University of Heidelberg

Department of Economics

Discussion Paper Series | No. 543

Measuring Persistence in Volatility Spillovers

Christian Conrad and
Enzo Weber

April 2013

Measuring Persistence in Volatility Spillovers*

Christian Conrad† and Enzo Weber‡

†Heidelberg University, Germany

‡IAB Nuremberg and University of Regensburg, Germany

April 11, 2013

Abstract

This paper analyzes volatility spillovers in multivariate GARCH-type models. We show that the cross-effects between the conditional variances determine the persistence of the transmitted volatility innovations. In particular, the effect of a foreign volatility innovation on a conditional variance is even more persistent than the effect of an own innovation unless it is offset by an accompanying negative variance spillover of sufficient size. Moreover, ignoring a negative variance spillover causes a downward bias in the estimate of the initial impact of the foreign volatility innovation. Applying the concept to portfolios of small and large firms, we find that shocks to small firm returns affect the large firm conditional variance once we allow for (negative) spillovers between the conditional variances themselves.

Keywords: Multivariate GARCH, spillover, persistence, small and large firms.

JEL Classification: C32, C51, C52, C53, G10.

---

*Christian Conrad gratefully acknowledges financial support from the "Juniorprofessorenprogramm" of the state of Baden-Württemberg, Germany.

†Address for correspondence: Christian Conrad, Heidelberg University, Faculty of Economics and Social Studies, Bergheimer Strasse 58, 69115 Heidelberg, Germany. E-mail: `christian.conrad@awi.uni-heidelberg.de`, tel: +49 (0)6221 54 3173, fax: +49 (0)6221 54 3578.

‡Address for correspondence: Enzo Weber, University of Regensburg, Department of Economics and Econometrics, 93040 Regensburg, Germany. E-mail: `enzo.weber@wiwi.uni-regensburg.de`, tel: +49 (0)941 943 1952, fax: +49 (0)941 943 4917.

1 Introduction

The investigation of *volatility spillovers* in multivariate GARCH models has recently attracted considerable attention. A specification that is particularly suited for the analysis of volatility spillovers is the extended constant conditional correlation (ECCC) GARCH model proposed in Jeantheau (1998). In this model the conditional variance of one variable can be affected not only by its own lagged squared residuals and conditional variances but also by the lagged squared residuals and conditional variances of the *other* variables in the system. In the following, we will refer to the former as an ARCH spillover and to the latter as a GARCH spillover. For the ECCC GARCH model, He and Teräsvirta (2004) derive the fourth moments and correlation structure of the squared residuals, Nakatani and Teräsvirta (2009) suggest a Lagrange multiplier test for volatility transmission, Woźniak (2012) provides restrictions for second-order noncausality and Francq and Zakoïan (2012) consider quasi-maximum likelihood estimation of an asymmetric version of the model. Nakatani and Teräsvirta (2008) and Conrad and Karanasos (2010) consider the possibility of *negative* GARCH spillovers within the ECCC GARCH framework. In particular, for the N-dimensional model Conrad and Karanasos (2010) derive necessary and sufficient conditions for guaranteeing the positive definiteness of the conditional covariance matrix in the presence of ARCH and potentially negative GARCH spillovers.[1] They term this flexible multivariate specification the *unrestricted* ECCC (UECCC) GARCH model.

From an empirical perspective, in a seminal paper Conrad et al. (1991) analyze the existence of volatility spillovers in equity portfolios of small and large firms. They find 'asymmetric predictability of conditional variances' in the sense that the lagged squared residuals of large firms matter for the conditional variances of small firms, but the reverse effect remains insignificant. That is, the asymmetric predictability refers to the observation of one-directional ARCH spillovers from large to small firms but not vice-versa. Conrad et al. (1991, p.620) provide a potential explanation for this finding by referring to the argument in Ross (1989) that the variance of asset price changes is directly related to the rate of flow of information: "aggregate information first affects large firms and is then impounded with a lag in the prices of small capitalization companies". Although Conrad et al. (1991) employ a two-step estimation strategy, their model can be considered as a version of the UECCC GARCH model which allows for ARCH – but not for GARCH – spillovers. Subsequently, many other studies provided evidence for ARCH spillovers in a variety of financial market applications: e.g. Koutmos and Booth (1995), Laopodis (2003), Chelley-Steeley and Steeley (2005), Wong et al. (2005), Skintzi and Refenes (2006)

[1] The constraints derived in Conrad and Karanasos (2010) are a direct extension of the results for the univariate GARCH model in Tsai and Chan (2008).

and Savva et al. (2009). More recent studies that allow for ARCH as well as GARCH spillovers are McAleer and da Veiga (2008), Chang et al. (2010), Hakim and McAleer (2010), Weber (2010a), Arouri et al. (2011) and Rittler (2012).

At first sight, a somewhat puzzling finding in studies that allow for both ARCH and GARCH spillovers is that the GARCH spillover coefficients often take negative values (see, e.g., Nakatani and Teräsvirta, 2008, or Weber, 2010a). Conrad and Karanasos (2010) interpret this phenomenon simply as a trade-off between volatilities, i.e. an increase in one conditional variance leads to a decrease in another conditional variance. However, we show that there is a more appropriate explanation for the existence of negative GARCH spillovers. Throughout the paper we consider the case of a bivariate UECCC GARCH(1,1) model in order to simplify our arguments. This model has also received the most attention in empirical applications.

We first derive the univariate representation of each conditional variance in terms of its own and the foreign volatility innovation. Following Conrad and Karansos (2006) we then define the impulse response function of an own and foreign volatility innovation as the sequence of coefficients that describe how a shock in period t affects the forecast of the conditional variance in period $t+k$, for $k = 1, 2, \ldots$. In this representation the ARCH spillover coefficients simply measure the initial effect of the own and foreign volatility innovations on the conditional variances. Next, we show that GARCH spillovers can be interpreted as determinants of the persistence of volatility innovations that are transmitted between the variables. Our main result states that in the UECCC GARCH model a negative GARCH spillover is a necessary condition for ensuring that an own volatility innovation has a more persistent effect on a conditional variance than a foreign volatility innovation. More specifically, the necessary and sufficient condition requires that the initial effect of the foreign volatility innovation is subsequently offset by a negative GARCH spillover of at least the same size. Given this result we should consider the finding of negative GARCH spillovers as the rule rather than the exception in most empirical applications. Clearly, neglecting GARCH spillovers imposes a model property that is unintended in most cases.

Our theoretical result has another important implication. A neglected but relevant GARCH spillover evidently represents an omitted variable in the conditional variance equation. Since the omitted lagged conditional variance is clearly positively correlated with the lagged squared residual, the estimate of the ARCH spillover will be biased. Specifically, in cases of erroneously omitting a negative GARCH spillover, the size of the ARCH spillover will be underestimated. In sum, both the size and persistence of cross-effects of volatility innovations are not appropriately determined.

As an alternative to the UECCC GARCH model we also consider the multivariate exponential GARCH (EGARCH) model which – in contrast to the UECCC model – ensures the positive definiteness of the conditional covariance by construction without imposing any constraints on the model parameters. In case of the EGARCH model, we show that a negative GARCH spillover directly implies that the effect of an own volatility innovation is more persistent than the one of a foreign volatility innovation.

Finally, we follow Conrad et al. (1991) and apply both models to the returns of two equity portfolios consisting of large and small firms. In a preliminary step, we neglect potential GARCH spillovers and confirm their result of one-directional ARCH spillovers from large to small firms but not in the opposite direction. However, once we allow for GARCH spillovers the results clearly change. In the UECCC GARCH model we find strong evidence for bi-directional ARCH spillovers in combination with a negative GARCH spillover from small to large firms, which is strong enough to offset the initial effect of the small firm volatility innovation on the large firm conditional variance. That is, our results suggest that small firm volatility innovations do affect large firm volatility, but the effect is less persistent than in case of own large firm volatility innovations. On the other hand, we do not find a GARCH spillover from large to small firms. This suggests that the effect of large firm volatility innovations on the small firm conditional variance is even more persistent than the effect of own small firm volatility innovations. Finally, the results from the EGARCH estimation reconfirm our findings. We argue that these outcomes underline the importance of allowing for adequate flexibility in volatility models. Showing that optimal portfolio weights of small and large firm stocks depend substantially on including or excluding GARCH spillovers further strengthens our argument.

The remainder of the article is organized as follows. In the next section, we introduce the multivariate GARCH and EGARCH models, derive the persistence properties and discuss the consequences of neglecting relevant GARCH spillovers. Section 3 presents the application to stock portfolio returns. The last section concludes.

2 The Model

Let $\mathbf{y}_t = (y_{1,t} \;\, y_{2,t})'$ represent a 2×1 vector of stock returns. Further, let $\mathcal{F}_{t-1} = \sigma(\mathbf{y}_{t-1}, \mathbf{y}_{t-2}, \ldots)$ be the filtration generated by the information available up through time $t-1$. We consider the bivariate process

$$\mathbf{y}_t = \mathbf{E}[\mathbf{y}_t | \mathcal{F}_{t-1}] + \boldsymbol{\varepsilon}_t, \tag{1}$$

where the residual vector $\boldsymbol{\varepsilon}_t = (\varepsilon_{1,t} \;\, \varepsilon_{2,t})'$ is defined as

$$\boldsymbol{\varepsilon}_t = \mathbf{z}_t \odot \mathbf{h}_t^{\wedge 1/2}, \tag{2}$$

with $\mathbf{h}_t = (h_{1,t}\ h_{2,t})'$ being $\mathcal{F}_{t-1}$ measurable and the symbols $\odot$ and $^\wedge$ denote the Hadamard product and the elementwise exponentiation respectively. The stochastic vector $\mathbf{z}_t = (z_{1,t}\ \ z_{2,t})'$ is assumed to be independently and identically distributed (*i.i.d.*) with mean zero, finite second moments, and 2×2 correlation matrix $\mathbf{R} = [\rho_{ij}]_{i,j=1,2}$ with diagonal elements equal to one and off-diagonal elements less than one in absolute value. Thus, we have $\mathbf{E}[\boldsymbol{\varepsilon}_t|\mathcal{F}_{t-1}] = \mathbf{0}$ and $\mathbf{H}_t = \mathbf{E}[\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t'|\mathcal{F}_{t-1}] = diag\{\mathbf{h}_t\}^{1/2}\ \mathbf{R}\ diag\{\mathbf{h}_t\}^{1/2}$. The constant conditional correlation is given by $\rho_{12} = h_{12,t}/\sqrt{h_{1,t}h_{2,t}}$.[2] A meaningful specification for the conditional variances $h_{i,t}$, $i = 1, 2$, must ensure that $\mathbf{H}_t$ is positive definite almost surely for all t. Since $\mathbf{R}_t$ is positive definite by assumption, the specification for the conditional variances has to guarantee that $h_{i,t} > 0$ for all t. In Section 2.1, we model the conditional variances as a UECCC GARCH process and in Section 2.2 as an EGARCH.

2.1 UECCC GARCH

In the first specification we impose the UECCC GARCH(1,1) structure introduced in Conrad and Karanasos (2010) on the conditional variances:

$$\mathbf{h}_t = \boldsymbol{\mu} + \mathbf{A}\boldsymbol{\varepsilon}_{t-1}^{\wedge 2} + \mathbf{B}\mathbf{h}_{t-1}, \tag{3}$$

where $\boldsymbol{\mu} = [\mu_i]_{i=1,2}$, $\mathbf{A} = [a_{ij}]_{i,j=1,2}$ and $\mathbf{B} = [b_{ij}]_{i,j=1,2}$. The UECCC specification allows for non-zero off-diagonal elements in the $\mathbf{A}$ and $\mathbf{B}$ matrices. The coefficients a_{12} and a_{21} measure the ARCH spillovers, i.e. the effects of the squared shocks $\varepsilon_{2,t-1}^2$ and $\varepsilon_{1,t-1}^2$ on the conditional variances $h_{1,t}$ and $h_{2,t}$, respectively. Similarly, the coefficients b_{12} and b_{21} represent the GARCH spillovers, i.e. the effects of the conditional variances $h_{2,t-1}$ and $h_{1,t-1}$ on $h_{1,t}$ and $h_{2,t}$, respectively.

In Bollerslev's (1990) original CCC model, $\mathbf{A}$ and $\mathbf{B}$ are assumed to be diagonal matrices and, hence, the model does neither allow for ARCH nor for GARCH spillovers. Conrad et al. (1991) employ a formulation of the model which allows for non-zero off-diagonal elements of the $\mathbf{A}$ matrix, but restricts $\mathbf{B}$ to be diagonal. That is, their model allows for ARCH but not for GARCH spillovers.[3] Finally, the specification considered in Jeantheau (1998) allows for both ARCH and GARCH spillovers, but under the constraint that $a_{ij} \geq 0$ and $b_{ij} \geq 0$. While these non-negativity restrictions on the entries of the $\mathbf{A}$

[2] Although, for simplicity, we focus on the case of a constant conditional correlation, our results on the persistence of the effects of volatility innovations also hold for specifications which allow for dynamic conditional correlations.

[3] More precisely, Conrad et al. (1991) employ a two-step procedure. In the first step, they estimate univariate GARCH models for each return series. In the second step, they use the squared fitted residual from the first (second) return series as an additional regressor in the second (first) return's conditional variance equation.

and $\mathbf{B}$ matrices ensure that the conditional covariance matrix $\mathbf{H}_t$ is positive definite almost surely for all t, they rule out the possibility of *negative* GARCH spillovers. However, as shown in Conrad and Karanasos (2010), this is only a sufficient but not a necessary condition for the positive definiteness of $\mathbf{H}_t$. Conrad and Karanasos (2010) assume that the model is identified in the sense of Jeantheau (1998, Definition 3.3) and invertible. Defining $\mathbf{B}(L) = \mathbf{I} - \mathbf{B}L$ and $\beta(L) = det[\mathbf{B}(L)] = 1 - \beta_1 L - \beta_2 L^2$, where $\beta_1 = b_{11} + b_{22}$, $\beta_2 = b_{12}b_{21} - b_{11}b_{22}$, L denotes the lag operator and $\mathbf{I}$ is the 2×2 identity matrix, the invertibility condition requires that the inverse roots of $\beta(z)$, denoted by ϕ_1 and ϕ_2, lie inside the unit circle and without loss of generality are ordered as $|\phi_1| \geq |\phi_2|$. Under these two assumptions Conrad and Karanasos (2010) derive the following result:

In the bivariate UECCC GARCH(1,1) the necessary and sufficient conditions for $h_{i,t} > 0$, $i = 1, 2$, *almost surely for all* t *are given by: (i)* $(1-b_{22})\mu_1 + b_{12}\mu_2 > 0$ *and* $(1-b_{11})\mu_2 + b_{21}\mu_1 > 0$, *(ii)* ϕ_1, ϕ_2 *are real and* $\phi_1 \geq |\phi_2|$, *(iii)* $\mathbf{A} \geq 0$ *and (iv)* $[\mathbf{B} - max(\phi_2, 0)\mathbf{I}]\mathbf{A} \geq 0$.

Note that condition (iii) implies that only positive ARCH spillovers are possible. However, as shown in Conrad and Karanasos (2010), Corollary 4, at least one of the two GARCH spillover parameters can take a negative value.[4] In the following derivations we consider a situation with b_{12} being this unrestricted parameter.

Assumption A1 (GARCH spillover) *We assume that* $b_{11} > 0, b_{22} > 0, b_{21} \geq 0$, *while* b_{12} *is unrestricted. Moreover, we assume that* $det(\mathbf{B}) \neq 0$.

The assumption that $det(\mathbf{B}) \neq 0$ is required by Jeantheau's (1998) identifiability condition. Since $det(\mathbf{B}) = -\beta_2$ this condition ensures that $\phi_2 \neq 0$.

Next, it is important to distinguish between the squared shocks $\boldsymbol{\varepsilon}_t^{\wedge 2}$ and the *volatility innovations* $\mathbf{v}_t$ which we define as the squared shock minus its conditional expectation, i.e. $\mathbf{v}_t = \boldsymbol{\varepsilon}_t^{\wedge 2} - \mathbf{h}_t$, such that $\mathbf{E}[\mathbf{v}_t|\mathcal{F}_{t-1}] = \mathbf{0}$. That is, a squared shock $\varepsilon_{i,t}^2$ either implies a positive or negative volatility innovation depending on whether it is bigger or smaller than expected. Using this definition we can rewrite the model as

$$\mathbf{C}(L)\mathbf{h}_t = \boldsymbol{\mu} + \mathbf{A}\mathbf{v}_{t-1}, \tag{4}$$

where $\mathbf{C}(L) = \mathbf{I} - \mathbf{C}L$ and $\mathbf{C} = [c_{ij}]_{i,j=1,2} = \mathbf{A} + \mathbf{B}$. Next, we define the polynomial $\gamma(L) = det[\mathbf{C}(L)] = 1 - \gamma_1 L - \gamma_2 L^2$ with $\gamma_1 = c_{11} + c_{22}$ and $\gamma_2 = c_{12}c_{21} - c_{11}c_{22}$. The following assumption guarantees the weak and strict stationarity of the UECCC GARCH process (see He and Teräsvirta, 2004, or Conrad and Karanasos, 2010, p.859).

[4] In models of dimension higher than two this restriction is further relaxed.

Assumption A2 (Stationarity GARCH) *The inverse roots θ_1 and θ_2 of $\gamma(z)$ are real, satisfy the condition $|\theta_1| < 1$, $|\theta_2| < 1$ and without loss of generality are ordered as $|\theta_1| \geq |\theta_2|$.*[5]

Under Assumption A2, the model can be rearranged as

$$\gamma(L)\mathbf{h}_t = adj[\mathbf{C}(1)]\boldsymbol{\mu} + adj[\mathbf{C}(L)]\mathbf{A}\mathbf{v}_{t-1}, \tag{5}$$

where $adj[\mathbf{C}(L)]$ denotes the adjoint of the matrix $\mathbf{C}(L)$ and we use that $adj[\mathbf{C}(L)] = \gamma(L)[\mathbf{C}(L)]^{-1}$. The unconditional variances of the elements of $\boldsymbol{\varepsilon}_t$ are then given by

$$\mathbf{E}[\boldsymbol{\varepsilon}_t^{\wedge 2}] = \frac{1}{\gamma(1)} adj[\mathbf{C}(1)]\boldsymbol{\mu}. \tag{6}$$

Next, the univariate GARCH(2,2) representation of the bivariate UECCC GARCH model in terms of the volatility innovations $\mathbf{v}_t$ is obtained as

$$\gamma(L)\mathbf{h}_t = adj[\mathbf{C}(1)]\boldsymbol{\mu} + \boldsymbol{\alpha}^{(1)}\mathbf{v}_{t-1} + \boldsymbol{\alpha}^{(2)}\mathbf{v}_{t-2}, \tag{7}$$

where

$$\boldsymbol{\alpha}^{(1)} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \tag{8}$$

$$\boldsymbol{\alpha}^{(2)} = \begin{pmatrix} a_{21}(a_{12}+b_{12}) - a_{11}(a_{22}+b_{22}) & a_{22}b_{12} - a_{12}b_{22} \\ a_{11}b_{21} - a_{21}b_{11} & a_{12}(a_{21}+b_{21}) - a_{22}(a_{11}+b_{11}) \end{pmatrix}. \tag{9}$$

Note that in the univariate representation, $h_{1,t}$ ($h_{2,t}$) depends on the first and second lag of the volatility innovations $v_{1,t}$, $v_{2,t}$ and on two lags of $h_{1,t}$ ($h_{2,t}$). However, $h_{1,t}$ ($h_{2,t}$) does no longer depend on the lagged values of $h_{2,t}$ ($h_{1,t}$) as it was the case in equation (3).

Definition 1 *In analogy to Conrad and Karanasos (2006), we define the impulse response function (IRF) as the effect of an own, $v_{i,t-k}$, or foreign, $v_{j,t-k}$, volatility innovation in $t-k$ on the conditional variance, $h_{i,t}$, i.e. as the sequence of impulse response coefficients*

$$\lambda_{ii}^{(k)} = \frac{\partial h_{i,t}}{\partial v_{i,t-k}} \qquad \text{and} \qquad \lambda_{ij}^{(k)} = \frac{\partial h_{i,t}}{\partial v_{j,t-k}} \tag{10}$$

for $k = 1, 2, \ldots$.

[5] The assumption that the roots are real is not necessary for the stationarity of the process. However, it is realistic in most empirical applications and simplifies the subsequent analysis.

The impulse response coefficients can be obtained from the expansion

$$\mathbf{h}_t = \gamma(1)^{-1} adj[\mathbf{C}(1)]\boldsymbol{\mu} + \boldsymbol{\Lambda}(L)\mathbf{v}_t, \tag{11}$$

where $\boldsymbol{\Lambda}(L) = [\lambda_{ij}(L)]_{i,j=1,2}$ with

$$\lambda_{ij}(L) = \frac{\alpha_{ij}^{(1)} L + \alpha_{ij}^{(2)} L^2}{\gamma(L)} = \sum_{k=1}^{\infty} \lambda_{ij}^{(k)} L^k. \tag{12}$$

Note that each $\lambda_{ij}(L)$ takes the form of a GARCH(2,2) kernel. For illustrative purposes we compare the effects of an own, $v_{1,t-k}$, and a cross, $v_{2,t-k}$, volatility innovation on $h_{1,t}$. The first two impulse response coefficients are given by

$$\begin{aligned} \lambda_{11}^{(1)} &= \alpha_{11}^{(1)} = a_{11} \quad \text{and} \quad \lambda_{11}^{(2)} = \gamma_1 \alpha_{11}^{(1)} + \alpha_{11}^{(2)} = (a_{11}+b_{11})a_{11} + a_{21}(a_{12}+b_{12}) & (13)\\ \lambda_{12}^{(1)} &= \alpha_{12}^{(1)} = a_{12} \quad \text{and} \quad \lambda_{12}^{(2)} = \gamma_1 \alpha_{12}^{(1)} + \alpha_{12}^{(2)} = (a_{11}+b_{11})a_{12} + a_{22}(a_{12}+b_{12}). & (14)\end{aligned}$$

From equation (7) the interpretation of $\lambda_{12}^{(2)}$, for example, is straightforward. A one unit shock in $v_{2,t-2}$ affects $h_{1,t}$ directly by $\alpha_{12}^{(2)}$. In addition, it affects $h_{1,t}$ indirectly via $h_{1,t-1}$ with $\gamma_1 \alpha_{12}^{(1)}$. The combined effect is given by $\lambda_{12}^{(2)} = \gamma_1 \alpha_{12}^{(1)} + \alpha_{12}^{(2)}$.

Hence, we can think of the ARCH parameters a_{ij} as determinants of the size of the *initial impacts* or *first-order effects* of the innovations $v_{1,t-1}$, $v_{2,t-1}$ on $h_{1,t}$ and $h_{2,t}$. Similarly, the GARCH spillovers b_{ij} can be thought of as part of the *second-order effects* of $v_{1,t-2}$, $v_{2,t-2}$ on $h_{1,t}$ and $h_{2,t}$. For example, if Assumption 1 is satisfied and all ARCH spillovers are non-negative (as required by the conditions in Conrad and Karanasos, 2010), the effect of $v_{2,t-2}$ on $h_{1,t}$ will be dampened if $b_{12} < 0$ (see $\lambda_{12}^{(2)}$ in equation (14)).

In general, we can recursively express each $\lambda_{ij}^{(k)}$ sequence as

$$\lambda_{ij}^{(k)} = \gamma_1 \lambda_{ij}^{(k-1)} + \gamma_2 \lambda_{ij}^{(k-2)} \quad \text{for } k \geq 3, \tag{15}$$

where $\lambda_{ij}^{(1)} = \alpha_{ij}^{(1)}$ and $\lambda_{ij}^{(2)} = \gamma_1 \alpha_{ij}^{(1)} + \alpha_{ij}^{(2)}$ (see Conrad and Karanasos, 2010, p.846).

Note that in Bollerslev's (1990) diagonal model ($a_{12} = a_{21} = b_{12} = b_{21} = 0$) we have $\alpha_{ii}^{(1)} L + \alpha_{ii}^{(2)} L^2 = (1 - c_{jj}L)a_{ii}L$, $\alpha_{ij}^{(1)} L + \alpha_{ij}^{(2)} L^2 = 0$, $\gamma(L) = (1 - c_{11}L)(1 - c_{22}L)$ and, hence, the IRFs in equation (12) reduce to

$$\lambda_{11}^{(k)} = a_{11}(a_{11}+b_{11})^{k-1}, \quad \lambda_{22}^{(k)} = a_{22}(a_{22}+b_{22})^{k-1} \quad \text{and} \quad \lambda_{12}^{(k)} = \lambda_{21}^{(k)} = 0 \tag{16}$$

for $k = 1, 2, \ldots$. That is, the cross IRFs are equal to zero and the own IRFs correspond to the ones in the univariate case with rate of decay being governed by $a_{ii} + b_{ii}$. The Lagrange multiplier test suggested in Nakatani and Teräsvirta (2009) considers exactly this case in which $\mathbf{A}$ and $\mathbf{B}$ are diagonal matrices under the null hypothesis.

As discussed before, we can view a GARCH spillover as a second-order effect of the initial ARCH spillover. Since we have assumed that b_{12} is the unrestricted parameter in the $\mathbf{B}$ matrix, in the following we assume that $a_{12} > 0$. In addition, it is meaningful to make an assumption about the relation between the size of the initial impacts a_{ij} of the different volatility innovations. It would be natural to assume that the initial impact of a one unit $v_{1,t}$ innovation on the *own* conditional variance $h_{1,t+1}$ is at least as strong as its initial impact on the *cross* conditional variance $h_{2,t+1}$. Similarly, the initial impact of $v_{2,t}$ on $h_{2,t+1}$ should be at least as strong as its initial impact on $h_{1,t+1}$:

$$\frac{a_{21}}{a_{11}} \leq 1 \leq \frac{a_{22}}{a_{12}}. \tag{17}$$

For deriving our main result, we impose a slightly modified condition.

Assumption A3 (Initial Impact) *We assume that* $a_{11} > 0$*,* $a_{22} > 0$*,* $a_{12} > 0$ *and* $a_{21} \geq 0$*. In addition, the ARCH coefficients satisfy the following condition:*

$$\frac{a_{21}}{a_{11}} < \frac{a_{22}}{a_{12}}. \tag{18}$$

The strict inequality is due to Jeantheau's (1998) identifiability condition which requires that $det(\mathbf{A}) \neq 0$.

If $a_{12} > 0$ and $a_{21} > 0$, we can consider the ratios a_{11}/a_{21} and a_{22}/a_{12} as 'impact ratios'.[6] Then, Assumption A3 can be interpreted as requiring that the geometric mean of the impact ratios, i.e. the average impact ratio, is greater than one:

$$0 < \sqrt{\frac{a_{22}}{a_{12}} \cdot \frac{a_{11}}{a_{21}}} - 1 \tag{19}$$

Since in most empirical applications the size of a_{12} will be quite small in comparison to the size of a_{11} and these two coefficients determine the initial level of the IRFs $\lambda_{11}^{(k)}$ and $\lambda_{12}^{(k)}$, it is inconvenient to directly compare the own and cross IRFs. Instead, we introduce the concept of a *relative* IRF.

Definition 2 *The relative IRFs for the effect of* $v_{1,t-k}$ *and* $v_{2,t-k}$ *on* $h_{1,t}$ *are given by* $\tilde{\lambda}_{11}^{(k)} = \lambda_{11}^{(k)}/\lambda_{11}^{(1)}$ *and* $\tilde{\lambda}_{12}^{(k)} = \lambda_{12}^{(k)}/\lambda_{12}^{(1)}$ *for* $k = 1, 2, \ldots$*.*

The relative IRFs measure the volatility response in relation to the *initial impact* of a volatility innovation. Alternatively, we can think of a relative IRF as *normalizing* the size of the own and foreign volatility innovations such that the initial impact is one.

[6] By the same arguments as before, $a_{21} = 0$ would only make sense if $b_{21} = 0$ as well. This case would imply that there is no second-order causality from the first to the second equation (see Woźniak, 2012).

That is, $\tilde{\lambda}_{11}^{(k)}$ represents the volatility response to a normalized shock $\tilde{v}_{1,t-k}$ of size $1/a_{11}$. Similarly, $\tilde{\lambda}_{12}^{(k)}$ represents the volatility response to a normalized shock $\tilde{v}_{2,t-k}$ of size $1/a_{12}$. By definition, we have that $\tilde{\lambda}_{11}^{(1)} = \tilde{\lambda}_{12}^{(1)} = 1$.

In analogy to the interpretation of the $\lambda_{ij}^{(2)}$, the $\tilde{\lambda}_{11}^{(2)}$ and $\tilde{\lambda}_{12}^{(2)}$ measure the *second-order effects* of the standardized shocks $\tilde{v}_{1,t-2}$ and $\tilde{v}_{2,t-2}$ on $h_{1,t}$. It is then natural to say that the *second-order effect* of the own volatility innovation, $\tilde{v}_{1,t-2}$, on $h_{1,t}$ is at least as strong as the *second-order effect* of the foreign volatility innovation, $\tilde{v}_{2,t-2}$, if $\tilde{\lambda}_{11}^{(2)} \geq \tilde{\lambda}_{12}^{(2)}$. Next, we define a measure for comparing the *persistence* of the effects of own and foreign (standardized) volatility innovations.

Definition 3 *We say that the (relative) effect of an own volatility innovation, $v_{1,t-k}$, on $h_{1,t}$ is at least as persistent as the (relative) effect of a foreign volatility innovation, $v_{2,t-k}$, iff $\tilde{\lambda}_{11}^{(2)} > 0$ and $\tilde{\lambda}_{11}^{(k)} \geq \tilde{\lambda}_{12}^{(k)}$ as long as $\tilde{\lambda}_{11}^{(k)}$ stays positive.*

Intuitively, the definition of persistence requires that – starting from the initial impact of unity – the effect of the own volatility innovation is positive on the second-order, i.e. does not vanish immediately, and – as long as it stays positive – is stronger than the effect of the foreign volatility innovation. For the UECCC GARCH it is straightforward to show that the non-negativity conditions derived in Conrad and Karanasos (2010) directly imply that $\lambda_{11}^{(2)} > 0$ and, hence, $\tilde{\lambda}_{11}^{(2)} > 0$.[7]

Conrad and Karanasos (2010) interpreted a negative GARCH spillover, say $b_{12} < 0$, simply as a situation in which an increase in $h_{2,t}$ leads to a decrease in $h_{1,t+1}$. Such a relationship can be meaningful if, for example, economic theory suggests that there should be a trade-off between the volatilities of two variables such as the trade-off between inflation and output volatility considered in the empirical example in Conrad and Karanasos (2010). Nevertheless, empirically negative GARCH spillovers have been observed in many situations in which a trade-off between volatilities does not appear to be the most plausible explanation.

The following theorem states our main result for the UECCC GARCH model and allows for a new interpretation of negative GARCH spillovers.

Theorem 1 *If Assumptions A1, A2, A3 and the non-negativity conditions hold, then in the UECCC GARCH model the effect of an own volatility innovation $v_{1,t-k}$ on $h_{1,t}$ is at least as persistent as the effect of a foreign volatility innovation $v_{2,t-k}$ on $h_{1,t}$, iff*

$$a_{12} + b_{12} \leq 0. \tag{20}$$

[7] The non-negativity conditions imply that all ARCH(∞) coefficients are non-negative. The ARCH(∞) coefficient that corresponds to $\lambda_{11}^{(2)}$ is given by $\psi_{11}^{(2)} = a_{11}b_{11} + a_{21}b_{12}$. Since $\lambda_{11}^{(2)} = a_{11}^2 + a_{12}a_{21} + \psi_{11}^{(2)}$, it follows that $\lambda_{11}^{(2)} > 0$ if $\psi_{11}^{(2)} \geq 0$.

The proof of Theorem 1 shows that the condition $a_{12} + b_{12} \leq 0$ is already required to ensure a stronger *second-order effect* of the own volatility innovation, i.e. $\tilde{\lambda}_{11}^{(2)} \geq \tilde{\lambda}_{12}^{(2)}$. Then, $\tilde{\lambda}_{11}^{(2)} \geq \tilde{\lambda}_{12}^{(2)}$ combined with the other assumptions implies that $\tilde{\lambda}_{11}^{(k)} \geq \tilde{\lambda}_{12}^{(k)}$ for all k – independent of the sign of $\tilde{\lambda}_{11}^{(k)}$.

Since $a_{12} > 0$ by assumption, the condition $a_{12} + b_{12} \leq 0$ can only be satisfied if $b_{12} < 0$. That is, in the UECCC GARCH model a negative GARCH spillover is a necessary condition for ensuring that the effect of an own volatility innovation is at least as persistent as the effect of a foreign one. Since this is a quite natural situation, Theorem 1 provides a justification for the common finding of negative GARCH spillovers.

In order illustrate our result, we discuss an empirical example taken from Nakatani and Teräsvirta (2008).

Example 1 *Nakatani and Teräsvirta (2008) assume that the conditional variances of two Japanese stock return series can be modeled as a bivariate UECCC GARCH(*$1,1$*) process. They obtain the following estimates of the* $\mathbf{A}$ *and* $\mathbf{B}$ *matrices*

$$\widehat{\mathbf{A}} = \begin{pmatrix} 0.0394 & 0.0341 \\ 0.0350 & 0.1018 \end{pmatrix} \quad \text{and} \quad \widehat{\mathbf{B}} = \begin{pmatrix} 0.9627 & -0.0467 \\ 0.0353 & 0.8093 \end{pmatrix}, \tag{21}$$

where, in order to be in line with our notation, we have interchanged the ordering of the variables. Standard errors are omitted but can be found in Nakatani and Teräsvirta (2008). The estimated parameters clearly satisfy the non-negativity conditions provided in Conrad and Karanasos (2010) and, at the same time, we observe a negative GARCH spillover. Figure 1 shows the corresponding IRFs and relative IRFs. Since $\hat{a}_{12} + \hat{b}_{12} < 0$*, the* $v_{1,t-k}$ *innovation should have a more persistent (relative) effect on* $h_{1,t}$ *than the* $v_{2,t-k}$ *innovation. This is confirmed by the behavior of the relative IRFs* $\tilde{\lambda}_{11}^{(k)}$ *and* $\tilde{\lambda}_{12}^{(k)}$ *in Figure 1, upper right. On the contrary, since* $\hat{b}_{21} > 0$*, the* $v_{2,t-k}$ *innovation has a less persistent (relative) effect on* $h_{2,t}$ *than the* $v_{1,t-k}$ *innovation. Hence, in Figure 1, lower right,* $\tilde{\lambda}_{22}^{(k)}$ *is always below* $\tilde{\lambda}_{21}^{(k)}$*.*

We conclude this section by considering four specific cases.

Case 1: $b_{12} = 0$. This is the type of model considered in Conrad et al. (1991) which allows for an ARCH spillover, $a_{12} > 0$, but not for a GARCH spillover. The model possesses the unpleasant property that – by construction – foreign volatility innovations are more persistent than own volatility innovations. In addition, if in the true model $b_{12} \neq 0$, then imposing the restriction $b_{12} = 0$, i.e. omitting $h_{2,t-1}$ from the first equation, will lead to an omitted variables bias in the estimate of a_{12}. Since $\varepsilon_{2,t}^2$ and $h_{2,t}$ are positively correlated, the estimate of a_{12} will be downward

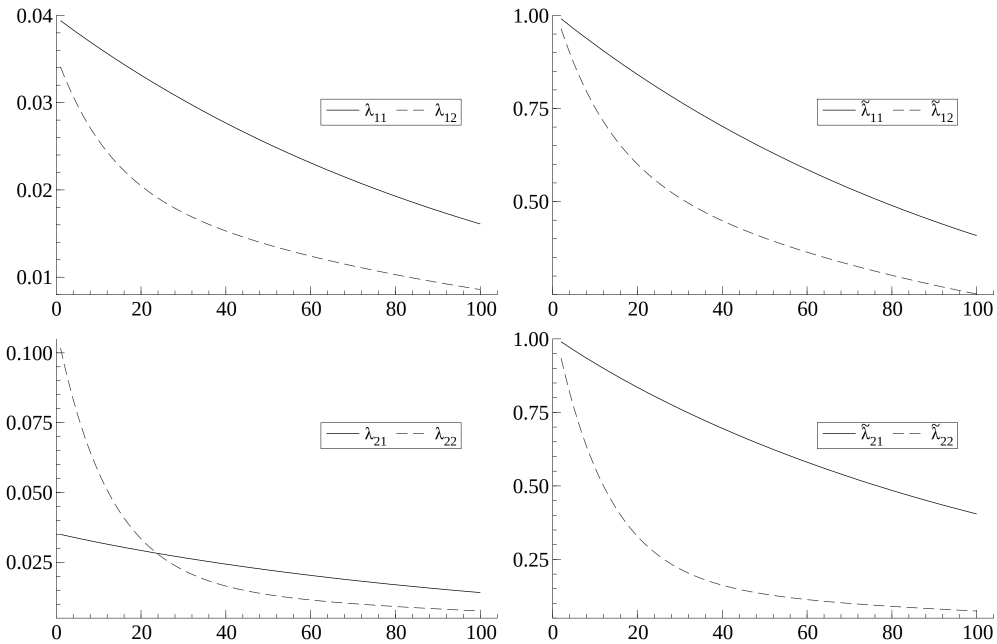

Figure 1: IRFs (upper and lower left figure) and relative IRFs (upper and lower right figure) for the empirical example from Nakatani and Teräsvirta (2008).

biased if the true b_{12} is negative. We will discuss this case in more detail in the empirical application in Section 3.

Case 2: $a_{12} + b_{12} = 0$. In this situation the IRFs in equation (12) reduce to $\lambda_{11}^{(k)} = a_{11}(a_{11}+b_{11})^{k-1}$ – which is the same as in Bollerslev's (1990) diagonal model – and $\lambda_{12}^{(k)} = a_{12}(a_{11} + b_{11})^{k-1}$. Obviously, the two relative IRFs are the same, i.e. $\tilde{\lambda}_{11}^{(k)} = \tilde{\lambda}_{12}^{(k)}$ for all k, meaning that the effect of an own volatility innovation is exactly *as persistent as* the effect of a foreign volatility innovation. Interestingly, although we have ARCH as well as GARCH spillovers, the unconditional variance of $\varepsilon_{1,t}$ is the same as without any spillovers from the second to the first equation ($a_{12} = 0$, $b_{12} = 0$), i.e. the first element in equation (6) reduces to

$$\mathbf{E}[\varepsilon_{1,t}^2] = \frac{\mu_1}{1 - a_{11} - b_{11}}.$$

Case 3: $a_{12} = a_{11}$. That is, the initial impact of the foreign volatility innovation $v_{2,t}$ on $h_{1,t+1}$ is *the same as* the initial impact of the own volatility innovation $v_{1,t}$. In this case we can directly compare the *original* IRFs $\lambda_{11}^{(k)}$ and $\lambda_{12}^{(k)}$ since they differ from the two relative IRFs only by the same factor.

Case 4: $a_{21} = 0$ and $b_{21} = 0$. Under these conditions there are no spillovers from the first to the second equation. This case is equivalent to assuming no second-order causality from the first to the second equation (see Woźniak, 2012). As in Case 2, the $\lambda_{11}^{(k)}$ IRF in equation (12) reduces to the one of the diagonal model $a_{11}(a_{11}+b_{11})^{k-1}$. While $\lambda_{12}^{(1)}$ and $\lambda_{12}^{(2)}$ remain as in equation (14), the recursion given by equation (15) now holds with $\gamma_1 = c_{11} + c_{22}$ and $\gamma_2 = -c_{11}c_{22}$.

2.2 EGARCH

As an alternative to the UECCC GARCH process, we now consider a bivariate EGARCH specification. The EGARCH model has the advantage that it does not require any restrictions on the parameters to ensure the positive definiteness of $\mathbf{H}_t$. The EGARCH structure, as suggested by Weber (2010b), is given by

$$\widetilde{\mathbf{h}}_t = \boldsymbol{\mu} + \mathbf{A}(|\mathbf{z}_{t-1}| - \mathbf{E}[|\mathbf{z}_{t-1}|]) + \mathbf{B}\widetilde{\mathbf{h}}_{t-1}, \tag{22}$$

where $\widetilde{\mathbf{h}}_t = [\ln(h_{i,t})]_{i=1,2}$ and $\boldsymbol{\mu}$, $\mathbf{A}$ and $\mathbf{B}$ are defined as before. The absolute value operation is to be applied element by element. Note that the original univariate EGARCH formulation of Nelson (1991) additionally includes a term which depends on the sign of $\mathbf{z}_{t-1}$ and, hence, allows to model a *leverage effect*. For reasons of parsimony and comparability with the UECCC GARCH model we abstract from this asymmetric term in the following. Note that in equation (22), $\widetilde{\mathbf{h}}_t$ directly depends on the volatility innovations $\mathbf{v}_t = |\mathbf{z}_{t-1}| - \mathbf{E}[|\mathbf{z}_{t-1}|]$, which simplifies the subsequent derivations in comparison to the UECCC GARCH case.

Next, recall from Section 2.1 that we have defined $\mathbf{B}(L) = \mathbf{I} - \mathbf{B}L$ and the corresponding polynomial $\beta(L) = det[\mathbf{B}(L)] = 1 - \beta_1 L - \beta_2 L^2$ with $\beta_1 = b_{11} + b_{22}$ and $\beta_2 = b_{12}b_{21} - b_{11}b_{22}$. Since the EGARCH model does not require non-negativity constraints, both EGARCH spillover parameters can now remain unrestricted. That is, the EGARCH model does not exclude the possibility of negative volatility spillovers in *both* directions. We modify Assumption A1 as follows:

Assumption A4 (EGARCH spillover) *We assume that $b_{11} > 0$ and $b_{22} > 0$. The spillover parameters b_{12} and b_{21} are unrestricted. Moreover, we assume that $det(\boldsymbol{B}) \neq 0$.*

The following assumption states the stationarity condition for the EGARCH model.

Assumption A5 (Stationarity EGARCH) *The inverse roots ϕ_1 and ϕ_2 of $\beta(z)$ are real, satisfy the condition $|\phi_1| < 1$, $|\phi_2| < 1$ and without loss of generality are ordered as $|\phi_1| \geq |\phi_2|$.*

Assumption A5 ensures the existence of the ARCH(∞) representation of the EGARCH process. As an intermediate step, we first derive the univariate representation of the bivariate EGARCH(1,1) in terms of the volatility innovations:

$$\begin{aligned}\beta(L)\widetilde{\mathbf{h}}_t &= adj[\mathbf{B}(1)]\boldsymbol{\mu} + adj[\mathbf{B}(L)]\mathbf{A}|\mathbf{v}_{t-1}| \\ &= adj[\mathbf{B}(1)]\boldsymbol{\mu} + \boldsymbol{\alpha}_1|\mathbf{v}_{t-1}| + \boldsymbol{\alpha}_2|\mathbf{v}_{t-2}|, \end{aligned} \tag{23}$$

with

$$\boldsymbol{\alpha}^{(1)} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \quad \text{and} \quad \boldsymbol{\alpha}^{(2)} = \begin{pmatrix} a_{21}b_{12} - a_{11}b_{22} & a_{22}b_{12} - a_{12}b_{22} \\ a_{11}b_{21} - a_{21}b_{11} & a_{12}b_{21} - a_{22}b_{11} \end{pmatrix}. \tag{24}$$

As for the UECCC GARCH model, we obtain the impulse response functions

$$\lambda_{ii}^{(k)} = \frac{\partial \ln h_{i,t}}{\partial v_{i,t-k}} \quad \text{and} \quad \lambda_{ij}^{(k)} = \frac{\partial \ln h_{i,t}}{\partial v_{j,t-k}} \tag{25}$$

from the coefficients in the expansion

$$\widetilde{\mathbf{h}}_t = \beta(1)^{-1} adj[\mathbf{B}(1)]\boldsymbol{\mu} + \boldsymbol{\Lambda}(L)|\mathbf{v}_t|, \tag{26}$$

where $\boldsymbol{\Lambda}(L) = [\lambda_{ij}(L)]_{i,j=1,2}$ with

$$\lambda_{ij}(L) = \frac{\alpha_{ij}^{(1)}L + \alpha_{ij}^{(2)}L^2}{\beta(L)} = \sum_{k=1}^{\infty} \lambda_{ij}^{(k)} L^k. \tag{27}$$

Note that the IRF of the EGARCH differs from the one of the UECCC GARCH because we now measure the effect of a volatility innovation on the *log* of the conditional variance. In addition, in equation (27) we divide by $\beta(L)$ instead of $\gamma(L)$ and, further, the diagonal elements of $\boldsymbol{\alpha}^{(2)}$ are different. Hence, the first two impulse response coefficients for the effects of $v_{1,t-1}$ and $v_{2,t-1}$ on $h_{1,t}$ and $h_{1,t+1}$ are now given by

$$\lambda_{11}^{(1)} = a_{11} \quad \text{and} \quad \lambda_{11}^{(2)} = a_{11}b_{11} + a_{21}b_{12} \tag{28}$$

$$\lambda_{12}^{(1)} = a_{12} \quad \text{and} \quad \lambda_{12}^{(2)} = a_{12}b_{11} + a_{22}b_{12}. \tag{29}$$

That is, the initial impacts $\lambda_{11}^{(1)}$ and $\lambda_{12}^{(1)}$ of an own and foreign volatility innovation are formally the same as in the UECCC GARCH model. However, for $k \geq 2$ the effects are different. Similarly as before, we can recursively express each $\lambda_{ij}^{(k)}$ sequence as

$$\lambda_{ij}^{(k)} = \beta_1 \lambda_{ij}^{(k-1)} + \beta_2 \lambda_{ij}^{(k-2)} \quad \text{for } k \geq 3, \tag{30}$$

where $\lambda_{ij}^{(1)} = \alpha_{ij}^{(1)}$ and $\lambda_{ij}^{(2)} = \beta_1 \alpha_{ij}^{(1)} + \alpha_{ij}^{(2)}$.

In the EGARCH specification $\widetilde{\mathbf{h}}_t$ directly depends on $\mathbf{v}_t$ and equation (22) allows for an interpretation of the effect of the volatility innovations. For example, a volatility innovation $v_{2,t}$ affects $\tilde{h}_{1,t+1}$ directly with a_{12} and $\tilde{h}_{2,t+1}$ with a_{22}. Hence, it affects $\tilde{h}_{1,t+2}$ via $\tilde{h}_{1,t+1}$ with $a_{12}b_{11}$ and via $\tilde{h}_{2,t+1}$ with $a_{22}b_{12}$. The combined effect is given by: $\lambda_{12}^{(2)} = a_{12}b_{11} + a_{22}b_{12}$.

Since the initial effects are the same as in the UECCC GARCH model, we maintain Assumption A3 for the EGARCH model as well. Similarly, the definition of the relative IRFs remains as before. Since the EGARCH model does not require non-negativity conditions, in principle, it is possible that $\tilde{\lambda}_{11}^{(2)} \leq 0$, even if this case does not seem to be practically relevant. However, in order to be able to apply our definition of persistence we assume that $\tilde{\lambda}_{11}^{(2)} > 0$. Next, we present the theorem for the EGARCH model.

Theorem 2 *If Assumptions A3, A4 and A5 hold and $\tilde{\lambda}_{11}^{(2)} > 0$, then in the EGARCH model the effect of an own volatility innovation $v_{1,t-k}$ on $h_{1,t}$ is at least as persistent as the effect of a foreign volatility innovation $v_{2,t-k}$ on $h_{1,t}$, iff*

$$b_{12} \leq 0.$$

Note that for the EGARCH it suffices that $b_{12} \leq 0$. This is in contrast to the UECCC GARCH case where we need that $a_{12} + b_{12} \leq 0$.

Again, we close this section by considering some specific cases of interest.

Case 1: $b_{12} = 0$. This is the EGARCH version of the Conrad et al. (1991) model with an ARCH but no GARCH spillover. The two impulse response functions reduce to $\lambda_{11}^{(k)} = a_{11}b_{11}^{k-1}$ and $\lambda_{12}^{(k)} = a_{12}b_{11}^{k-1}$ for $k = 1, 2, \ldots$. Clearly, the two relative IRFs are the same. That is, in the EGARCH model the restriction of no GARCH spillover is equivalent to assuming that the persistence of a foreign volatility innovation is the *same* as the persistence of an own one. On the contrary, in the UECCC GARCH model $b_{12} = 0$ implies that the effect of the foreign innovation is *more* persistent. Finally, note that if $b_{12} = 0$ and/or $b_{21} = 0$, we have $\beta(L) = (1 - b_{11}L)(1 - b_{22}L)$ and, hence, $\phi_1 > 0$ and $\phi_2 > 0$ by Assumption A4.

Case 2: $a_{12} = a_{11}$. Since in this case $v_{2,t}$ and $v_{1,t}$ have the same initial impact on $h_{1,t+1}$, we can directly compare the IRFs.

Case 3: $a_{21} = b_{21} = 0$. If there are neither ARCH nor GARCH volatility spillovers from the first to the second equation, then $\lambda_{11}^{(k)} = a_{12}b_{11}^{k-1}$ as in Case 1. The $\lambda_{12}^{(k)}$ remain as in equation (30) but with $\beta(L) = (1 - b_{11}L)(1 - b_{22}L)$.

3 Empirical Application

3.1 Data

We apply our methodological considerations to a research field particularly occupied with variance spillovers: the literature on stock market transmission between small and large firms (Conrad et al. 1991, amongst others). For this purpose we employ the Fama-French size-sorted portfolios available at Kenneth French's homepage. These are constructed annually at the end of June by sorting all NYSE, AMEX, and NASDAQ stocks according to their June market equity. We obtained daily returns for the portfolios consisting of the bottom and top 30%, respectively. Furthermore, we conduct some robustness analysis for the bottom and top 20% and 10%. The sample is July 1, 1963 until December 31, 2009.

As argued by Conrad et al. (1991), data at the daily frequency might still be subject to microstructure effects like nontrading and bid-ask bounce. Therefore, these authors aggregated the daily series to weekly returns, where these issues can be safely ignored. We follow them in this point, but we use the daily data for robustness checks. Table 1 shows summary statistics of the weekly 30% quantiles.

Table 1: Summary statistics of Fama-French size-sorted weekly 30% portfolio returns.

Return	Mean	St.Dev.	Skewness	Kurtosis	ACF(1)
Small cap	0.227	2.404	-1.005	8.818	0.219
Large cap	0.195	2.165	-0.450	10.254	-0.036

We encounter well-known stylized facts: slightly higher mean and standard deviation of small firm returns, considerable excess kurtosis, clearly positive first-order autocorrelation (ACF(1)) of small cap and slightly negative autocorrelation of large cap returns. Figure 2 provides a graphical impression of the weekly data. The presence of pronounced volatility clustering is clearly visible.

3.2 UECCC GARCH Results

We specify $\mathbf{E}[\mathbf{y}_t|\mathcal{F}_{t-1}]$ as a VAR(p) and estimate the VAR-UECCC GARCH equations simultaneously by numerical quasi-maximum likelihood (QML) using the BHHH algorithm (see Bollerslev and Wooldrigde, 1992). The Schwarz criterion prefers one lag in the VAR. First, in order to replicate conventional findings such as in Conrad et al. (1991), we set $b_{12} = b_{21} = 0$ in the $\mathbf{B}$ matrix. Estimation of the bivariate VAR(1)-UECCC(1,1) GARCH model with large firm returns as the first variable leads to the following result for the

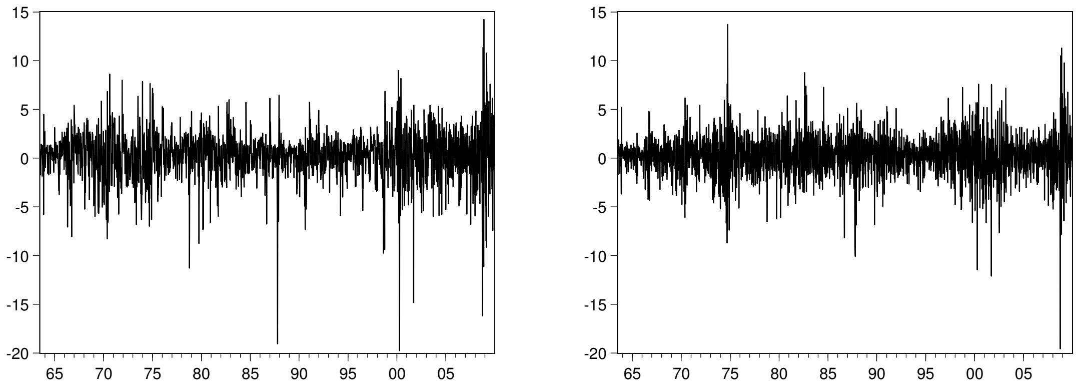

Figure 2: Returns of 30% portfolios of small (left) and large (right) firms.

conditional variance equations (numbers in parentheses are robust standard errors):

$$\begin{pmatrix} h_{L,t} \\ h_{S,t} \end{pmatrix} = \begin{pmatrix} \underset{(0.031)}{0.116} \\ \underset{(0.057)}{0.217} \end{pmatrix} + \begin{pmatrix} \underset{(0.028)}{0.142} & \underset{(0.008)}{-0.011} \\ \underset{(0.018)}{0.036} & \underset{(0.027)}{0.137} \end{pmatrix} \begin{pmatrix} \varepsilon^2_{L,t-1} \\ \varepsilon^2_{S,t-1} \end{pmatrix} + \begin{pmatrix} \underset{(0.030)}{0.853} & \underset{(-)}{0} \\ \underset{(-)}{0} & \underset{(0.033)}{0.799} \end{pmatrix} \begin{pmatrix} h_{L,t-1} \\ h_{S,t-1} \end{pmatrix} . \tag{31}$$

While a_{21} is positive and significant at the 5% level, the point estimate of a_{12} is slightly negative (which would violate the non-negativity constraints concerning the $\mathbf{A}$ matrix) but clearly insignificant. This tends to confirm the standard outcome that ARCH spillovers are running from large to small firms, but not vice versa. Next, we turn our attention to the model including GARCH spillovers:

$$\begin{pmatrix} h_{L,t} \\ h_{S,t} \end{pmatrix} = \begin{pmatrix} \underset{(0.039)}{0.184} \\ \underset{(0.080)}{0.315} \end{pmatrix} + \begin{pmatrix} \underset{(0.019)}{0.117} & \underset{(0.014)}{0.027} \\ \underset{(0.021)}{0.041} & \underset{(0.030)}{0.168} \end{pmatrix} \begin{pmatrix} \varepsilon^2_{L,t-1} \\ \varepsilon^2_{S,t-1} \end{pmatrix} + \begin{pmatrix} \underset{(0.016)}{0.890} & \underset{(0.019)}{-0.062} \\ \underset{(-)}{0} & \underset{(0.039)}{0.747} \end{pmatrix} \begin{pmatrix} h_{L,t-1} \\ h_{S,t-1} \end{pmatrix} . \tag{32}$$

In a preliminary step, the GARCH spillover b_{21} was eliminated due to clear insignificance. Based on the QML standard errors, both ARCH spillover coefficients are significant at the 10% level. However, we avoid relying exclusively on numerical QML standard errors by additionally conducting likelihood ratio (LR) tests. Both the null hypotheses of $a_{12} = 0$ and $a_{21} = 0$ are clearly rejected with LR statistics of 15.19 and 9.85, respectively. This shows that ARCH spillovers are indeed bi-directional, whereas the bulk of the literature failed to establish spillovers from small to large firms. Most importantly, b_{12} is highly significant and negative, highlighting two facts established in this paper: First, since $\hat{a}_{12} + \hat{b}_{12} < 0$, the persistence of small firm volatility innovations to large firm volatility is lower than the persistence of own large firm volatility innovations. Second, it is likely that the absence of empirical evidence in favor of small to large firm ARCH spillovers in the literature is due to an omitted variable bias caused by ignoring GARCH spillovers. Finally, the fact that $b_{21} = 0$ in combination with $\hat{a}_{21} > 0$ shows another important result: large firm volatility innovations have a more persistent effect on small firm volatility than small firm volatility innovations themselves.

It should be mentioned that the parameter estimates in model (32) violate the non-negativity conditions provided in Conrad and Karanasos (2010). This implies that forecasts based on these parameter estimates may generate conditional covariance matrices that are not positive definite. In order to avoid such problems, we prefer to rely on a model that guarantees positive definiteness of the covariance matrix by construction. For that purpose we also consider the EGARCH approach, to which we turn next.

3.3 EGARCH Results

Again, we begin with the restricted EGARCH model, setting $b_{12} = b_{21} = 0$.[8]

$$\begin{pmatrix} \tilde{h}_{L,t} \\ \tilde{h}_{S,t} \end{pmatrix} = \begin{pmatrix} \underset{(0.022)}{0.063} \\ \underset{(0.031)}{0.106} \end{pmatrix} + \begin{pmatrix} \underset{(0.073)}{0.268} & \underset{(0.036)}{0.008} \\ \underset{(0.044)}{0.091} & \underset{(0.058)}{0.257} \end{pmatrix} \begin{pmatrix} |z_{L,t-1}| \\ |z_{S,t-1}| \end{pmatrix} + \begin{pmatrix} \underset{(0.015)}{0.956} & \underset{(-)}{0} \\ \underset{(-)}{0} & \underset{(0.018)}{0.938} \end{pmatrix} \begin{pmatrix} \tilde{h}_{L,t-1} \\ \tilde{h}_{S,t-1} \end{pmatrix} \tag{33}$$

We obtain the same result as in the UECCC GARCH case: a significant large to small firm ARCH spillover (LR=9.71), but an insignificant small to large firm ARCH spillover (LR=0.10). Figure 3 shows the corresponding IRFs and relative IRFs.

In accordance with the insignificant a_{12} small firm volatility innovations have a negligible effect on large firm volatility ($\lambda_{12}^{(k)}$, Figure 3, upper left), while large firm volatility innovations have a considerable effect on small firm volatility ($\lambda_{21}^{(k)}$, Figure 3, lower left). Since $\hat{a}_{21} < \hat{a}_{22}$ the initial impact of small firm innovations on small firm volatility is stronger than the initial impact of large firm innovations. Finally, since $b_{12} = b_{21} = 0$ by assumption, the persistence of own and foreign volatility innovations is the same in both equations (Figure 3, upper and lower right).

Estimating the general system, where the insignificant b_{21} is again set to zero, leads to

$$\begin{pmatrix} \tilde{h}_{L,t} \\ \tilde{h}_{S,t} \end{pmatrix} = \begin{pmatrix} \underset{(0.027)}{0.105} \\ \underset{(0.037)}{0.144} \end{pmatrix} + \begin{pmatrix} \underset{(0.047)}{0.208} & \underset{(0.044)}{0.106} \\ \underset{(0.059)}{0.101} & \underset{(0.059)}{0.284} \end{pmatrix} \begin{pmatrix} |z_{L,t-1}| \\ |z_{S,t-1}| \end{pmatrix} + \begin{pmatrix} \underset{(0.005)}{0.986} & \underset{(0.017)}{-0.053} \\ \underset{(-)}{0} & \underset{(0.022)}{0.913} \end{pmatrix} \begin{pmatrix} \tilde{h}_{L,t-1} \\ \tilde{h}_{S,t-1} \end{pmatrix} . \tag{34}$$

Now, the null hypotheses of $a_{12} = 0$ and $a_{21} = 0$ are clearly rejected with LR statistics of 15.63 and 10.68. This reconfirms the presence of bi-directional ARCH spillover effects. The GARCH spillover from small firm to large firm volatility, b_{12}, is again significantly

[8] Assuming normally distributed $z_{i,t}$ we redefine the constant as $\tilde{\boldsymbol{\mu}} = \boldsymbol{\mu} - \mathbf{A}\mathbf{E}[|\mathbf{z}_{t-1}|] = \boldsymbol{\mu} - \mathbf{A}\mathbf{i}\sqrt{2/\pi}$, where $\mathbf{i} = (1,1)'$.

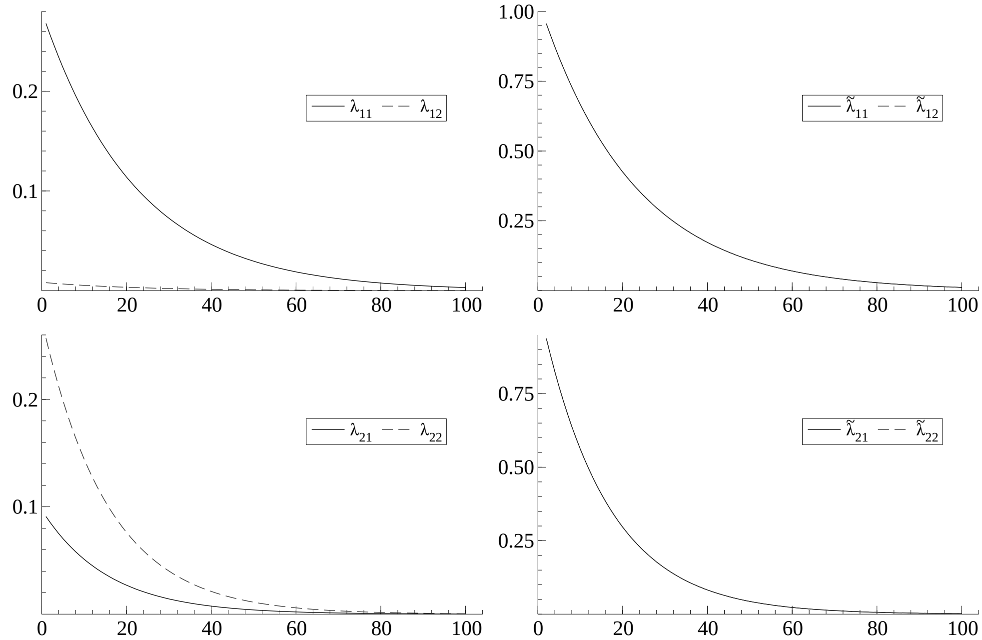

Figure 3: IRFs (upper and lower left figure) and relative IRFs (upper and lower right figure) for EGARCH estimates from equation (33).

negative. That is, own volatility innovations to the large firm portfolio are more persistent than small firm volatility innovations. The IRFs and relative IRFs in Figure 4, upper left and right, indeed show that a small firm volatility innovation first has a positive effect on the volatility of large firms, and that this effect vanishes rapidly. Furthermore, the effect even turns negative subsequently. Evidently, the existing literature did not only impose identical persistence ($b_{12} = 0$) of different volatility shocks, but it also neglected the possibility of negative dynamic responses.

3.4 Optimal Portfolio Choice

The empirical investigation has established the econometric relevance of our theoretical considerations on GARCH spillovers. In this section we shall corroborate the economic significance of our results. Since equity markets consist of firms of very different size, portfolio managers have a natural interest in finding an optimal mixture of stocks. This leads us to the problem of computing optimal fully invested portfolio holdings, where we impose the no-shorting constraint. We focus on the second moments and do not attempt to forecast returns themselves. Thus, we assume expected returns to be zero, making the problem equivalent to determining conditional risk-minimizing portfolio weights.

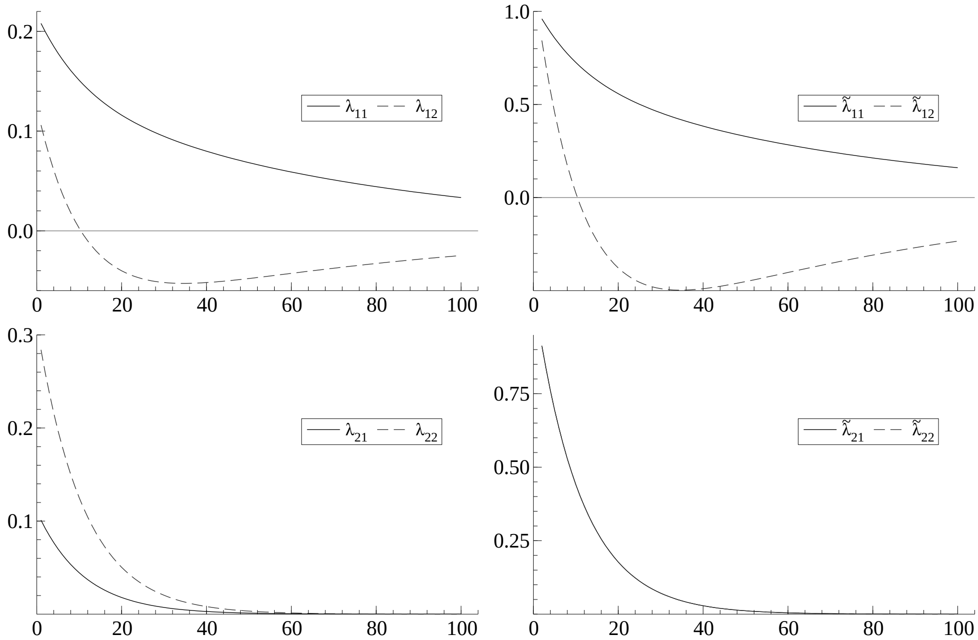

Figure 4: IRFs (upper and lower left figure) and relative IRFs (upper and lower right figure) for EGARCH estimates from equation (34).

Given a mean-variance utility function, one can derive the following optimal portfolio holdings w_t for the large firm (sub-)portfolio: $w_t = 0$ if $w_t^* < 0$, $w_t = 1$ if $w_t^* > 1$ and $w_t = w_t^*$ else. Therein, w_t^* is given by

$$w_t^* = \frac{h_{S,t} - h_{LS,t}}{h_{L,t} - 2h_{LS,t} + h_{S,t}} , \tag{35}$$

with $h_{LS,t}$ denoting the conditional covariance between the small and large firm portfolios. The optimal holdings of the small firm portfolio are $1 - w_t$.

We compare the choice of portfolio weights made by an investor using the misspecified EGARCH model (33) to the optimal choice based on the correctly specified model (34). Figure (5) plots the estimated weights against each other. While the positive relationship is not surprising, we see considerable deviations from the 45°-line. This implies that portfolio choices based on the two specifications of the volatility model are often substantially different. In particular, GARCH spillovers and small firm influences clearly prove to be relevant for decision-making in financial markets.

3.5 Robustness Analysis

As mentioned above we conducted several robustness checks:

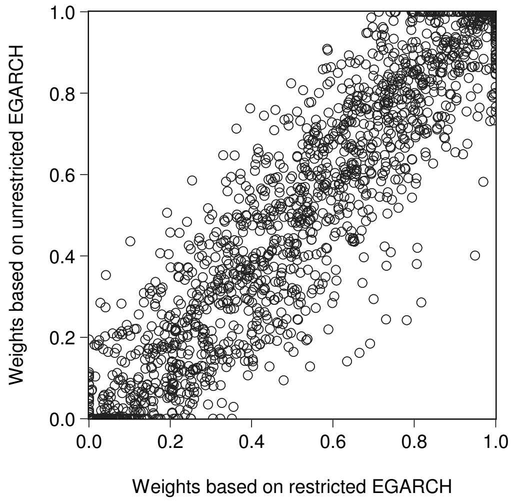

Figure 5: Optimal large firm portfolio weights with and without GARCH spillovers

- We corroborate the established effect in daily data, i.e. small to large firm ARCH spillovers can only be detected once the model takes the (negative) GARCH spillover into account.
- The results are confirmed taking the top and bottom deciles and quintiles from the Fama-French database instead of the 30% portfolios.
- Instead of the CCC specification we employed the DCC model of Engle (2002), thus allowing for time-varying correlations. The results for the volatility equations, and especially the spillovers, were qualitatively unchanged.
- Above, following the Schwarz criterion, we preferred a parsimonious VAR(1) specification for the mean equations, while other criteria, i.e. Hannan-Quinn and Akaike, naturally choose higher lags. This leaves the volatility equations largely unaffected.

The detailed estimation results are omitted for reasons of brevity, but are of course available from the authors upon request.

4 Conclusion

We reconsider the existence of volatility spillovers in multivariate GARCH models. In particular, we show that the existence of negative GARCH spillovers can be rationalized by the fact that a negative GARCH spillover is a necessary (and sufficient) condition in a multivariate GARCH (EGARCH) model to guarantee that the effect of an own volatility innovation is at least as persistent as the effect of a foreign volatility innovation. If the

GARCH spillover is constraint to be zero this imposes an unintended condition on the relation between the persistence of the effects of own and foreign volatility innovations. In addition, it leads to a bias in the estimate of the initial impact of the corresponding ARCH spillover.

We conclude that our main result represents a robust empirical fact: allowing for negative GARCH spillovers uncovers a significant impact of small firm volatility innovations on large firm portfolio volatility. While small to large firm volatility innovations have a less persistent effect than own large firm innovations, large to small firm innovations have a more persistent effect than own small firm innovations. As a general device, we suggest that investigations of volatility interactions should therefore always start from unrestricted multivariate models.

References

[1] Arouri, M.H., J. Jouini, D.K. Nguyen (2011): Volatility spillovers between oil prices and stock sector returns: implications for portfolio management. *Journal of International Money and Finance*, 30, 1387-1405

[2] Bollerslev, T. (1990): Modelling the coherence in short-run nominal exchange rates: a multivariate generalised ARCH model. *Review of Economics and Statistics*, 72, 498-505.

[3] Bollerslev, T., J.M. Wooldridge (1992): Quasi-Maximum Likelihood Estimation and Inference in Dynamic Models with Time Varying Covariances. *Econometric Reviews*, 11, 143-172.

[4] Chang, C.-L., M. McAleer, R. Tansuchat (2010): Analyzing and forecasting volatility spillovers, asymmetries and hedging in major oil markets. *Energy Economics*, 32, 1445-1455.

[5] Chelley-Steeley, P.L., J.M. Steeley (2005): The leverage effect in the UK stock market. *Applied Financial Economics*, 15, 409-423.

[6] Conrad, C., M. Karanasos (2006): The impulse response function of the long memory GARCH process. *Economics Letters*, 90, 34-41.

[7] Conrad, C., M. Karanasos (2010): Negative volatility spillovers in the unrestricted ECCC-GARCH model. *Econometric Theory*, 26, 838-862.

[8] Conrad, J., M.N. Gultekin, G. Kaul (1991): Asymmetric predictability of conditional variances. *Review of Financial Studies*, 4, 597-622.

[9] Engle, R.F. (2002): Dynamic conditional correlation: a simple class of multivariate generalized autoregressive conditional heteroscedasticity models. *Journal of Business and Economic Statistics*, 20, 339-350.

[10] Francq, C., J.-M. Zakoïan (2012): QML estimation of a class of multivariate asymmetric GARCH models. *Econometric Theory*, 28, 179-206.

[11] Hakim, A., M. McAleer (2010): Modelling the interactions across international stock, bond and foreign exchange markets. *Applied Economics*, 42, 825-850.

[12] He, C., T. Teräsvirta (2004): An extended constant conditional correlation GARCH model and its fourth-moment structure. *Econometric Theory*, 20, 904-926.

[13] Jeantheau, T. (1998): Strong consistency of estimators for multivariate ARCH models. *Econometric Theory*, 14, 70-86.

[14] Koutmos, G., G.G. Booth (1995): Asymmetric volatility transmission in international stock markets. *Journal of International Money and Finance*, 14, 747-762.

[15] Laopodis, N.T. (2003): Stochastic behaviour of Deutsche mark exchange rates within EMS. *Applied Financial Economics*, 13, 665-676.

[16] McAleer, M., B. da Veiga (2008): Forecasting Value-at-Risk with a parsimonious portfolio spillover GARCH (PS-GARCH) model. *Journal of Forecasting*, 27, 1-19.

[17] Nakatani, T., T. Teräsvirta (2008): Positivity constraints on the conditional variances in the family of conditional correlation GARCH models. *Finance Research Letters*, 5, 88-95.

[18] Nelson, D.B. (1991): Conditional heteroskedasticity in asset returns: a new approach. *Econometrica*, 59, 347-370.

[19] Ross, S.A. (1989): Information and volatility: the no-arbitrage martingale approach to timing and resolution irrelevancy. *Journal of Finance*, 44, 1-17.

[20] Rittler, D. (2012): Price discovery and volatility spillovers in the European Union emissions trading scheme: a high-frequency analysis. *Journal of Banking & Finance*, 36, 774-785.

[21] Savva, C.S., D.R. Osborn, L. Gill (2009): Spillovers and correlations between US and major European stock markets: the role of the euro. *Applied Financial Economics*, 19, 1595-1604.

[22] Skintzi, V.D., A.N. Refenes (2006): Volatility spillovers and dynamic correlation in European bond markets. *Journal of International Financial Markets, Institutions and Money*, 16, 23-40.

[23] Tsai, H., K.S. Chan (2008): A note on inequality constraints in the GARCH model. *Econometric Theory*, 24, 823-828.

[24] Weber, E. (2010a): Structural conditional correlation. *Journal of Financial Econometrics*, 8, 392-407.

[25] Weber, E. (2010b): Volatility and causality in Asia Pacific financial markets. *Applied Financial Economics*, 20, 1269-1292.

[26] Wong, H., W. Li, S. Ling (2005): Joint modeling of cointegration and conditional heteroscedasticity with applications. *Annals of the Institute of Statistical Mathematics*, 57, 83-103.

[27] Wozniak, T. (2012): Testing causality between two vectors in multivariate GARCH models. Economics Working Papers ECO2012/20, European University Institute.

A Proofs

Proof of Theorem 1.

We denote the difference in the two relative impulse response functions by $\delta_{12}(L) = \tilde{\lambda}_{11}(L) - \tilde{\lambda}_{12}(L)$. Using Assumption 2, we obtain the following expression

$$\begin{aligned}
\delta_{12}(L) = \tilde{\lambda}_{11}(L) - \tilde{\lambda}_{12}(L) &= \frac{\alpha_{11}^{(1)} L + \alpha_{11}^{(2)} L^2}{\alpha_{11}^{(1)} \gamma(L)} - \frac{\alpha_{12}^{(1)} L + \alpha_{12}^{(2)} L^2}{\alpha_{12}^{(1)} \gamma(L)} && (36)\\
&= \frac{\left(\frac{\alpha_{11}^{(2)}}{\alpha_{11}^{(1)}} - \frac{\alpha_{12}^{(2)}}{\alpha_{12}^{(1)}}\right) L^2}{(1-\theta_1 L)(1-\theta_2 L)} && (37)\\
&= \frac{1}{(1-\theta_2 L)} \left(\frac{\alpha_{11}^{(2)}}{\alpha_{11}^{(1)}} - \frac{\alpha_{12}^{(2)}}{\alpha_{12}^{(1)}}\right) L^2 \sum_{k=0}^{\infty} \theta_1^k L^k && (38)\\
&= \frac{1}{(1-\theta_2 L)} \delta_{12}^{(1)}(L), && (39)
\end{aligned}$$

where we define

$$\delta_{12}^{(1)}(L) = \sum_{k=0}^{\infty} \delta_{12,k}^{(1)} L^k \qquad (40)$$

with $\delta_{12,0}^{(1)} = 0$, $\delta_{12,1}^{(1)} = 0$ and

$$\begin{aligned}
\delta_{12,2}^{(1)} &= \left(\frac{\alpha_{11}^{(2)}}{\alpha_{11}^{(1)}} - \frac{\alpha_{12}^{(2)}}{\alpha_{12}^{(1)}}\right) && (41)\\
\delta_{12,k}^{(1)} &= \theta_1 \delta_{12,k-1}^{(1)} = \ldots = \theta_1^{k-2} \delta_{12,2}^{(1)} \quad \text{for} \quad k \geq 3. && (42)
\end{aligned}$$

In the next step we obtain a recursive representation of the coefficients, $\delta_{12,k}$, in

$$\delta_{12}(L) = \left(\sum_{k=0}^{\infty} \theta_2^k L^k\right) \delta_{12}^{(1)}(L) = \sum_{k=0}^{\infty} \delta_{12,k} L^k$$

with

$$\delta_{12,k} = \sum_{l=0}^{k} \theta_2^l \delta_{12,k-l}^{(1)}.$$

The recursive representation is given by

$$\begin{aligned}
\delta_{12,k} &= \sum_{l=0}^{k} \theta_2^l \delta_{12,k-l}^{(1)} = \sum_{l=1}^{k} \theta_2^l \delta_{12,k-l}^{(1)} + \delta_{12,k}^{(1)}\\
&= \theta_2 \sum_{l=0}^{k-1} \theta_2^l \delta_{12,k-l-1}^{(1)} + \delta_{12,k}^{(1)}\\
&= \theta_2 \delta_{12,k-1} + \delta_{12,k}^{(1)}. && (43)
\end{aligned}$$

Next, we show that under the Assumptions stated in Theorem 1 all $\delta_{12,k}$ are non-negative.

Consider the sequence $\delta_{12,k}^{(1)}$ for $k = 2, 3, \ldots$. Using equation (8) we can simplify $\delta_{12,2}^{(1)}$ to

$$\begin{aligned} \delta_{12,2}^{(1)} &= \left(\frac{\alpha_{11}^{(2)}}{\alpha_{11}^{(1)}} - \frac{\alpha_{12}^{(2)}}{\alpha_{12}^{(1)}} \right) = \left[\frac{a_{21}}{a_{11}}(a_{12} + b_{12}) - (a_{22} + b_{22}) \right] - \left[\frac{a_{22}}{a_{12}} b_{12} - b_{22} \right] \\ &= \left(\frac{a_{21}}{a_{11}} - \frac{a_{22}}{a_{12}} \right) (a_{12} + b_{12}). \end{aligned} \tag{44}$$

Since $a_{21}/a_{11} - a_{22}/a_{12} < 0$ by Assumption A3, $\delta_{12,2}^{(1)} \geq 0$ if and only if

$$c_{12} = a_{12} + b_{12} \leq 0.$$

Note that Assumptions A1 and A3 imply that $\gamma_1 = c_{11} + c_{22} > 0$ and $c_{21} \geq 0$. In addition, if $c_{12} = a_{12} + b_{12} \leq 0$ then $\gamma_2 = c_{12}c_{21} - c_{11}c_{22} < 0$. Since $\gamma_1 = \theta_1 + \theta_2$ and $\gamma_2 = -\theta_1\theta_2$, it follows that $\theta_1 > 0$ and $\theta_2 > 0$ in this case. Because $\theta_1 > 0$, it directly follows that $\delta_{12,k}^{(1)} = \theta_1^{k-2}\delta_{12,2}^{(1)} \geq 0$ for $k \geq 3$.

Finally, note that $\delta_{12,0} = \delta_{12,1} = 0$ and $\delta_{12,2} = \delta_{12,2}^{(1)} \geq 0$. Hence, the recursive representation

$$\delta_{12,k} = \theta_2 \delta_{12,k-1} + \delta_{12,k}^{(1)}$$

with $\theta_2 > 0$ implies that $\delta_{12,k} \geq 0$ for $k \geq 3$. ■

Proof of Theorem 2.

Theorem 2 can be proven analogously to Theorem 1 when three small adjustments are made. First, replace $\gamma(L)$, θ_1 and θ_2 by $\beta(L)$, ϕ_1 and ϕ_2. Second, because the $\boldsymbol{\alpha}^{(2)}$ matrix is now given by equation (24), equation (44) changes to

$$\begin{aligned} \delta_{12,2}^{(1)} &= \left(\frac{\alpha_{11}^{(2)}}{\alpha_{11}^{(1)}} - \frac{\alpha_{12}^{(2)}}{\alpha_{12}^{(1)}} \right) = \left[\frac{a_{21}b_{12}}{a_{11}} - b_{22} \right] - \left[\frac{a_{22}b_{12}}{a_{12}} - b_{22} \right] \\ &= \left(\frac{a_{21}}{a_{11}} - \frac{a_{22}}{a_{12}} \right) b_{12}. \end{aligned} \tag{45}$$

Hence, under Assumption A3, in the EGARCH $\delta_{12,2}^{(1)} \geq 0$ if and only if

$$b_{12} \leq 0.$$

Note that Assumption A4 implies that $\beta_1 = b_{11} + b_{22} = \phi_1 + \phi_2 > 0$. In combination with Assumption A5 it is clear that $\phi_1 > 0$ and, hence, $\delta_{12,k}^{(1)} = \phi_1^{k-2}\delta_{12,2}^{(1)} \geq 0$ for $k \geq 3$. Since in the EGARCH model b_{21} is unrestricted, we now have to distinguish two cases.

Case 1: If $b_{12} \leq 0$ and $b_{21} \geq 0$, then $\beta_2 = b_{12}b_{21} - b_{11}b_{22} < 0$. Since $\beta_2 = -\phi_1\phi_2$, this implies that $\phi_2 > 0$ and we can proceed as in the proof of Theorem 1.

Case 2: If $b_{12} \leq 0$ and $b_{21} < 0$, then the sign of β_2 is unclear. If $\phi_2 > 0$, i.e. $\beta_2 < 0$, we are back in Case 1. If, on the other hand, $\phi_2 < 0$, i.e. $\beta_2 > 0$, we rewrite the recursion in equation (43) for $k \geq 3$ as

$$\begin{aligned}
\delta_{12,k} &= \phi_2\delta_{12,k-1} + \delta^{(1)}_{12,k} = \phi_2\delta_{12,k-1} + \phi_1\delta^{(1)}_{12,k-1} \\
&= \phi_2\delta_{12,k-1} + \phi_1(\delta_{12,k-1} - \phi_2\delta_{12,k-2}) \\
&= (\phi_1 + \phi_2)\delta_{12,k-1} - \phi_1\phi_2\delta_{12,k-2} \\
&= \beta_1\delta_{12,k-1} + \beta_2\delta_{12,k-2}.
\end{aligned}$$

Since $\beta_1 > 0$, $\beta_2 > 0$, $\delta_{12,0} = \delta_{12,1} = 0$ and $\delta_{12,2} = \delta^{(1)}_{12,2} \geq 0$ it follows that $\delta_{12,k} \geq 0$ for all k. ■