pmatrix} q \\ \omega \end{pmatrix} + \left(\overline{N}\beta^* N'\right) h_{t-1} + \left(\overline{N}\alpha^* N'\right) e_{t-1}$$

which is of the form $h_t = c + E h_{t-1} + A e_{t-1}$ for

$$c = \left(b : \overline{N}\right) c^*, \qquad A = \overline{N}\alpha^* N' \qquad E = \overline{N}\beta^* N', \tag{39}$$

with $c^* := (q' : \omega')'$.

B Appendix B: Derivatives of the likelihood function

In this Appendix we give details on the calculation of derivatives for the most general diagonal SEARCH model. When $g(x) := (g_1(x) : \ldots g_n(x))'$ is a $n \times 1$ differentiable function of the $m \times 1$ vector $x := (x_1 : \ldots : x_m)'$, we indicate by $\partial g(x)/\partial x'$ the $n \times m$ matrix with lq element given by $\partial g_l(x)/\partial x_q$. When ε_t is assumed conditionally Gaussian, the log-likelihood function is $\ln L(\theta) = \sum_{t=1}^{T} \ln f_t(\theta)$, where $\ln f_t(\theta) := \ln f\left(y_t | \mathcal{I}_{t-1}, \theta\right)$ is given by

$$\ln f_t(\theta) = -\frac{1}{2}\left(\ln\det\left(\Sigma_t\right) + \operatorname{tr}\left(\Sigma_t^{-1}\varepsilon_t\varepsilon_t'\right)\right) = -\frac{1}{2}\left(f_{1t} + f_{2t}\right), \text{ say} \tag{40}$$

with

$$f_{1t} := -2\ln\det\left(I - S\right) + \ln\det\left(R\right), \qquad f_{2t} := 2\ln\det\left(D_t\right) + \operatorname{tr}\left(\Sigma_t^{-1}\varepsilon_t\varepsilon_t'\right)$$

[5] One possible generalization of the present factor structure could be to assume that the matrices a^* and ζ^* are not necessarily diagonal.

where

$$\text{tr}\left(\Sigma_t^{-1}\varepsilon_t\varepsilon_t'\right) = \text{tr}\left(\Gamma_t^{-1}\left(I-S\right)\varepsilon_t\varepsilon_t'(I-S')\right) = \text{tr}\left(R^{-1}D_t^{-1}\eta_t\eta_t'D_t^{-1}\right) = \text{tr}\left(R^{-1}\psi_t\psi_t'\right).$$

We find derivatives of $\ln f_t(\theta)$ for the diagonal S, E, A, parameterizations, which nest the scalar ones, and for the general R specification (25). The scalar S, E, A, specifications are obtained as restrictions of the corresponding diagonal specifications. In fact let θ_u indicate the $m \times 1$ subvector of parameters contained on the diagonal of a $m \times m$ matrix; examples of θ_u are θ_{s_i}, $\theta_{\beta_{ij,q}}$, $\theta_{\alpha_{ij,q}}$. The scalar specifications are obtained by constraining θ_u to $\theta_u = \iota_m\varphi_u$ where ι_m is the $m \times 1$ vector of ones and φ_u is a scalar parameter, so that $\mathrm{d}\theta_u = \iota_m \mathrm{d}\varphi_u$. Hence $\partial \ln f_t/\partial\varphi_u = (\partial \ln f_t/\partial\theta_u)\,\iota_m$, i.e. the derivatives wrt φ_u are obtained by summing the corresponding derivatives wrt θ_u.

Similarly the case $\rho_{ij,q} = \rho_{ij}^q$ is obtained as a restriction of (25). One finds $\mathrm{d}\rho_{ij,q} = q\rho_{ij}^{q-1}\mathrm{d}\rho_{ij}$ and hence the derivatives wrt ρ_{ij} are weighted derivatives wrt $\rho_{ij,q}$, $\partial \ln f_t(\theta)/\partial\rho_{ij} = \sum_{q=1}^{n_{ij}-1} q\rho_{ij}^{q-1}\ \partial \ln f_t(\theta)/\partial\rho_{ij,q}$. For the spatial EA specification, we use the definition $W_{ij,q}^* = U_{ij}^q + U_{ij}^{q\prime}$, see (19). The derivatives for the diagonal EA specification with $W_{ij,q}^* = U_{ij}^q$ in (18) is analogous, except that the first elements in $\theta_{\beta_{ij,q}}$ is constrained to 0, i.e. $\theta_{\beta_{ij,q}} = (0 : I_{n_{ij}-1})'\varphi_u$, where φ_u is a $n_{ij}-1$ vector of free parameters, so that $\partial \ln f_t/\partial\varphi_u = \left(\partial \ln f_t/\partial\theta_{\beta_{ij,q}}\right)(0 : I_{n_{ij}-1})'$.

B.1 Parametrization

The model we consider is given by the diagonal SEARCH specification with

$$S := \sum_{i=1}^{q} s_i W_i, \qquad c := \sum_{i=1}^{k}\sum_{j=1}^{\ell} \mathcal{M}_{ij}c_{ij}, \qquad E := \sum_{i=1}^{k}\sum_{j=1}^{\ell}\sum_{q=0}^{m_{E_{ij}}} \mathcal{M}_{ij}\beta_{ij,q}W_{ij,q}^*\mathcal{M}_{ij}',$$

$$A := \sum_{i=1}^{k}\sum_{j=1}^{\ell}\sum_{q=0}^{m_{A_{ij}}} \mathcal{M}_{ij}\alpha_{ij,q}W_{ij,q}^*\mathcal{M}_{ij}', \qquad R := \sum_{i=1}^{k}\sum_{j=1}^{\ell}\sum_{q=0}^{m_{R_{ij}}} \mathcal{M}_{ij}\rho_{ij,q}W_{ij,q}^*\mathcal{M}_{ij}'$$

where $\mathcal{M}_{ij} := (0 : I_{n_{ij}} : 0)'$ is a $n \times n_{ij}$ matrix with all zero entries except for the block corresponding to $\mathcal{C}_{ij}$, with corresponding block entries equal to $I_{n_{ij}}$.

Let also $\mathcal{H}$ denote the elimination matrix that satisfies $\mathcal{H}'\text{vec}(D) = \text{vecd}(D)$ or $\text{vec}(D) = \mathcal{H}\,\text{vecd}(D)$, with D a $n \times n$ diagonal matrix, see Magnus (1988), section 7.3 p. 109. Similarly let $\mathcal{H}_{ij}$ be the same type of matrix when D is $n_{ij} \times n_{ij}$. Let moreover $\mathcal{R}_{ij} := \mathcal{M}_{ij} \otimes \mathcal{M}_{ij}$.

The parameters θ are grouped into subvectors θ_l, where $l \in \mathcal{A} := \{\mu, S, c, E, A, R\}$. The typical parameter subvector of θ_S is indicated as θ_{s_i}, where $s_i := \text{diag}(\theta_{s_i})$; $\theta_{c_{ij}}$ is a typical element in θ_c, where $c_{ij} := \mathcal{F}_{ij}\theta_{c_{ij}}$; $\theta_{\beta_{ij,q}}$ is a typical subvector of θ_E, where $\beta_{ij,q} := \text{diag}\left(\theta_{\beta_{ij,q}}\right)$, $\theta_{\alpha_{ij,q}}$ is a subvector of θ_A, with $\alpha_{ij,q} := \text{diag}\left(\theta_{\alpha_{ij,q}}\right)$ and $\rho_{ij,q}$ is a scalar element of θ_R.

Consider derivatives of $\text{vec}\mu'$, $\text{vec}S$, c, $\text{vec}E$, $\text{vec}A$, $\text{vec}R$ with respect to the typical parameters θ_μ, θ_{s_i}, $\theta_{c_{ij}}$, $\theta_{\beta_{ij,q}}$, $\theta_{\alpha_{ij,q}}$, $\rho_{ij,q}$. One immediately finds from definitions above that

$$
\begin{aligned}
\frac{\partial \mathrm{vec}\left(\mu'\right)}{\partial \theta'_{\mu}} &= I_n, \qquad \frac{\partial \mathrm{vec}\left(S\right)}{\partial \theta'_{s_i}} = \left(W'_i \otimes I_n\right)\mathcal{H}, \\
\frac{\partial c}{\partial \theta'_{c_{ij}}} &= \mathcal{M}_{ij}\frac{\partial c_{ij}}{\partial \theta'_{c_{ij}}}, \qquad \frac{\partial \mathrm{vec}\left(E\right)}{\partial \theta'_{\beta_{ij,q}}} = \mathcal{R}_{ij}\frac{\partial \mathrm{vec}\left(E_{ij}\right)}{\partial \theta'_{\beta_{ij,q}}}, \\
\frac{\partial \mathrm{vec}\left(A\right)}{\partial \theta'_{\alpha_{ij,q}}} &= \mathcal{R}_{ij}\frac{\partial \mathrm{vec}\left(A_{ij}\right)}{\partial \theta'_{\alpha_{ij,q}}}, \qquad \frac{\partial \mathrm{vec}\left(R\right)}{\partial \rho_{ij,q}} = \mathcal{R}_{ij}\frac{\partial \mathrm{vec}\left(R_{ij}\right)}{\partial \rho_{ij,q}}.
\end{aligned} \tag{41}
$$

where

$$
\begin{aligned}
\frac{\partial c_{ij}}{\partial \theta'_{c_{ij}}} &= \mathcal{F}_{ij}, \qquad \frac{\partial \mathrm{vec}\left(E_{ij}\right)}{\partial \theta'_{\beta_{ij,q}}} = \left(W^{*\prime}_{ij,q} \otimes I_{n_{ij}}\right)\mathcal{H}_{ij}, \\
\frac{\partial \mathrm{vec}\left(A_{ij}\right)}{\partial \theta'_{\alpha_{ij,q}}} &= \left(W^{*\prime}_{ij,q} \otimes I_{n_{ij}}\right)\mathcal{H}_{ij}, \qquad \frac{\partial \mathrm{vec}\left(R_{ij}\right)}{\partial \rho_{ij,q}} = \mathrm{vec}\left(W^{*}_{ij,q}\right).
\end{aligned} \tag{42}
$$

Moreover we wish to calculate $\partial \mathrm{vec}\left(\Sigma_t\right)/\partial \theta'_q$ for a generic $q \in \mathcal{A}$. Define

$$
\mathrm{vec}\left(\Sigma_t\right) = \mathrm{vec}\left(\left(I - S\right)^{-1} D_t R D_t \left(I - S'\right)^{-1}\right) =: g^*\left(\varphi, h_t\right)
$$

where φ are the parameters that enter g^*_t for fixed h_t, see the section on first derivatives. We first calculate $\mathrm{d}g^*_t\left(\varphi, \mathrm{d}h_t\right)$ as follows

$$
\begin{aligned}
\mathrm{d}g^*_t\left(\varphi, \mathrm{d}h_t\right) &= \frac{1}{2}\mathrm{vec}\left(\left(I - S\right)^{-1}\left(\mathrm{diag}\left(\mathrm{d}h_t\right) D_t^{-2}\Gamma_t + \Gamma_t D_t^{-2}\,\mathrm{diag}\left(\mathrm{d}h_t\right)\right)\left(I - S'\right)^{-1}\right) \\
&= \frac{1}{2}\left(\left(\left(I - S\right)^{-1} \otimes \left(I - S\right)^{-1}\right)\left(\Gamma_t D_t^{-2} \otimes I + I \otimes \Gamma_t D_t^{-2}\right)\right)\mathcal{H}\mathrm{d}h_t \\
&=: H_{1t}\mathrm{d}h_t, \text{ say.}
\end{aligned}
$$

We next calculate $\mathrm{d}g^*\left(\mathrm{d}\varphi, h_t\right)$ as follows

$$
\begin{aligned}
\mathrm{d}g^*_t\left(\mathrm{d}\varphi, h_t\right) &= \mathrm{vec}(\left(I - S\right)^{-1}\mathrm{d}S\Sigma_t + \Sigma_t \mathrm{d}S'\left(I - S\right)^{-1} + \left(I - S\right)^{-1} D_t \mathrm{d}R D_t \left(I - S'\right)^{-1}) \\
&= \left(\left(\left(I - S\right)^{-1} \otimes \Sigma_t\right)\mathcal{K}_{nn} + \left(\Sigma_t \otimes \left(I - S\right)^{-1}\right)\right)\mathrm{d}\left(\mathrm{vec}\left(S\right)\right) - \\
&\left(\left(I - S\right)^{-1} D_t \otimes \left(I - S\right)^{-1} D_t\right)\mathrm{d}\left(\mathrm{vec}\left(R\right)\right) \\
&= H_{2t}\mathrm{d}\left(\mathrm{vec}\left(S\right)\right) + H_{3t}\mathrm{d}\left(\mathrm{vec}\left(R\right)\right), \text{ say.}
\end{aligned}
$$

Finally, collecting terms and considering a generic parameter θ_q one finds

$$
\frac{\partial \mathrm{vec}\left(\Sigma_t\right)}{\partial \theta'_q} = H_{1t}\frac{\partial h_t}{\partial \theta'_q} + H_{2t}\frac{\partial \mathrm{vec}\left(S\right)}{\partial \theta'_q} + H_{3t}\frac{\partial \mathrm{vec}\left(R\right)}{\partial \theta'_q}, \tag{43}
$$

where H_{1t}, H_{2t}, H_{3t} are defined above.

B.2 First order derivatives

In (40) $f_{1t} := f_{1t}\left(\zeta_1\right)$ is the time-invariant part of $\ln f_t\left(\theta\right)$, where ζ_1 collects the parameters θ_S, θ_R that appear in f_{1t}. Conversely f_{2t} is the time-varying part of

$\ln f_t(\theta)$; observe that f_{2t} depends directly on some parameters ζ_2 and also on h_t, which is itself a function of θ (except for θ_R); we write $f_{2t} := f_{2t}(\zeta_2, h_t)$.

We next introduce notation for differentials, which follows Magnus and Neudecker (1999). The differential of f_{2t} is for instance indicated as $\mathrm{d}f_{2t}(\zeta_2, h_t; \mathrm{d}\zeta_2, \mathrm{d}h_t)$; this by definition can be decomposed into

$$\mathrm{d}f_{2t}(\zeta_2, h_t; \mathrm{d}\zeta_2, \mathrm{d}h_t) = \mathrm{d}f_{2t}(\zeta_2, h_t; \mathrm{d}\zeta_2, 0) + \mathrm{d}f_{2t}(\zeta_2, h_t; 0, \mathrm{d}h_t).$$

Obviously this decomposition applies to all partitions of the parameter vector θ. In the following we use the shorthands $\mathrm{d}f_{2t}(\mathrm{d}\zeta_2, h_t)$ or $\mathrm{d}f_{2t}(\mathrm{d}\zeta_2)$ for $\mathrm{d}f_{2t}(\zeta_2, h_t; \mathrm{d}\zeta_2, 0)$ and $\mathrm{d}f_{2t}(\zeta_2, \mathrm{d}h_t)$ or $\mathrm{d}f_{2t}(\mathrm{d}h_t)$ for $\mathrm{d}f_{2t}(\zeta_2, h_t; 0, \mathrm{d}h_t)$. We use similar shorthands for other functions and subvectors of parameters, like in $\mathrm{d}f_{1t}(\mathrm{d}\mu)$. We find

$$\begin{aligned}
\mathrm{d}\ln f_t(\theta) &= -\frac{1}{2}\left(\mathrm{d}f_{1t}(\mathrm{d}\zeta_1) + \mathrm{d}f_{2t}(\zeta_2, \mathrm{d}h_t) + \mathrm{d}f_{2t}(\mathrm{d}\zeta_2, h_t)\right) = \\
&= \sum_{l\in\mathcal{A}} g'_{l,t}\mathrm{d}\theta_i + \operatorname{tr}\left(B_t \operatorname{diag}(\mathrm{d}h_t)\right) \text{ with} \\
\mathrm{d}f_{2t}(\zeta_2, \mathrm{d}h_t) &= \operatorname{tr}\left(\left(I - \psi_t\psi_t' R^{-1}\right) D_t^{-2}\operatorname{diag}(\mathrm{d}h_t)\right) =: -2\operatorname{tr}\left(B_t \operatorname{diag}(\mathrm{d}h_t)\right) \qquad (44)
\end{aligned}$$

and $\sum_{l\in\mathcal{A}} g'_{l,t}\mathrm{d}\theta_i := -2^{-1}\left(\mathrm{d}f_{1t}(\mathrm{d}\zeta_1) + \mathrm{d}f_{2t}(\mathrm{d}\zeta_2, h_t)\right)$, $B_t := -2^{-1}\left(I - \psi_t\psi_t' R^{-1}\right) D_t^{-2}$. In (44) we have used the definition $h_t := \operatorname{vecd}(D_t^2)$, which implies $\mathrm{d}D_t^2 = \operatorname{diag}(\mathrm{d}h_t)$, and

$$\mathrm{d}D_t = \frac{1}{2}D_t^{-1}\mathrm{d}D_t^2 = \frac{1}{2}D_t^{-1}\operatorname{diag}(\mathrm{d}h_t)$$

which follows from the $\mathrm{d}D_t^2 = 2D_t\mathrm{d}D_t$ (thanks to diagonality of D_t).

In Subsection B.2.1 we describe the contributions $g_{l,t}$ to the various derivatives that appear in the differential above; in Subsection B.2.2 we calculate the differential of h_t, which is the used in Subsection B.2.3 to derive recursions for the first order derivatives.

B.2.1 Contributions $g_{l,t}$

We first consider $\theta_\mu := \operatorname{vec}(\mu')$; one finds $\mathrm{d}f_{1t}(\mathrm{d}\mu) = 0$, $\mathrm{d}f_{2t}(\mathrm{d}\mu, h_t) = -2\operatorname{tr}\left(\Sigma_t^{-1}\varepsilon_t w_t'\mathrm{d}\mu'\right)$ and hence

$$g'_{\mu,t}\mathrm{d}\theta_\mu = \operatorname{tr}\left(\Sigma_t^{-1}\varepsilon_t w_t'\mathrm{d}\mu'\right) = \operatorname{vec}\left(w_t\varepsilon_t'\Sigma_t^{-1}\right)'\mathrm{d}\theta_\mu. \qquad (45)$$

Consider next θ_{s_i} where $s_i := \operatorname{diag}(\theta_{s_i})$ appears in $S := \sum_{i=1}^{m_S} s_i W_i$. Hence $\mathrm{d}S = \operatorname{diag}(\mathrm{d}\theta_{s_i}) W_i$ and one finds $\mathrm{d}f_{1t}(\mathrm{d}S) = 2\operatorname{tr}\left(\mathrm{d}S\,(I - S)^{-1}\right)$, $\mathrm{d}f_{2t}(\mathrm{d}S, h_t) = -2\operatorname{tr}\left(\mathrm{d}S\varepsilon_t\varepsilon_t'\Sigma_t^{-1}(I - S)^{-1}\right)$ and hence

$$\begin{aligned}
g'_{s_i,t}\mathrm{d}\theta_{s_q} &= \operatorname{tr}\left(\operatorname{diag}(\mathrm{d}\theta_{s_i}) W_i \left(\left(\varepsilon_t\varepsilon_t'\Sigma_t^{-1} - I\right)(I - S)^{-1}\right)\right) \qquad (46) \\
&= \operatorname{vecd}\left(W_i\left(\left(\varepsilon_t\varepsilon_t'\Sigma_t^{-1} - I\right)(I - S)^{-1}\right)\right)'\mathrm{d}\theta_{s_i}
\end{aligned}$$

Consider next $\rho_{ij,q}$; $\mathrm{d}f_{1t}(\mathrm{d}R) = \operatorname{tr}(R^{-1}\mathrm{d}R)$, $\mathrm{d}f_{2t}(\mathrm{d}R, h_t) = -\operatorname{tr}(R^{-1}\psi_t\psi_t'R^{-1}\mathrm{d}R)$ and hence

$$\begin{aligned}
g_{\rho_{ij,q}}\mathrm{d}\rho_{ij,q} &= -\frac{1}{2}\operatorname{tr}\left(\left(R^{-1} - R^{-1}\psi_t\psi_t'R^{-1}\right)\mathcal{M}_{ij}W^*_{ij,q}\mathcal{M}'_{ij}\mathrm{d}\rho_{ij,q}\right) = \\
&= \frac{1}{2}\left(\psi'_{ij,t}R_{ij}^{-1}W^*_{ij,q}R_{ij}^{-1}\psi_{ij,t} - \operatorname{tr}\left(R_{ij}^{-1}W^*_{ij,q}\right)\right)\mathrm{d}\rho_{ij,q} \qquad (47)
\end{aligned}$$

where also here we have used the block-diagonal structure of R. In (47) the term $\mathrm{d}h_t$ does not appear because h_t is not a function of R.

We note that $g'_{E,t}\mathrm{d}\theta_E = g'_{A,t}\mathrm{d}\theta_A = g'_{c,t}\mathrm{d}\theta_c = 0$, because $\ln f_t$ depends on θ_E, θ_A, θ_c only through h_t.

B.2.2 Contributions from h_t

Consider the differential of h_t with respect to μ, S, c_{ij}, E_{ij}, A_{ij}

$$\begin{aligned} \mathrm{d}h_t &= E\mathrm{d}h_{t-1} - 2A\,\mathrm{vecd}\left(\eta_{t-1}w'_{t-1}\mathrm{d}\mu'\left(I-S'\right)\right), & (48)\\ \mathrm{d}h_t &= E\mathrm{d}h_{t-1} - 2A\,\mathrm{vecd}\left(\mathrm{d}S\varepsilon_{t-1}\eta'_{t-1}\right) & (49)\\ \mathrm{d}h_{ij,t} &= E_{ij}\mathrm{d}h_{ij,t-1} + \mathrm{d}c_{ij}, & (50)\\ \mathrm{d}h_{ij,t} &= E_{ij}\mathrm{d}h_{ij,t-1} + \mathrm{d}E_{ij}h_{ij,t-1}, & (51)\\ \mathrm{d}h_{ij,t} &= E_{ij}\mathrm{d}h_{ij,t-1} + \mathrm{d}A_{ij}e_{ij,t-1}, & (52) \end{aligned}$$

where we have used the block-diagonal structure of E and A.

We next introduce further notation, in order to write derivatives in a compact form. Recall that the differential (44) is of the form $\mathrm{d}\ln f_t = \mathrm{tr}\left(B_t\,\mathrm{diag}\left(\mathrm{d}h_t\right)\right) + \sum_{l\in\mathcal{A}} g'_{l,t}\mathrm{d}\theta_i$. Consider a subvector θ_l, $l\in\mathcal{A}$, and apply the chain rule one obtains the $1\times v_l$ vector

$$\frac{\partial \ln f_t}{\partial \theta'_l} = g'_{l,t} + \mathrm{vec}(B'_t)'\mathcal{H}\frac{\partial h_t}{\partial \theta'_l} =: g'_{l,t} + \mathrm{vec}(B'_t)'\mathcal{H}K_{l,t}, \tag{53}$$

where $K_{l,t} := \partial h_t/\partial\theta'_l$ is $n\times v_l$. Note that $K_{l,t} := \partial h_t/\partial\theta'_l$ may depend on a single block $\mathcal{C}_{ij}$ when θ_l belongs to the EA or R specifications. Hence we let

$$K_{l,t} =: \mathcal{N}_l K^*_{l,t}$$

where $\mathcal{N}_l$ has dimensions $n\times n_{k,l}$ and $K^*_{l,t}$ is of dimension $n_{k,l}\times v_l$. For $l=\mu$, S one has $n_{k,l}=n$, $\mathcal{N}_l = I_n$, $K^*_{l,t} = \partial h_t/\partial\theta'_l$. For $l\in\{c, E, A, R\}$ one has $n_{k,l}=n_{ij}$, $\mathcal{N}_l = \mathcal{M}_{ij}$, $K^*_{l,t} = \partial h_{ij,t}/\partial\theta'_l$. We further note that eq. (48) to (52) imply that $K^*_{l,t}$ satisfies recursions

$$K^*_{l,t} = F_l K^*_{l,t-1} + P_{l,t} \tag{54}$$

where $P_{l,t}$ is of the same dimensions of $K^*_{l,t}$, i.e. $n_{k,l}\times v_l$. We here derive F_l, $P_{l,t}$ for the various parameters θ_l. Note that $F_{\rho_{ij,q}} = P_{\rho_{ij,q},t} = 0$.

B.2.3 Terms F_l and $P_{l,t}$

For θ_μ, from (48) one has $\mathcal{N}_\mu = I_n$, $F_\mu = E$ and

$$P_{\mu,t} = -2A\mathcal{H}'\left(\left(I-S\right)\otimes\eta_{t-1}w'_{t-1}\right). \tag{55}$$

For θ_{s_i} from (49) one finds $\mathcal{N}_{s_i} = I_n$, $F_{s_i} = E$ and

$$P_{s_i,t} = -2A\left(I_n\odot\eta_{t-1}\varepsilon'_{t-1}W_i\right) = -2A\,\mathrm{dg}\left(\eta_{t-1}\varepsilon'_{t-1}W_i\right)$$

where the property $\mathcal{H}'\left(A\otimes B\right)\mathcal{H} = A\odot B$ has been used, see e.g. Magnus (1988) Theorem 7.7 page 113.

θ_i	$g_{i,t}$	F_i	$P_{i,t}$
θ_μ	$\text{vec}(w_{t-1}\varepsilon'_{t-1}\Sigma_t^{-1})$	E	$-2A\mathcal{H}'\left((I-S)\otimes\eta_{t-1}w'_{t-1}\right)$
θ_{s_i}	$\text{vecd}\left(W_i\left(\left(\varepsilon_t\varepsilon'_t\Sigma_t^{-1}-I\right)(I-S)^{-1}\right)\right)$	E	$-2A\,\text{dg}\left(\eta_{t-1}\varepsilon'_{t-1}W_i\right)$
$\theta_{c_{ij}}$	0	E_{ij}	$\mathcal{F}_{ij}$
$\theta_{\beta_{ij,q}}$	0	E_{ij}	$\text{dg}\left(W^*_{ij,q}h_{ij,t-1}\right)$
$\theta_{A_{ij,q}}$	0	E_{ij}	$\text{dg}\left(W^*_{ij,q}e_{ij,t-1}\right)$
$\rho_{ij,q}$	$\frac{1}{2}\left(\psi'_{ij,t}R_{ij}^{-1}W_q^*R_{ij}^{-1}\psi_{ij,t}-\text{tr}\left(R_{ij}^{-1}W_q^*\right)\right)$	0	0

Table 8: First derivatives in the format (53) for a diagonal specification.

For the EA parameters, from (50), (51), (52) one finds that $\mathcal{N}_{\beta_{ij,q}} = \mathcal{N}_{\alpha_{ij,q}} = \mathcal{N}_{c_{ij}} := \mathcal{M}_{ij}$, $F_{\beta_{ij,q}} = F_{\alpha_{ij,q}} = F_{c_{ij}} = E_{ij}$. The elements $P_{l,t}$ are seen to be

$$P_{\beta_{ij,q},t} = \left(h'_{ij,t-1}\otimes I_{n_{ij}}\right)\frac{\partial\text{vec}\left(E_{ij}\right)}{\partial\theta'_{\beta_{ij,q}}},$$
$$P_{\alpha_{ij,q},t} = \left(e'_{ij,t-1}\otimes I_{n_{ij}}\right)\frac{\partial\text{vec}\left(A_{ij}\right)}{\partial\theta'_{\alpha_{ij,q}}},\qquad P_{c_{ij},t} = \frac{\partial c_{ij}}{\partial\theta'_{c_{ij}}},\tag{56}$$

where the partial derivatives are given in (42). Collecting terms and using the property of $\mathcal{H}_{ij}$ given e.g. in Magnus (1988) Theorem 7.7. (vi) page 113, one finds the results summarized in Table 8, where we report $g_{l,t}$, F_l and $P_{l,t}$ for various subvectors of parameters.

B.3 Second order derivatives

In this subsection we consider second order derivatives. Applying standard properties of the vec operator to (53) one sees that

$$\frac{\partial\ln f_t(\theta)}{\partial\theta_i\partial\theta'_j} = \left(\text{vec}\left(B'_t\right)'\mathcal{H}\mathcal{N}_i\otimes I_{v_i}\right)\mathcal{K}_{n_{k,i},v_i}V_{ij,t} + K'_{i,t}\mathcal{H}'\frac{\partial\text{vec}\left(B'_t\right)}{\partial\theta'_j} + \frac{\partial g_{i,t}}{\partial\theta'_j}.\tag{57}$$

where $V_{ij,t} := \partial k_{i,t}/\partial\theta'_j$, $k_{i,t} := \text{vec}(K^*_{i,t})$. Given recursions (54), see Table 8, one finds for $p_{it} := \text{vec}\left(P_{i,t}\right)$ that $k_{i,t} = \left(I_{v_i}\otimes F_i\right)k_{i,t}+p_{i,t} = \left(K^{*\prime}_{i,t-1}\otimes I_{n_{k,i}}\right)\text{vec}\left(F_i\right)+p_{i,t}$, from which

$$\text{d}k_{i,t} = \left(I_{v_i}\otimes F_i\right)\text{d}k_{i,t} + \left(K^{*\prime}_{i,t-1}\otimes I_{n_{k,i}}\right)\text{vec}\left(\text{d}F_i\right) + \text{d}p_{i,t}$$

and thus

$$V_{ij,t} := \left(I_{v_i}\otimes F_i\right)V_{ij,t-1} + \left(K^{*\prime}_{i,t-1}\otimes I_{n_{k,i}}\right)\frac{\partial\text{vec}\left(F_i\right)}{\partial\theta'_j} + \frac{\partial p_{i,t}}{\partial\theta'_j}.$$

We note that F_l equals E, E_{ij} or 0, see Table 8. Hence all partial derivatives $\partial\text{vec}\left(F_i\right)/\partial\theta'_j = 0$ for $j\in\{\mu,\, S,\, c,\, A,\, R\}$. For $\beta_{ij,q}$, the expression $\partial\text{vec}\left(E_{ij}\right)/\partial\theta'_{\beta_{ij,q}}$ is given in (42).

In order to calculate (57) we hence need to calculate $\partial\text{vec}\left(B'_t\right)/\partial\theta'_j$, $\partial p_{i,t}/\partial\theta'_j$, and $\partial g_{i,t}/\partial\theta'_j$ for all pairs θ_i, θ_j, where $i,\, j\in\mathcal{A}$.

B.3.1 Derivatives of B_t

Consider the differential $B_t' := 2^{-1} D_t^{-2} \left(R^{-1} \psi_t \psi_t' - I \right)$

$$
\begin{aligned}
\operatorname{vec}\left(\mathrm{d}B_t'\right) = & \frac{1}{2} \operatorname{vec}(D_t^{-2} R^{-1} \left(\mathrm{d}\psi_t \psi_t' + \psi_t \mathrm{d}\psi_t' - \mathrm{d}R R^{-1} \psi_t \psi_t' \right) - \\
& - \left(R^{-1} \psi_t \psi_t' - I \right) D_t^{-2} \operatorname{diag}\left(\mathrm{d}h_t\right) D_t^{-2}) \\
& = C_{1t} \mathrm{d}\psi_t + C_{2t} \mathrm{d}\left(\operatorname{vec} R\right) + C_{3t} \mathrm{d}h_t, \text{ say,}
\end{aligned}
$$

where

$$
\begin{aligned}
C_{1t} &:= 2^{-1} \left(\left(\psi_t \otimes D_t^{-2} R^{-1} \right) + \left(D_t^{-2} R^{-1} \psi_t \otimes I_n \right) \right) \\
C_{2t} &:= -2^{-1} \left(\psi_t \psi_t' R^{-1} \otimes D_t^{-2} R^{-1} \right), \\
C_{3t} &:= 2^{-1} \left(D_t^{-2} \otimes \left(I - R^{-1} \psi_t \psi_t' \right) D_t^{-2} \right) \mathcal{H}.
\end{aligned}
$$

Recall that $\psi_t := D_t^{-1} \left(I - S \right) \varepsilon_t = D_t^{-1} \eta_t$ so that

$$
\begin{aligned}
\mathrm{d}\psi_t &= -D_t^{-1} \mathrm{d}D_t \psi_t - D_t^{-1} \mathrm{d}S \varepsilon_t - D_t^{-1} \left(I - S \right) \mathrm{d}\mu w_{t-1} = \\
&= -\frac{1}{2} \operatorname{diag}\left(\mathrm{d}h_t\right) D_t^{-2} \psi_t - D_t^{-1} \mathrm{d}S \varepsilon_t - D_t^{-1} \left(I - S \right) \mathrm{d}\mu w_{t-1} = \qquad (58) \\
&= -\frac{1}{2} \operatorname{diag}\left(D_t^{-2} \psi_t \right) \mathrm{d}h_t - \left(\varepsilon_t' \otimes D_t^{-1} \right) \mathrm{d}\left(\operatorname{vec} S\right) - \left(w_{t-1}' \otimes D_t^{-1} \left(I - S \right) \right) \mathcal{K}_{n_w, n} \mathrm{d}\left(\operatorname{vec} \mu'\right) \\
&= C_{4t} \mathrm{d}h_t + C_{5t} \mathrm{d}\left(\operatorname{vec} S\right) + C_{6t} \mathrm{d}\left(\operatorname{vec} \mu'\right), \text{ say.}
\end{aligned}
$$

where $\operatorname{diag}\left(D_t^{-2} \psi_t \right) \mathrm{d}h_t = \left(\psi_t' \otimes D_t^{-2} \right) \mathcal{H} \mathrm{d}h_t$. Substituting

$$
\operatorname{vec}\left(\mathrm{d}B_t'\right) = \left(C_{1t} C_{4t} + C_{3t} \right) \mathrm{d}h_t + C_{1t} C_{5t} \mathrm{d}\left(\operatorname{vec} S\right) + C_{1t} C_{6t} \mathrm{d}\left(\operatorname{vec} \mu'\right) + C_{2t} \mathrm{d}\left(\operatorname{vec} R\right)
$$

Hence, naming $C_{*t} := C_{1t} C_{4t} + C_{3t}$, one finds

$$
\begin{aligned}
&\frac{\partial \operatorname{vec}\left(B_t'\right)}{\partial \theta_\mu'} = C_{*t} \frac{\partial h_t}{\partial \theta_\mu'} + C_{1t} C_{6t}, \qquad \frac{\partial \operatorname{vec}\left(B_t'\right)}{\partial \theta_{s_i}'} = C_{*t} \frac{\partial h_t}{\partial \theta_{s_i}'} + C_{1t} C_{5t} \frac{\partial \operatorname{vec}\left(S\right)}{\partial \theta_{s_i}'} \\
&\frac{\partial \operatorname{vec}\left(B_t'\right)}{\partial \theta_l'} = C_{*t} \frac{\partial h_t}{\partial \theta_l'}, \qquad l \in \left\{ \alpha_{ij,q}, \beta_{ij,q}, c_{ij} \right\} \\
&\frac{\partial \operatorname{vec}\left(B_t'\right)}{\partial \rho_{ij,q}} = C_{2t} \frac{\partial \operatorname{vec}\left(R\right)}{\partial \rho_{ij,q}}
\end{aligned}
$$

where $\partial h_t / \partial \theta_i'$ appears in Subsections B.2.2 and B.2.3 and basic derivatives are given in (41) (42).

B.3.2 Derivatives of $p_{l,t}$

Consider $P_{\mu,t} = -2 A \mathcal{H}' \left(\left(I - S \right) \otimes \eta_{t-1} w_{t-1}' \right)$, with differential

$$
\begin{aligned}
\mathrm{d}P_{\mu,t} = & -2 \mathrm{d}A \mathcal{H}' \left(\left(I - S \right) \otimes \eta_{t-1} w_{t-1}' \right) + 2 A \mathcal{H}' \left(\mathrm{d}S \otimes \eta_{t-1} w_{t-1}' \right) + \\
& + 2 A \mathcal{H}' \left(\left(I - S \right) \otimes \mathrm{d}S \varepsilon_{t-1} w_{t-1}' \right) - 2 A \mathcal{H}' \left(\left(I - S \right) \otimes \left(I - S \right) \mathrm{d}\mu w_{t-1} w_{t-1}' \right)
\end{aligned}
$$

Vectorizing, using Theorem 10 in Magnus and Neudecker (1988), p. 47 or Chapter 13.14, we find

$$\frac{\partial p_{\mu,t}}{\partial \theta'_{\alpha_{ij,q}}} = C_{7t}\frac{\partial \mathrm{vec}\,(A)}{\partial \theta'_{\alpha_{ij,q}}}, \qquad \frac{\partial p_{\mu,t}}{\partial \theta'_{s_i}} = C_{8t}\frac{\partial \mathrm{vec}\,(S)}{\partial \theta'_{s_i}}, \qquad \frac{\partial p_{\mu,t}}{\partial \theta'_{\mu}} = C_{9t}$$

where

$$\begin{aligned}
C_{7t} &:= -2\left(\left((I-S')\otimes w_{t-1}\eta'_{t-1}\right)\mathcal{H}\otimes I_n\right),\\
C_{8t} &:= \left(I_{v_\mu}\otimes 2A\mathcal{H}'\right)\left(I_n\otimes \mathcal{K}_{n_w n}\otimes I_n\right)\left(\left(I_{n_2}\otimes \mathrm{vec}\left(w_{t-1}\eta'_{t-1}\right)\right)+\right.\\
&\quad \left.+\left(\mathrm{vec}\,(I-S)\otimes I_{nn_w}\right)\right)\left(w_{t-1}\varepsilon'_{t-1}\otimes I_n\right),\\
C_{9t} &:= -\left(I_{v_\mu}\otimes 2A\mathcal{H}'\right)\left(I_n\otimes \mathcal{K}_{n_w n}\otimes I_n\right)\left(\mathrm{vec}\,(I-S)\otimes I_{nn_w}\right)\left(w_{t-1}w'_{t-1}\otimes (I-S)\right)\mathcal{K}_{n_w n},
\end{aligned}$$

see also (41) (42). All other derivatives are zero.

Consider next $P_{s_i,t} = -2A\,\mathrm{dg}\left(\eta_{t-1}\varepsilon'_{t-1}W_i\right)$, with differential

$$\begin{aligned}
\mathrm{d}P_{s_i,t} &= 2A\,\mathrm{dg}\left(\mathrm{d}S\varepsilon_{t-1}\varepsilon'_{t-1}W_i\right) + 2A\,\mathrm{dg}\left(W'_i\varepsilon_{t-1}w'_{t-1}\mathrm{d}\mu'\,(I-S') + \eta_{t-1}w'_{t-1}\mathrm{d}\mu' W_i\right) -\\
&\quad - 2\mathrm{d}A\,\mathrm{dg}\left(\eta_{t-1}\varepsilon'_{t-1}W_i\right).
\end{aligned}$$

Using the property $\mathcal{H}\mathcal{H}'\mathrm{vec}(A) = \mathrm{vec}\,(\mathrm{dg}\,(A))$, see Theorem 7.3 (ii) in Magnus (1998) p. 110 one finds

$$\frac{\partial p_{s_i,t}}{\partial \theta'_{\alpha_{ij,q}}} = C_{12t}\frac{\partial \mathrm{vec}\,(A)}{\partial \theta'_{\alpha_{ij,q}}}, \qquad \frac{\partial p_{s_i,t}}{\partial \theta'_{s_j}} = C_{10t}\frac{\partial \mathrm{vec}\,(S)}{\partial \theta'_{s_j}}, \qquad \frac{\partial p_{s_i,t}}{\partial \theta'_{\mu}} = C_{11t},$$

where

$$\begin{aligned}
C_{10t} &:= 2\left(I_n\otimes A\right)\mathcal{H}\mathcal{H}'\left(W_i\varepsilon_{t-1}\varepsilon'_{t-1}\otimes I_n\right),\\
C_{11t} &:= 2\left(I_n\otimes A\right)\mathcal{H}\mathcal{H}'\left(\left((I-S)\otimes W_i\varepsilon_{t-1}w'_{t-1}\right)+\left(W_i\otimes \eta_{t-1}w'_{t-1}\right)\right),\\
C_{12t} &:= 2\left(\mathrm{dg}\left(\eta_{t-1}\varepsilon'_{t-1}W_i\right)\otimes I_n\right),
\end{aligned}$$

and see (41) (42). Next observe that $P_{c,t}$ and $P_{\rho_{ij,q},t}$ do not depend on parameters, so that $\partial p_{i,t}/\partial \theta'_j = 0$ for all $j \in \mathcal{A}$, $i = c$, R.

Consider next $P_{\beta_{ij,q},t} = \mathrm{dg}\left(W^*_{ij,q}h_{ij,t-1}\right)$, with differential $\mathrm{d}P_{\beta_{ij,q},t} = \mathrm{dg}\left(W^*_{ij,q}\mathrm{d}h_{ij,t-1}\right)$. We obtain for a generic parameter subvector θ_l

$$\frac{\partial p_{\beta_{ij,q},t}}{\partial \theta'_l} = \mathcal{H}_{ij}W^*_{ij,q}\frac{\partial h_{ij,t}}{\partial \theta'_l}$$

where the derivatives $\partial h_t/\partial \theta'_l$ appears in Subsections B.2.2 and B.2.3.

Consider next $P_{\alpha_{ij,q},t} = \mathrm{dg}\left(W^*_{ij,q}e_{ij,t-1}\right)$, with differential $\mathrm{d}P_{\alpha_{ij,q},t} = \mathrm{dg}\left(W^*_{ij,q}\mathrm{d}e_{ij,t-1}\right)$ where $\mathrm{d}e_{ij,t-1} = 2\,\mathrm{vecd}\left(\mathrm{d}\eta_{ij,t-1}\eta'_{ij,t-1}\right)$ and

$$\mathrm{d}\eta_{ij,t-1} = \mathcal{M}'_{ij}\mathrm{d}\eta_{t-1} = -\mathcal{M}'_{ij}\left(\mathrm{d}S\varepsilon_{t-1} + (I-S)\,\mathrm{d}\mu w_{t-1}\right),$$

so that

$$\mathrm{d}e_{ij,t-1} = -2\mathcal{H}_{ij}' \left(\eta_{ij,t-1} \otimes I_{n_{ij}}\right) \mathcal{M}_{ij}' \left(\left(\varepsilon_{t-1}' \otimes I_n\right) \mathrm{vec}\left(\mathrm{d}S\right) + \left(\left(I-S\right) \otimes w_{t-1}'\right) \mathrm{vec}\left(\mathrm{d}\mu'\right)\right) \tag{59}$$

We hence obtain

$$\begin{aligned}
\frac{\partial p_{\alpha_{ij},q,t}}{\partial \theta_{s_l}'} &= -2\mathcal{H}_{ij} W_{ij,q}^{*} \mathcal{H}_{ij}' \left(\eta_{ij,t-1} \otimes I_{n_{ij}}\right) \mathcal{M}_{ij}' \left(\varepsilon_{t-1}' \otimes I_n\right) \frac{\partial \mathrm{vec}\left(S\right)}{\partial \theta_{s_l}'} \\
\frac{\partial p_{\alpha_{ij},q,t}}{\partial \theta_{\mu}'} &= -2\mathcal{H}_{ij} W_{ij,q}^{*} \mathcal{H}_{ij}' \left(\eta_{ij,t-1} \otimes I_{n_{ij}}\right) \mathcal{M}_{ij}' \left(\left(I-S\right) \otimes w_{t-1}'\right).
\end{aligned}$$

and all other derivatives are zero.

B.3.3 Derivatives of $g_{l,t}$

In order to calculate derivatives of $g_{l,t}$ we need $\partial \mathrm{vec}\left(\Sigma_t^{-1}\right) / \partial \theta_q'$ for a generic θ_q; one has

$$\begin{aligned}
\frac{\partial \mathrm{vec}\left(\Sigma_t^{-1}\right)}{\partial \theta_q'} &= -\left(\Sigma_t^{-1} \otimes \Sigma_t^{-1}\right) \frac{\partial \mathrm{vec}\left(\Sigma_t\right)}{\partial \theta_q'} \\
&=: G_{1t} \frac{\partial h_t}{\partial \theta_q'} + G_{2t} \frac{\partial \mathrm{vec}\left(S\right)}{\partial \theta_q'} + G_{3t} \frac{\partial \mathrm{vec}\left(R\right)}{\partial \theta_q'},
\end{aligned} \tag{60}$$

where $G_{it} := -\left(\Sigma_t^{-1} \otimes \Sigma_t^{-1}\right) H_{it}$ and H_{it} are defined in (43).

Consider first $g_{\mu,t} = \mathrm{vec}\left(w_{t-1}\varepsilon_{t-1}\Sigma_t^{-1}\right)$, with differential

$$\mathrm{d}g_{\mu,t} = \mathrm{vec}\left(-w_{t-1}w_{t-1}'\mathrm{d}\mu'\Sigma_t^{-1} + w_{t-1}\varepsilon_{t-1}'\mathrm{d}\Sigma_t^{-1}\right).$$

We obtain

$$\begin{aligned}
\frac{\partial g_{\mu,t}}{\partial \theta_q'} &= -\left(\Sigma_t^{-1} \otimes w_{t-1}w_{t-1}'\right) \frac{\partial \mathrm{vec}\left(\mu'\right)}{\partial \theta_q'} + \left(I_n \otimes w_{t-1}\varepsilon_{t-1}'\right) \frac{\partial \mathrm{vec}\left(\Sigma_t^{-1}\right)}{\partial \theta_q'} \\
&= G_{4t} \frac{\partial \mathrm{vec}\left(\mu'\right)}{\partial \theta_q'} + G_{5t} \frac{\partial \mathrm{vec}\left(\Sigma_t^{-1}\right)}{\partial \theta_q'}, \text{ say.}
\end{aligned}$$

Substituting and letting $G_{6t} := G_{5t}G_{1t}$, one finds

$$\begin{aligned}
&\frac{\partial g_{\mu,t}}{\partial \theta_{\mu}'} = G_{6t} \frac{\partial h_t}{\partial \theta_{\mu}'} + G_{4t}, \qquad \frac{\partial g_{\mu,t}}{\partial \theta_{s_i}'} = G_{6t} \frac{\partial h_t}{\partial \theta_{s_i}'} + G_{5t}G_{2t} \frac{\partial \mathrm{vec}\left(S\right)}{\partial \theta_{s_i}'} \\
&\frac{\partial g_{\mu,t}}{\partial \theta_l'} = G_{6t} \frac{\partial h_t}{\partial \theta_l'}, \; l \in \left\{\alpha_{ij,q}, \beta_{ij,q}, c_{ij}\right\} \\
&\frac{\partial g_{\mu,t}}{\partial \rho_{ij,q}} = G_{5t}G_{3t} \frac{\partial \mathrm{vec}\left(R\right)}{\partial \theta_q'}
\end{aligned}$$

Consider now $g_{s_i,t} := \text{vecd}\left(W_i\left(\left(\varepsilon_t\varepsilon_t'\Sigma_t^{-1} - I\right)(I-S)^{-1}\right)\right)$, with differential

$$\begin{aligned}
\mathrm{d}g_{s_i,t} &:= \text{vecd}(W_i(\mathrm{d}\varepsilon_t\varepsilon_t'\Sigma_t^{-1}(I-S)^{-1} + \varepsilon_t\mathrm{d}\varepsilon_t'\Sigma_t^{-1}(I-S)^{-1} + \varepsilon_t\varepsilon_t'\mathrm{d}\Sigma_t^{-1}(I-S)^{-1} + \\
&\quad + \left(\varepsilon_t\varepsilon_t'\Sigma_t^{-1} - I\right)(I-S)^{-1}\mathrm{d}S(I-S)^{-1})) \\
&= \mathcal{H}'\left((I-S')^{-1}\otimes W_i\right)\left(\left(\Sigma_t^{-1}\varepsilon_t w_t'\otimes I\right)\mathcal{K}_{n_w n} + \left(\Sigma_t^{-1}\otimes\varepsilon_t w_t'\right)\right)\mathrm{d}\left(\text{vec}\left(\mu'\right)\right) + \\
&\quad + \mathcal{H}'\left((I-S')^{-1}\otimes W_i\varepsilon_t\varepsilon_t'\right)\mathrm{d}\left(\text{vec}\left(\Sigma_t^{-1}\right)\right) + \\
&\quad + \mathcal{H}'\left((I-S')^{-1}\otimes W_i\left(\varepsilon_t\varepsilon_t'\Sigma_t^{-1} - I\right)(I-S)^{-1}\right)\mathrm{d}\left(\text{vec}\left(S\right)\right) \\
&= G_{7t}\mathrm{d}\left(\text{vec}\left(\mu'\right)\right) + G_{8t}\mathrm{d}\left(\text{vec}\left(\Sigma_t^{-1}\right)\right) + G_{9t}\mathrm{d}\left(\text{vec}\left(S\right)\right), \text{ say.}
\end{aligned}$$

Substituting from (60) one finds, naming $G_{10t} := G_{8t}G_{1t}$,

$$\begin{aligned}
&\frac{\partial g_{s_i,t}}{\partial\theta_\mu'} = G_{10t}\frac{\partial h_t}{\partial\theta_\mu'} + G_{7t}, \qquad \frac{\partial g_{s_i,t}}{\partial\theta_{s_j}'} = G_{10t}\frac{\partial h_t}{\partial\theta_{s_j}'} + \left(G_{8t}G_{2t} + G_{9t}\right)\frac{\partial\text{vec}\left(S\right)}{\partial\theta_{s_j}'} \\
&\frac{\partial g_{s_i,t}}{\partial\theta_l'} = G_{10t}\frac{\partial h_t}{\partial\theta_l'},\ l\in\{\alpha_{ij,q},\beta_{ij,q},c_{ij}\} \\
&\frac{\partial g_{s_i,t}}{\partial\rho_{lj,q}} = G_{8t}G_{3t}\frac{\partial\text{vec}\left(R\right)}{\partial\rho_{lj,q}}.
\end{aligned}$$

Let now $a_{ij} := R_{ij}^{-1}W_{ij,q}^*R_{ij}^{-1}$, and consider $g_{\rho_{ij,q},t} := 2^{-1}\left(\psi_{ij,t}'R_{ij}^{-1}W_q^*R_{ij}^{-1}\psi_{ij,t} - \text{tr}\left(R_{ij}^{-1}W_q^*\right)\right)$, with differential,

$$\begin{aligned}
\mathrm{d}g_{\rho_{ij,q},t} &:= \psi_{ij,t}'a_{ij}\mathrm{d}\psi_{ij,t} + \psi_{ij,t}'R_{ij}^{-1}\mathrm{d}R_{ij}a_{ij}\psi_{ij,t} - \text{tr}\left(\mathrm{d}R_{ij}a_{ij}\right) \\
&= \left(\left(\left(\psi_{ij,t}'a_{ij}\right)\otimes\psi_{ij,t}'R_{ij}^{-1}\right) + \text{vec}\left(a_{ij}\right)'\right)\left(\text{vec}\left(\mathrm{d}R_{ij}\right)\right) + \psi_{ij,t}'a_{ij}\mathrm{d}\psi_{ij,t} \\
&= G_{11t}\mathrm{d}\left(\text{vec}R_{ij}\right) + G_{12t}\mathrm{d}\psi_{ij,t},\ \text{say}
\end{aligned}$$

where from (58) one has

$$\mathrm{d}\psi_{ij,t} = \mathcal{M}_{ij}'\mathrm{d}\psi_t = \mathcal{M}_{ij}'C_{4t}\mathrm{d}h_t + \mathcal{M}_{ij}'C_{5t}\mathrm{d}\left(\text{vec}S\right) + \mathcal{M}_{ij}'C_{6t}\mathrm{d}\left(\text{vec}\mu'\right).$$

Setting $G_{13t} := G_{12t}\mathcal{M}_{ij}'C_{4t}$, substituting one finds

$$\begin{aligned}
&\frac{\partial g_{\rho_{ij,q},t}}{\partial\theta_\mu'} = G_{13t}\frac{\partial h_t}{\partial\theta_\mu'} + G_{12t}\mathcal{M}_{ij}'C_{6t}, \qquad \frac{\partial g_{\rho_{ij,q},t}}{\partial\theta_{s_l}'} = G_{13t}\frac{\partial h_t}{\partial\theta_{s_l}'} + G_{12t}\mathcal{M}_{ij}'C_{5t}\frac{\partial\text{vec}\left(S\right)}{\partial\theta_{s_l}'} \\
&\frac{\partial g_{\rho_{ij,q},t}}{\partial\theta_l'} = G_{13t}\frac{\partial h_t}{\partial\theta_l'}, \qquad l\in\{\alpha_{ij,q},\beta_{ij,q},c_{ij}\} \\
&\frac{\partial g_{\rho_{ij,q},t}}{\partial\rho_{ij,q}} = G_{11t}\frac{\partial\text{vec}\left(R_{ij}\right)}{\partial\rho_{ij,q}}.
\end{aligned}$$

We note that $\partial g_{l,t}/\partial\theta_q'$ is null for all q when l is c, E, A, see Table 8.