Structural Volatility Impulse Response Analysis*

Matthias Fengler‡ Jeannine Polivka§

Version: January 31, 2022

Abstract

We make three contributions to the volatility impulse response function (VIRF) of Hafner and Herwartz (2006). Firstly, we derive the law of the VIRF in the BEKK model. Secondly, we present a structural embedding of the VIRF by relying on recent developments for identification in MGARCH models. This broadens the use case of the VIRF, to date limited to historical analyses, by allowing for counterfactual and out-of-sample scenario analyses. Thirdly, we show how to endow the VIRF with a causal interpretation. We illustrate the merits of a structural VIRF analysis by investigating the impacts of historical and counterfactual shock events as well as the consequences of well-defined future shock scenarios on the U.S. equity, government bond and foreign exchange market.

Keywords: causality in volatility, multivariate GARCH models, proxy identification, scenario analysis, structural identification, volatility impulse response functions

*We are grateful for financial support by the Swiss National Science Foundation (SNF Grant No: 176684, Project "Structural Models of Volatility").

‡School of Economics and Political Science, Department of Economics, University of St. Gallen, Bodanstrasse 6, 9000 St. Gallen, Switzerland. Email: matthias.fengler@unisg.ch.

§School of Economics and Political Science, Department of Economics, University of St. Gallen, Bodanstrasse 6, 9000 St. Gallen, Switzerland. Email: jeannine.polivka@unisg.ch.

1 Introduction

The impulse response function (IRF) is the tool of choice to analyze how dynamic multivariate systems respond to shocks. It imparts how an unanticipated perturbation impacts the modelled variables and exhibits how sign, magnitude and persistence of the response evolve over time. In its standard use case in vector autoregressions, the focus of IRF analysis rests on learning about the feedback in the mean process (Inoue and Kilian, 2013; Lütkepohl, 2010). By contrast, in volatility models of high-frequent speculative asset returns, where the mean equation is often mundane, it is the second-order response that is of commanding interest. For example, a mutual fund manager may be concerned about how the variance matrix of certain asset classes may react to an unforeseen monetary policy shock. In this situation, the volatility impulse response function (VIRF) affords her the relevant insights.

The VIRF as first conceptualized by Hafner and Herwartz (2006) extends ideas of the generalized IRF of Koop et al. (1996) to second-order moments. It conditions on past information and an exogenous shock component and traces the nonlinear effects of the shocks on volatility dynamics. Among extant VIRF specifications, the VIRF of Hafner and Herwartz (2006) is especially attractive because it admits a closed-form expression within the BEKK(p, q) model, which is among the most general multivariate GARCH models. In this work, we take a fresh look at the VIRF and add to the extant literature in three ways: Firstly, we derive the asymptotic distribution of the VIRF in the BEKK model, which is most frequently applied in the empirical VIRF literature. More specifically, we show that, like the VIRF, its asymptotic variance matrix can be written as a function of the forecast horizon in a compact recursive form. This allows for an efficient numerical evaluation of confidence intervals, which to date can only be obtained by time-consuming simulation techniques, such as the bootstrap. Secondly, we endow the VIRF with a structural interpretation by relying on recent advances for identification in MGARCH models (Hafner et al., 2020; Fengler and Polivka, 2021). This new interpretation broadens the use cases of the VIRF, which to date has been limited to historical analyses, materially. A structural volatility model featuring identified and labeled structural shocks raises the curtain for proper scenario analyses as is common in traditional structural VAR analyses (Amisano and Giannini, 2012). It allows one, e.g., to define counterfactual scenarios or to obtain insights into the average volatility impact of certain families of well-defined shock scenarios. We coin the term "scenario VIRF" for this novel use case. Our third contribution is to empower the VIRF with a causal interpretation. To this end, we build on recent advances on causal inference in the time series context by Rambachan and Shephard (2020, 2021). This allows us to use the microeconometricians' notion of causality when analyzing scenarios relevant for risk management purposes, e.g., the causal effects of tail events in a specific financial market on the asset return system.

Identifying the impact of structural shocks on financial volatility is crucial for models of asset price dynamics, risk management and portfolio optimization and has been investigated, among others, by Gallant et al. (1993), Lin (1997), Hafner and Herwartz (2006) and more recently by Liu (2018). However, despite this long history, structural advances of volatility impulse response analysis have remained in their infancies. The conditional moment profile developed by Gallant et al. (1993) suffers from difficulties in choosing realistic shocks and a sensible baseline for setting the conditional volatility profile into context. In volatility analysis – in contrast to impulse response analysis for the (conditional) mean – the choice of a baseline shock to returns is non-trivial as it cannot represent the long-run volatility state (Hafner and Herwartz, 2006). Lin (1997) bases his univariate volatility impulse response analysis on reduced-form GARCH models such that the shocks lack structural interpretation. Hafner and Herwartz (2006) tackle the issues of Gallant et al. (1993) by developing a VIRF concept in the spirit of the GIRF of

Koop et al. (1996). However, their shock identification strategy – while not building on a reduced form model approach as in Lin (1997) – relies on statistical concepts. Thus, their modeling framework lacks economic interpretability. Its ambit is therefore limited to retrospective historical analyses with model-implied past shocks. Moreover, the majority of the aforementioned studies do not provide asymptotic properties for statistical inference of their VIRFs. While Liu (2018) embarks upon deriving confidence intervals, his VIRF framework is defined for DCC models, which are generally considered to be less flexible than BEKK models. Furthermore, he applies his VIRF to shocks which are identified statistically via time-varying heteroscedasticity and thus lack a direct economic interpretation.

The remainder of the paper is structured as follows. Section 2 presents the concept of the VIRF including rigorous mathematical derivations, derives its asymptotic distribution and establishes connections to structural and causal modeling. Section 3 shortly presents the structural model we use for identification of the asset return system covering equity, fixed income and foreign exchange markets. It then showcases historical as well as out-of-sample scenario VIRFs for well-defined risk scenarios. Section 4 concludes.

2 Volatility impulse response analysis

2.1 Modeling framework

We consider the system of n speculative (log) returns given by

$$r_t = \mu_t + \varepsilon_t \quad (t \in \mathbb{Z}) \tag{1}$$

where $\mu_t = \mathrm{E}[r_t|\mathcal{F}_{t-1}]$ and $\mathcal{F}_t = \sigma(\{\varepsilon_s : s \leq t\})$ denotes the information available up to time t, generating the filtration $\mathcal{F}$. The n-dimensional innovation vector ε_t is assumed to be square integrable, satisfies $\mathrm{E}[\varepsilon_t|\mathcal{F}_{t-1}] = 0$ ($\mathrm{E}[|\varepsilon_t|] < \infty$) and has the conditional covariance matrix $\mathrm{Var}[\varepsilon_t|\mathcal{F}_{t-1}] = H_t$. The process $(H_t)_{t\in\mathbb{Z}}$ is assumed to be almost surely symmetric, positive definite for all t and covariance stationary and is by definition $\mathcal{F}$-adapted. We denote by $H_t^{1/2} \in \mathbb{R}^{n\times n}$ the principal matrix square root[1] of H_t which exists for all t and is uniquely defined. Furthermore, we define $\tilde{R}$ to be a structural rotation matrix[2] and $\xi = (\xi_t)_{t\in\mathbb{Z}}$ an n-dimensional real-valued white noise process of structural shocks with zero mean and identity covariance matrix, i.e., $\xi_t \sim WN(0, I_n)$.

Definition 2.1. The process $(\varepsilon_t)_{t\in\mathbb{Z}}$ follows a structural multivariate GARCH process if it satisfies:

$$\varepsilon_t = H_t^{1/2}\tilde{R}\xi_t. \tag{2}$$

If $\xi_t \sim SWN(0, I_n)$, $(\varepsilon_t)_{t\in\mathbb{Z}}$ is said to follow a strong structural multivariate GARCH process.

The structural rotation matrix $\tilde{R}$ in (2) endows the model with a structural interpretation and specifies the propagation channels of the structural shocks. For example, choosing $\tilde{R}$ to be the identity matrix implies a symmetric volatility spillover mechanism as the principal matrix square root preserves the positive definiteness as well as the symmetry of H_t. In contrast, if one chooses the rotation which transforms the principal matrix square root into, e.g., a Cholesky decomposition, the volatility spillovers obey a recursive ordering principle imposed through the triangular structure of the decomposition. Although both are popular ad-hoc decompositions, neither specification for $\tilde{R}$ rests on solid eco-

[1] Any real symmetric $(n \times n)$ matrix M can be factorized as $M = \Gamma\Lambda\Gamma^\top$ where Γ is an orthogonal $(n \times n)$ matrix with the normalized eigenvectors of M as columns, and Λ the diagonal matrix of the eigenvalues. The principal matrix square root of M is defined as $\Gamma\Lambda^{1/2}\Gamma^\top$ where $\Lambda^{1/2}$ denotes the diagonal matrix of the square root of the eigenvalues of M. It is the unique matrix square root which has non-negative eigenvalues, see Horn and Johnson (2012, Theorem 7.2.6).

[2] A rotation matrix R is a real $(n \times n)$ matrix satisfying $R^\top R = RR^\top = I_n$ (orthogonality) and $\det(R) = +1$.

nomic grounds for asset return systems. In contrast to this, in structural volatility models one estimates $\tilde{R}$ and thus identifies the economic propagation channels of labeled, interpretable structural shocks to the asset return system, for example backed by external data (Fengler and Polivka, 2021).

2.2 Volatility impulse response functions

In order to assess the impact of a shock ξ_t on volatility given $\mathcal{F}_{t-1}$, Hafner and Herwartz (2006) define the h-step ahead volatility impulse response function (VIRF) as the difference between the expected h-step-ahead covariance conditioning on the shock and past information and the expected h-step-ahead covariance conditioning on past information only. This choice of conditioning sets is in the tradition of the generalized impulse response function (GIRF) developed by Koop et al. (1996) to adress the problems of history, shock, and compositional dependence of impulse responses in non-linear multivariate models.

Definition 2.2. We denote by $\tilde{\mathcal{F}}_t = \sigma(\{\xi_t, \varepsilon_s, s \leq t-1\})$. Let $h \in \mathbb{N}$. The h-step ahead VIRF is given by:

$$\mathrm{V}_{t+h}(\xi_t) := \mathrm{E}[\mathrm{vech}(H_{t+h})|\tilde{\mathcal{F}}_t] - \mathrm{E}[\mathrm{vech}(H_{t+h})|\mathcal{F}_{t-1}]. \tag{3}$$

Here, $\mathrm{vech}(\cdot)$ denotes the operator stacking the lower triangular part of a (symmetric) $(n \times n)$ matrix in an n^*-dimensional vector where $n^* = \frac{n(n+1)}{2}$ such that $\mathrm{V}_{t+h}(\xi_t)$ is an n^*-dimensional vector of impulse responses of the conditional (co-)variances. Note that in spite of making reference to "volatility", the VIRF is in fact a (co-)variance IRF. The definition of the VIRF based on $\tilde{\mathcal{F}}_t$ and $\mathcal{F}_{t-1}$ is motivated by the fact that, in contrast to impulse responses for conditional means, there is no natural baseline for shocks to volatility. In IRFs for conditional means a baseline ε_t^0 corresponding to the long-run mean of the process, is economically natural. A fixed return baseline, however, cannot represent the steady state of volatility, because $\varepsilon_t^0 \varepsilon_t^{0\top}$ has rank one with probability one. Thus, it cannot coincide with the average volatility state. Hence, instead of artificially adding a shock δ to an arbitrarily chosen baseline level of volatility, $\varepsilon_t^* = \varepsilon_t^0 + \delta$, to generate impulse responses as proposed by Gallant et al. (1993), Hafner and Herwartz (2006) consider responses to shocks ξ_t.

The VIRF in (3) is attractive for two further reasons. On the one hand, for the analysis of responses of volatility to historical shocks, it is sufficient to employ the principal square root of H_t, because the VIRF as quadratic form is independent of the underlying structural model.[3] This invariance allows one to quickly and efficiently infer past shocks and to calculate corresponding historical VIRFs by means of (2) and (3) using $\tilde{R} = I_n$ – making any additional computational effort related to structural modeling redundant in the historical set-up. On the other hand however, in scenario analysis, constructing VIRFs based on structural shocks is highly appealing. To begin with, employing a structural model solves the so-called "composition effect" problem (Koop et al., 1996): firstly, in multivariate models, it is a priori not clear how a shock of interest should be chosen as it will often have contemporaneous effects on several outcome variables; secondly, it is unrealistic to observe pertubations in solely one shock while keeping the others fixed because impulse responses generally depend on compositions of shocks. A structural volatility model, similar to structural VAR models, see e.g. (Kilian, 2013), specifies the shock composition mechanism and provides an economically meaningful multivariate shock time series. This solves the composition effect problem and allows one to investigate the impacts of specific future market scenarios defined by the labeled structural shocks. Such

[3] The VIRF traces the effects of shocks on the covariance matrix H_t and not on its structural decomposition $H_t^{1/2}\tilde{R}$. When considering the square of the matrix decomposition, the effect of the structural model vanishes as $\tilde{R}\tilde{R}^\top = 1$.

investigations using scenario VIRFs have not been possible to date, as they critically hinge on the interpretability of the shocks. We stress this aspect further in Section 2.5.

2.3 The VIRF in the BEKK(p, q) model

To derive closed form expressions for structural VIRFs, we need to choose a model for the dynamics of the conditional covariance matrix process. Amongst these, the parametric VEC and BEKK GARCH models enjoy great popularity (Bauwens et al., 2006) and are most frequently employed in the applied VIRF literature (see, e.g., Jin et al., 2012 and Olson et al., 2014). The n-dimensional process $(\varepsilon_t)_{t\in\mathbb{Z}}$ admits a BEKK(p, q) specification (Engle and Kroner, 1995) if H_t satisfies for all $t \in \mathbb{Z}$:

$$H_t = CC^\top + \sum_{i=1}^{p} A_i^\top \varepsilon_{t-i}\varepsilon_{t-i}^\top A_i + \sum_{j=1}^{q} B_j^\top H_{t-j} B_j \quad (p, q \in \mathbb{N}) \tag{4}$$

where C is a lower triangular matrix and A_i and B_j are coefficient matrices in $\mathbb{R}^{n\times n}$. The intercept matrix $CC^\top$ is symmetric and positive semi-definite by construction, and strictly positive definite if C has full rank. The latter property ensures positive definiteness of $(H_t)_{t\in\mathbb{Z}}$. Boussama et al. (2011, Theorem 2.4) show that under weak regularity conditions on $(\xi_t)_{t\in\mathbb{Z}}$, the multivariate BEKK GARCH process is ergodic, strictly and weakly stationary and invertible if the eigenvalues of $\sum_{i=1}^{p} A_i \otimes A_i + \sum_{j=1}^{q} B_j \otimes B_j$ are less than one in modulus. Hafner and Preminger (2009) provide conditions to establish consistency as well as asymptotic normality of the QML estimator, assuming inter alia the existence of second-order moments of $(\xi_t)_{(t\in\mathbb{Z})}$ and finite sixth-order moments of $(\varepsilon_t)_{t\in\mathbb{Z}}$. Due to the quadratic structure of the BEKK(p, q) model, the parameter matrices are only identified up to sign. For the BEKK(1, 1), the model is uniquely identified assuming the diagonal elements of C and the first matrix entries of A_1, $a_{11(1)}$, and B_1, $b_{11(1)}$, to be positive.

For the BEKK(p, q) model, there is a closed-form expression for the h-step ahead VIRF, which can be further simplified for the BEKK(1, 1) case. While Hafner and Herwartz (2006) have stated these results, we provide here a rigorous derivation for the sake of completeness, which will be of additional value for our proofs of the asymptotic properties of the structural VIRF in Section 2.4. Because the BEKK-VIRF of Hafner and Herwartz (2006) does not involve any structural parameters, we furthermore adjust the formulas to allow for structural models of type (2). Let $\text{vec}(\cdot)$ denote the operator stacking the columns of a matrix on top of each other.

Proposition 2.1. *For a BEKK(p, q) representation of H_t in (2) with parameter vector $\eta = \left(\text{vec}(C)^\top, \text{vec}(A_i)^\top, \text{vec}(B_j)^\top\right)^\top$, $(i = 1, \dots, p;\ j = 1, \dots, q;\ p, q \in \mathbb{N})$ and with VMA(∞) representation with coefficients $(\Psi_i)_{i\in\mathbb{N}}$ given in Proposition A.1, the h-step ahead VIRF given a structural shock ξ_t is given by*

$$\mathrm{V}_{t+h}(\xi_t, \eta) = \Psi_h\, D_n^+ \left(H_t^{1/2} \otimes H_t^{1/2}\right) D_n \left(\text{vech}(\tilde{R}\xi_t\xi_t^\top \tilde{R}^\top - I_n)\right) \tag{5}$$

where D_n denotes the duplication matrix (see Equation (20)) and D_n^+ its Moore-Penrose inverse.

Proof. See Proof (2.1) in Appendix A.2. □

Notably, the impulse response function is a nonlinear, but even function of the structural shock. As the conditional volatility at the time of the shock occurrence enters the VIRF, it is an $\tilde{\mathcal{F}}_t$-adapted process. The persistence of a shock to volatility is governed by the moving average matrices Ψ_h.

Proposition 2.2. *For the BEKK$(1, 1)$ model with parameter vector η and with VMA(∞) coefficients $\Psi_0 = I_{n^*}$, $\Psi_1 = \tilde{A}_1$ and $\Psi_i = \left(\tilde{A}_1 + \tilde{B}_1\right)\Psi_{i-1} = \left(\tilde{A}_1 + \tilde{B}_1\right)^{i-1} \tilde{A}_1$ $(i \geq 2)$, the VIRF given an initial shock ξ_t and a corresponding variance level H_t reduces to the recursion:*

$$\begin{aligned} &\mathrm{V}_{t+h}(\xi_t, \eta) = \left(\tilde{A}_1 + \tilde{B}_1\right)^{h-1} \tilde{A}_1 D_n^+ \left(H_t^{1/2} \otimes H_t^{1/2}\right) D_n\, \text{vech}(\tilde{R}\xi_t\xi_t^\top \tilde{R}^\top - I_n) \quad (h \geq 1) \\ \Rightarrow\ &\mathrm{V}_{t+h}(\xi_t, \eta) = \left(\tilde{A}_1 + \tilde{B}_1\right) V_{t+h-1}(\xi_t) \quad (h \geq 2). \end{aligned} \tag{6}$$

Proof. For a derivation of the vech form of the BEKK(1, 1) model see Proposition A.1, Equation (22). Inserting the VMA(∞) coefficients in (5) yields the claim. □

2.4 Inference for the VIRF

2.4.1 Consistency of the VIRF

Proposition 2.3. *Under the regularity conditions of Hafner and Preminger (2009) and if the VIRF is a continuous function of the parameter vector η of the MGARCH model, $V_{t+h}(\xi_t, \eta)$ can be estimated consistently.*

$$\hat{\mathrm{V}}_{t+h}(\xi_t, \eta) \xrightarrow{p} \mathrm{V}_{t+h}(\xi_t, \eta) \quad (h \in \mathbb{N}) \tag{7}$$

Proof. Given that the VIRF is a continuous function of the parameter vector of the underlying MGARCH model η, the result follows by the continuous mapping theorem. □

Remark. When modeling the dynamics of the conditional covariance matrix by a BEKK(p, q) model, this model is continuous in the parameter vector η.

2.4.2 Asymptotics of the BEKK-VIRF

Theorem 1. *Under the regularity conditions of Hafner and Preminger (2009) and given continuous differentiability of the BEKK-VIRF as function of $\eta \in \mathbb{R}^m$, $\mathrm{V}_{t+h}(\xi_t, \eta)$ is asymptotically normally distributed with:*

$$\sqrt{T}\left(\hat{\mathrm{V}}_{t+h}(\xi_t, \eta) - \mathrm{V}_{t+h}(\xi_t, \eta)\right) \xrightarrow{d} N(0, \boldsymbol{V}_\eta(\mathrm{E}[\boldsymbol{\mathcal{H}}(\eta)])^{-1}\boldsymbol{\mathcal{I}}(\mathrm{E}[\boldsymbol{\mathcal{H}}(\eta)])^{-1}\boldsymbol{V}_\eta^\top) \tag{8}$$

where $\boldsymbol{V}_\eta = \frac{\partial \mathrm{V}_{t+h}(\xi_t,\eta)}{\partial \eta^\top}$ denotes the $n^* \times m$ Jacobian matrix of the VIRF with respect to the m-dimensional parameter vector η, $\boldsymbol{\mathcal{H}}(\eta) = \frac{\partial^2 \log(l_t(\eta))}{\partial \eta \partial \eta^\top}$ the Hessian matrix of the log likelihood function $\log(l_t)$ and $\boldsymbol{\mathcal{I}} = \mathrm{E}\left[\left(\frac{\partial \log(l_t(\eta))}{\partial \eta}\right)\left(\frac{\partial \log(l_t(\eta))}{\partial \eta^\top}\right)\right]$ the Fisher information matrix.

Proof. The result follows by an application of the Delta method together with the asymptotic normality of the QML estimator and Proposition 2.4. □

Without the Jacobian, the previous result is of little use. We therefore derive:

Proposition 2.4. *Let $h \in \mathbb{N}$. The $(n^* \times m)$ Jacobian of the VIRF with respect to $\eta \in \mathbb{R}^m$ which appears in the asymptotic distribution of the VIRF under the BEKK(p,q) model is given by*

$$\frac{\partial \mathrm{V}_{t+h}(\xi_t,\eta)}{\partial \eta^\top} = \left(\mathrm{V}_t^\top \otimes I_{\frac{n(n+1)}{2}}\right)\frac{\mathrm{vec}(\Psi_h)}{\partial \eta^\top} + \left(I_{\frac{n(n+1)}{2}} \otimes \Psi_h\right)\frac{\partial \mathrm{V}_t}{\partial \eta^\top} \tag{9}$$

where

$$\begin{aligned}\frac{\partial \mathrm{V}_t(\xi_t,\eta)}{\partial \eta^\top} = D^+ \Bigg[&\left(\left(H_t^{1/2}\tilde{R}\xi_t\xi_t^\top\tilde{R}^\top \otimes I_n\right) + \left(I_n \otimes H_t^{1/2}\tilde{R}\xi_t\xi_t^\top\tilde{R}^\top\right)\right) \\ &\times \left[\left(H_t^{1/2} \otimes I_n\right) + \left(I_n \otimes H_t^{1/2}\right)\right]^{-1} - I_{n^2}\Bigg]\frac{\partial\,\mathrm{vec}(H_t)}{\partial \eta^\top}\end{aligned} \tag{10}$$

and $(\Psi_i)_{i\in\mathbb{N}}$ denote the coefficients of the VMA(∞) representation of the BEKK(p,q) model. For the BEKK$(1,1)$ model the expression

$$\frac{\partial \mathrm{V}_{t+h}(\xi_t,\eta)}{\partial \eta^\top} = \left(\mathrm{V}_{t+h-1}^\top \otimes I_{n^*}\right)\frac{\partial\,\mathrm{vec}(\tilde{A} + \tilde{B}\mathbb{1}_{\{h>1\}})}{\partial \eta^\top} + \left(\tilde{A} + \tilde{B}\mathbb{1}_{\{h>1\}}\right)\frac{\partial \mathrm{V}_{t+h-1}}{\partial \eta^\top} \tag{11}$$

is available where $\mathbb{1}_{\{\bullet\}}$ denotes the indicator function being one when the condition in the subscript is satisfied and zero otherwise.

Proof. See Proof (2.4) in Appendix A.3. □

The asymptotic distribution of the structural VIRF estimator allows us to construct simultaneous confidence intervals for impulse responses of individual assets returns to structural shocks; see, e.g., Sims and Zha (1999) and Lütkepohl et al. (2015) for discussions of confidence interval construction for impulse responses. For our application, we choose a large sample approximation of Hotelling's T^2 to compute simultaneous confidence intervals for all components of the VIRF for a given significance level α. This allows us to assess the statistical significance of the responses of the (co-)variances of all assets in our speculative return system to a structural shock.

2.5 Structural VIRFs

In unidentified MGARCH models, volatility impulse response analysis is limited to a retrospective inspection of the impact of past financial, economic or political events. We accordingly refer to this VIRF as historical VIRF. By taking the estimated residual $\hat{\varepsilon}_t$ and volatility state $\hat{H}_t$ at time of the shock occurrence, Hafner and Herwartz (2006) define their structural shocks $\hat{\xi}_t = \hat{H}_t^{-1/2}\hat{\varepsilon}_t$ by means of the principal square root. However, Hafner et al. (2020) and Fengler and Polivka (2021) provide evidence that the principal matrix square root and its associated shocks do not comply with empirical results on asymmetry of volatility spillovers and suggest alternative matrix decompositions for H_t.

Nonetheless, the VIRF is unaffected by this identification problem because it is invariant to rotations in the historical context such that the results are independent of the structural model. To see this more clearly, let $Q_t = H_t^{1/2}\tilde{R}$ denote a structural decomposition of H_t. Given the return realized on the day of interest t, the h-step ahead VIRF is given by:

$$\begin{aligned} V_{t+h}(\xi_t) &= \Psi_h \left(\text{vech}(Q_t(\xi_t\xi_t^\top - I_k)Q_t^\top) \right) \\ &= \Psi_h \left(\text{vech}\left(H_t^{1/2}\tilde{R}\left(\tilde{R}^\top H_t^{-1/2}\varepsilon_t\varepsilon_t^\top H_t^{-1/2}\tilde{R} - I_k \right) \tilde{R}^\top H_t^{1/2} \right) \right) \\ &= \Psi_h \, \text{vech}(\varepsilon_t\varepsilon_t^\top - H_t) \end{aligned} \tag{12}$$

Thus, when ε_t is known, at least in retrospect, this expression is independent of the structural mechanism decoded by $\tilde{R}$.

However, recent contributions to the literature on structural volatility models open up new usage possibilities for VIRFs. Firstly, in out of sample settings, different structural decompositions unveil their full power as they allow to process responses to synthetic but economically meaningful interpretable structural shocks. Secondly, by investigating the outcomes of labeled shocks, structural VIRFs allow one to conduct counterfactual analyses based on defining specific scenarios. The profit of a structural approach is thus the full economic interpretability of the input shocks and correspondingly the full interpretability of the structural VIRF output. We coin the term "scenario VIRF" for this new application framework. We illustrate the merits of using a structural model in the computation of VIRFs in the empirical application in Section 3.

2.6 Causality

Our volatility impulse response analysis connects as well to recent advances on causality in times series. Rambachan and Shephard (2020, 2021) show that the generalized impulse response function of Koop et al. (1996) is, under certain assumptions, endowed with a causal interpretation. In contrast to the GIRF, the VIRF does not trace the dynamic effects of a shock not on the conditional mean but on the conditional covariance matrix. It does not have a causal content a priori, as it is simply the difference of two conditional expectations. In contrast, a causal effect traces the effects of changes in treatments. We show that the VIRF can be endowed with a causal interpretation as well. To this end, we adapt assumptions postulated by Rambachan and Shephard (2020, 2021) to model (2). To define subsequences of our entire time series, we adopt the time series path notation which indicates the start and end points of a series in time by subscripts separated by a colon. We start by rephrasing our variables of interest using terms stemming from causal inference:

Definition 2.3. Let $\xi := (\xi_t)_{t \in I}$, $(I = \{1, \dots, T\})$, denote the stochastic treatment path with realizations $\bar{\xi}_t \in \mathcal{W} \subseteq \mathbb{R}^n$ and let the potential outcome path for any deterministic trajectory $\bar{\xi} := (\bar{\xi}_t)_{t \in I}$ be given by $X := X_t(\bar{\xi})_{t \in I} = (X_1(\bar{\xi}), X_2(\bar{\xi}), \dots, X_T(\bar{\xi}))$ where $X_t := \text{vech}(\varepsilon_t \varepsilon_t^\top)$ and define $\widetilde{\mathcal{F}}_t := \sigma(\{X_s, \xi_s \ (s \leq t)\})$.

The demeaned return vectors respectively their outer products are the observable, continously valued, multidimensional outcomes for which the VMA(∞) representation of the BEKK model is available. Our treatment variable features continuous parameter values as opposed to the classical setting with a binary or finitely discrete treatment variable. We observe only one realization of the stochastic treatment path. Note that $\widetilde{\mathcal{F}}_t \subseteq \sigma(\varepsilon_s,\, \xi_s, s \leq t\})$ as the outer product of two random vectors is a Borel-measurable function. Due to the predictability of H_t, the measurability of the principal matrix square root operator and the fact that $\tilde{R}$ is known it even holds that $\widetilde{\mathcal{F}}_t = \tilde{\mathcal{F}}_t$. Let furthermore ξ and X satisfy the following assumptions:

Assumption 2.1 (Time series non-interference)**.** For each $t \in I$ and all deterministic $(\bar{\xi}_t)_{t\in I}$, $(\bar{\xi}'_t)_{t\in I}$ with $\bar{\xi}_t, \bar{\xi}'_t \in \mathcal{W}$:

$$X_t(\bar{\xi}_{1:t}, \bar{\xi}_{t+1:T}) = X_t(\bar{\xi}_{1:t}, \bar{\xi}'_{t+1:T}) \text{ almost surely.}$$

Assumption 2.1 allows the potential outcomes to depend on past and contemporaneous treatments but excludes dependence on future treatments. It links the potential outcomes and treatments in such a way that $X_t(\bar{\xi})_{t\in I}{=}(X_1(\xi_1), X_2(\xi_{1:2}), \ldots, X_T(\xi_{1:T}))^\top$ and acts as the time series analogon of SUTVA (Cox, 1958; Rubin, 1980). Our outcome series satisfies Assumption 2.1, which follows from the definition $\varepsilon_t = Q_t\xi_t$, where $Q_t = H_t^{1/2}\tilde{R}$ is $\widetilde{\mathcal{F}}_{t-1}$-measurable, and the fact that the sequence of structural shocks ξ is assumed to be white noise.

Assumption 2.2 (Time series unconfoundedness)**.** For each $t \in I$ and all $h > 0$:

$$\xi_t \perp\!\!\!\perp (\xi_{t+1:t+h}, \{X_{t+h}(\bar{\xi}_{1:t-1}, (\bar{\xi}_s)_{t\leq s\leq t+h}) : \bar{\xi}_s \in \mathcal{W}\}) \,|\widetilde{\mathcal{F}}_{t-1}$$

Assumption 2.2 encompasses non-anticipating treatment paths conditional on the information available up to time $t-1$. With ξ_t being iid, Assumption 2.2 is fulfilled due to the serial independence of the structural shocks and the predictability of the conditional covariance matrix in the MGARCH model given the past information contained in the returns. This guarantees the independence of the treatment ξ_t of future treatments and the associated potential outcomes. Francq and Zakoïan (2010, Thm 11.5) provide the conditions for which vector GARCH models admit a strictly stationary and non-anticipative solution.[4]

Assumption 2.3 (Positivity/Sequential overlap)**.** Let $\mathcal{T} \subset \mathcal{W}$ be a Borel set with positive measure. For each $t \in \mathbb{Z}$ the stochastic treatment path satisfies

$$0 < \Pr(\xi_t \in \mathcal{T}) < 1$$

almost surely for all $\mathcal{T} \subset \mathcal{W}$.

Assumption 2.3 mirrors the overlap assumption of cross-sectional causality analyses in the time series setting and is essential for the proof steps of Corollary 1.1 to be well-defined. As we are handling a continuous treatment variable, the overlap assumption is stated for Borel sets which can be chosen as ϵ-neighbourhoods of the structural shock of interest $\bar{\xi}_t$: $\mathcal{T} = \{\tilde{\xi} \in \mathcal{W} : d(\bar{\xi}_t, \tilde{\xi}) < \epsilon\}$ for some metric d on $\mathbb{R}^n$ and $\epsilon > 0$, which can be arbitrarily small. In empirical practice, the set-based definition aligns well with the concept of scenario and sensitivity analysis. As the BEKK model and thus the BEKK-VIRF are continuous in the structural shock, applying them to a set $\mathcal{T}_t$ results in a connected Borel set of conditional covariance matrices $\mathcal{H}_t$, outer return products $\mathcal{X}_t$ and VIRF vectors.

Based on these assumptions, we can now show that the VIRF applied to an ϵ-neighbourhood $\mathcal{T}_t$ of a structural shock of interest $\bar{\xi}_t$ can be decomposed into a filtered treatment effect, i.e. a filtered causal volatility impulse response, and a selection bias term which vanishes under time series unconfoundedness.

[4] A non-anticipative solution is defined as a process $(\varepsilon_t)_{t\in\mathbb{Z}}$ such that ε_t is a measurable function of ξ_{t-s} $(s \geq 0)$ with $H_t^{1/2} \perp\!\!\!\perp \sigma(\xi_{t+h}, h \geq 0)$ and $\varepsilon_t \perp\!\!\!\perp \sigma(\xi_{t+h}, h > 0)$.

Corollary 1.1. *Let Assumptions 2.1 – 2.3 hold, let $h \geq 0$ and assume that for any deterministic $\bar{\xi}_t \in \mathcal{W}$ and any $\tilde{\xi}_t \in \mathcal{T} = \{\tilde{\xi} \in \mathcal{W} : d(\bar{\xi}_t, \tilde{\xi}) < \epsilon\}$ for some metric d on $\mathbb{R}^n$ and $\epsilon > 0$: $\mathrm{E}[X_{t+h}(\tilde{\xi}_t) - X_{t+h}|\tilde{\mathcal{F}}_{t-1}] < \infty$. Then it holds:*

$$\mathrm{V}_{t+h}(\mathcal{T}_t) = \mathrm{E}[\mathcal{X}_{t+h} - X_{t+h}|\tilde{\mathcal{F}}_{t-1}] + \Delta_{t+h}(\mathcal{T}_t|\tilde{\mathcal{F}}_{t-1})$$

where $\Delta_{t+h}(\mathcal{T}_t|\tilde{\mathcal{F}}_{t-1}) = \frac{\mathrm{cov}\left[\mathrm{vech}(\mathcal{X}_{t+h}), \mathbb{1}_{\{\xi_t \in \mathcal{T}_t\}}|\mathcal{F}_{t-1}\right]}{\mathrm{E}\left[\mathbb{1}_{\{\xi_t \in \mathcal{T}_t\}}\right]}$ is a selection bias which vanishes under Assumption 2.2.

Proof.

$$\begin{aligned}
\mathrm{V}_{t+h}(\mathcal{T}_t) &= \mathrm{E}[\mathrm{vech}(H_{t+h}(\mathcal{T}_t))|\tilde{\mathcal{F}}_t] - \mathrm{E}[\mathrm{vech}(H_{t+h})|\mathcal{F}_{t-1}] \\
&= \mathrm{E}[\mathrm{vech}(\mathcal{H}_{t+h})|\mathcal{F}_{t-1}, \mathcal{T}_t] - \mathrm{E}[\mathrm{vech}(H_{t+h})|\mathcal{F}_{t-1}] \\
&= \mathrm{E}[\mathrm{vech}(\mathcal{X}_{t+h})|\mathcal{F}_{t-1}, \mathcal{T}_t] - \mathrm{E}[\mathrm{vech}(X_{t+h})|\mathcal{F}_{t-1}] \\
&= \frac{\mathrm{E}[\mathrm{vech}(X_{t+h}(\mathcal{T}_t))\mathbb{1}_{\{\xi_t \in \mathcal{T}_t\}}|\mathcal{F}_{t-1}]}{\mathrm{E}\left[\mathbb{1}_{\{\xi_t \in \mathcal{T}_t\}}|\mathcal{F}_{t-1}\right]} - \mathrm{E}[\mathrm{vech}(X_{t+h})|\mathcal{F}_{t-1}] \\
&= \frac{\mathrm{E}\left[\mathrm{vech}(X_{t+h}(\mathcal{T}_t)|\mathcal{F}_{t-1}\right]\mathrm{E}\left[\mathbb{1}_{\{\xi_t \in \mathcal{T}_t\}}|\mathcal{F}_{t-1}\right] + \mathrm{cov}(\mathrm{vech}(X_{t+h}(\mathcal{T}_t)), \mathbb{1}_{\{\xi_t \in \mathcal{T}_t\}}|\mathcal{F}_{t-1})}{\mathrm{E}\left[\mathbb{1}_{\{\xi_t \in \mathcal{T}_t\}}|\mathcal{F}_{t-1}\right]} \\
&\quad - \mathrm{E}[\mathrm{vech}(X_{t+h})|\mathcal{F}_{t-1}] \\
&= \mathrm{E}[\mathrm{vech}(X_{t+h}(\mathcal{T}_t)) - \mathrm{vech}(X_{t+h})|\mathcal{F}_{t-1}] + \underbrace{\frac{\mathrm{cov}(\mathrm{vech}(X_{t+h}(\mathcal{T}_t)), \mathbb{1}_{\{\xi_t \in \mathcal{T}_t\}}|\mathcal{F}_{t-1})}{\mathrm{E}[\mathbb{1}_{\{\xi_t \in \mathcal{T}_t\}}]}}_{:=\Delta_{t+h}(\mathcal{T}_t|\tilde{\mathcal{F}}_{t-1})}
\end{aligned}$$

and under Assumption 2.2 which asserts that the contemporaneous treatment ξ_t is jointly independent of all future treatments and potential outcomes: $\Delta_{t+h}(\mathcal{T}_t|\tilde{\mathcal{F}}_{t-1}) = 0$. Hence, the VIRF identifies the filtered impulse causal effect. □

Remark. The filtered treatment effect resembles the causal response function defined by Rambachan and Shephard (2020) which compares the effects of two finitely discrete treatments $\bar{\xi}_t$ and $\bar{\xi}_t^*$ given past information. In the VIRF, the occurence of the shock $\bar{\xi}_t$ given past information is however contrasted with the model results given only the past information instead of conditioning on an artifical base shock $\bar{\xi}_t^*$. By integrating out the effect of the shock $\bar{\xi}^*$, the integrated causal response function coincides with the VIRF. This gives a causal meaning to the VIRF as vectorized integrated causal response function applied to the outer product of returns.

Thus, the VIRF allows us not only to investigate the effects of structural shocks, but to assess the causal impacts of these interpretable structural "treatment" shocks on volatility.

3 Empirical Application

We illustrate the merits of the structural VIRF approach by analysing a system of speculative daily asset returns covering three important asset classes for portfolio optimization and key ingredients in financial systemic stress analysis by the ECB (Kremer et al., 2012): equity, fixed income and the foreign exchange markets. To obtain an economically directly interpretable structural model, we employ the structural proxy-MGARCH approach of Fengler and Polivka (2021) from which we identify labeled structural shocks for our volatility impulse response analyses.

3.1 Structural volatility model

The structural proxy-MGARCH model of Fengler and Polivka (2021) derives the structural rotation matrix by means of a proxy variable scheme; see also Stock and Watson (2012) and Mertens and Ravn (2013). Assume there exists a centered $(n-1)$-dimensional instrument process $Z = (Z_t)_{t \in I}$ such that, for all $i = 1, \ldots, n-1$,

$$\mathrm{E}[\xi_{it} Z_{it}] = \phi_i \in \mathbb{R} \setminus \{0\} \qquad \text{(relevance)} \qquad (13)$$

$$\mathrm{E}[\xi_t^{i*} Z_{it}] = \mathbf{0}_{(n-1)\times 1} \qquad \text{(exogeneity)} \qquad (14)$$

where the product process $(\xi_t Z_{it})_{(t=1,\dots,T)}$ is weakly stationary and the index i relates to the i-th vector element, the superscript $i*$ denotes all vector elements apart from the i-th element. Based on these assumptions, the columns of the rotation matrix are given by

$$\tilde{R}_{\bullet i} = \pm \mathrm{E}[u_t Z_{it}] \left(\mathrm{E}[Z_{it} u_t^\top] \, \mathrm{E}[u_t Z_{it}] \right)^{-1/2} \tag{15}$$

and one can estimate the full rotation matrix by means of a recurrence scheme based on the chaining of Givens rotations (Fengler and Polivka, 2021). Formally, the first k columns of the rotation matrix can be expressed analytically without involving the rotation angles using the Hadamard product $\odot$ by

$$\tilde{R}_{n\times k} = \mathrm{E}[u_t Z_t^\top] \left[I_{n\times k} \cdot \left(\mathrm{E}[u_t^\top \odot Z_t] \, \mathrm{E}[Z_t^\top \odot u_t] \odot I_k \right)^{-1/2} \right] \tag{16}$$

where the last column is given up to sign by the orthonormality property of the rotation matrix. Consistent estimation of the rotation matrix and asymptotic inference is possible by replacing expectations with their sample mean analogues.

We borrow the empirical analysis from Fengler and Polivka (2021) and study daily price data ranging from 1/1/1998 to 12/31/2014 taken from Bloomberg. Our asset triple consists of the S&P 500 Composite Index (SP500), the yield of the U.S. constant maturity 10 year treasury note (FRTCM10) and the Finex U.S. Dollar Index (NDXCS00). We compute daily log returns r_t for each asset; see Figure 1.

Typical proxy variables employed for identification are of narrative nature, for example records of monetary policy interventions, such that the resulting structural shocks $\xi_t = \tilde{R}^\top H_t^{-1/2} \varepsilon_t$ are endowed with a direct economic interpretation. For our high frequent asset triple we use news data taken from Thomson Reuters MarketPsych Indices (TRMI) to proxy for the underlying structural shocks. The TRMIs are available on a daily level. As in Fengler and Polivka (2021), we choose the U.S. stock index sentiment and the U.S. bond sentiment and distill the unexpected innovations to the proxy series by fitting

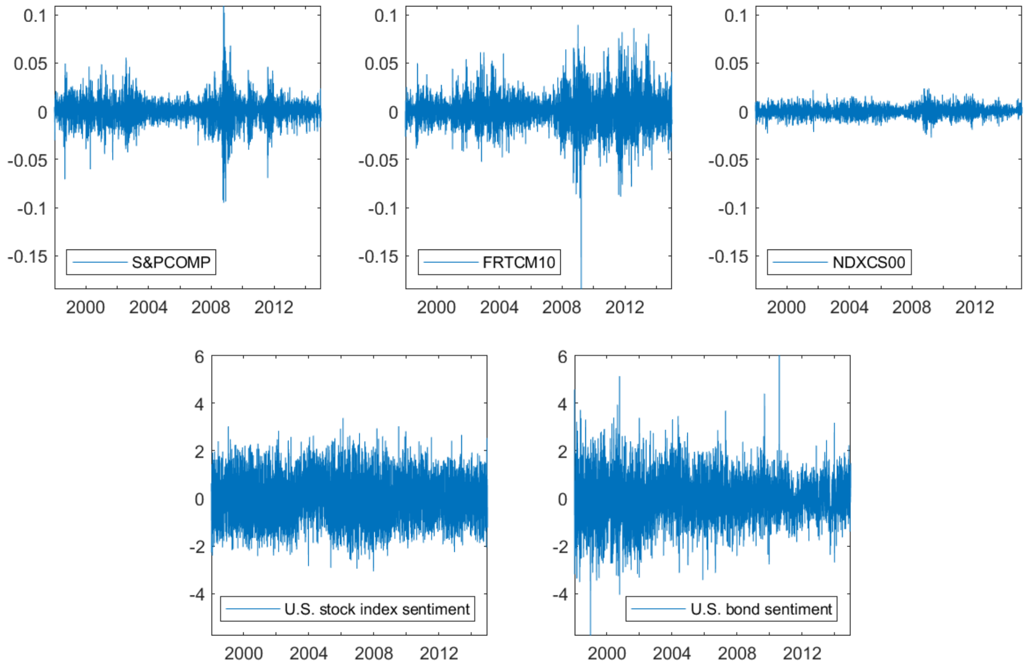

Figure 1: Demeaned daily log returns of the S&P 500 Composite Index (SP500), the yield of the U.S. constant maturity 10 year treasury note (FRTCM10) and the Finex U.S. Dollar Index (NDXCS00) from 1/1/1998 to 12/31/2014.

flexible ARMA models to identify an equity market and an unconventional monetary policy shock. As bond yields move in opposite direction to bond prices, a structural shock identified by means of the bond market sentiment is connected to the treasury yield with reversed sign.

Following Fengler and Polivka (2021), we estimate the structural model using a BEKK(1, 1) specification. The results of Table 1 show that the rotation angles of the structural rotation differ from zero. This suggests a deviation from the volatility spillover symmetry imposed by the spectral decomposition $H_t^{1/2}$ obtained when $\tilde{R} = I_n$. This can be nicely seen, as the structural model shifts mass from the unit diagonal entries of the identity matrix of no rotation in an asymmetric fashion to the off-diagonal matrix elements of the rotation matrix in Table 1. Furthermore, the third section of Table 1 shows that the stock market index Z_1 and bond market sentiment Z_2 seem to be relevant instruments. Fengler and Polivka (2021) conduct a narrative corroboration of the first percentiles of the structural shocks with the major financial market news issued on the day of the shock occurrence. They can connect each structural shock to specific economic and financial turmoil events reported in the news and identify ξ_1 as equity shock, ξ_2 as unconventional monetary policy shock and refer to ξ_3 as currency shock.

		proxy-MGARCH model		
$(\hat{\theta}_{12}, \hat{\theta}_{13}, \hat{\theta}_{23})^\top$		0.3811	−0.1885	−2.9164
$\hat{\hat{R}}$		0.9118	0.3238	−0.2526
		0.3654	−0.9204	0.1393
		−0.1874	−0.2193	−0.9575
		ξ_1	ξ_2	ξ_3
correlations	Z_1	0.3347	0.0000	−0.0000
	Z_2	0.0052	0.1936	0.0000
p-values	Z_1	0	1.0000	1.0000
	Z_2	0.7286	0	1.0000
Wald test	distribution	statistic	critical value	p-value
symmetric spillovers	$\chi^2_{(4)}$	107.85	9.4877	0.0000

Table 1: Estimation results of the structural MGARCH model of the demeaned daily log returns of the S&P 500 Composite Index (SP500), the yield of the U.S. constant maturity 10 year treasury note (FRTCM10) and the Finex U.S. Dollar Index (NDXCS00) from 1/1/1998 to 12/31/2014 when using the stock market index (Z_1) and bond market sentiment (Z_2) TRMIs as proxy variables. The table shows, from top to bottom, the estimated rotation angles, the estimated rotation matrix and the correlations of the proxies with the inferred structural shocks ξ_1, ξ_2 and ξ_3 including Wald test for symmetry of volatility spillovers.

3.2 Structural volatility impulse response analysis

As Bauwens et al. (2006) state, one of the most important applications of MGARCH models is the analysis of responses of volatility to shocks. With identified and labeled structural shocks at hand, we turn to an analysis of the structural VIRFs implied by our model starting with historical VIRFs followed by scenario VIRFs.

3.2.1 Historical VIRFs

We showcase three economic key events in our sample: Firstly, as an example of a structural 1% equity tail shock, we consider Freddie Mac's announcement to stop buying the most risky subprime mortgages and related securities on February 27, 2007. Secondly, we consider the bankruptcy of Lehman Brothers on September 15, 2008 as an example of an 1% equity and bond market tail shock. Thirdly, we examine the Fed FOMC statement to buy treasury securities worth 300B on March 18, 2009 which marked off the start of the first quantitative easing period after the financial crisis as example for an unconventional bond market shock. The corresponding shock vectors are documented in Table 2 and the resulting VIRFs are displayed in Figures 2 to 4. Note that the VIRF is of conditional nature, i.e. the impact of the structural shocks depends on the volatility state at the time of the shock occurrence. To compare the impacts of the three shocks directly, one has to conduct a counterfactual analysis.

The structural shock of the Freddie Mac announcement (see Figure 2) causes a statistically significant and strong response of the variance of the S&P 500 as well as of its covariances with the other assets. Even though the bond market shock is only about one standard deviation, there is as well a strong reaction of the 10 year treasury yield volatility. The conditional variance of the Finex U.S. Dollar Index increases as well, assuming a humped shape over the following months. Its conditional covariances with both, the S&P 500 and

	Event dates		
Structural shock components	02/27/2007	09/15/2008	03/18/2009
Equity	-7.2849	-3.7002	0.2395
Bond	1.0465	3.5972	7.3030
Currency	2.4019	0.1334	1.8027

Table 2: Structural shock vectors selected for historical VIRF analysis from the 1% quantiles with respect to each shock component identified by the proxy MGARCH model.

the treasury yield, increase in response to the shock with a short-lived effect with regard to the S&P 500 and a longer-lasting impact with regard to the conditional covariance with the treasury yield. This reflects that the near-term implications of this announcement may be negative for the U.S. Dollar due to possible flight from the currency and lack of confidence in the U.S. economy. Overall, the structural model indicates that the Freddie Mac announcement increases conditional variances and covariances in all market segments which aligns well with findings on tail events (Longin and Solnik, 1995).

Moving to the impact of the Lehman Brothers bankruptcy (see Figure 3), the structural shock causes an immediate sharp increase in the conditional variances of the S&P 500 and even more so in the 10 year treasury yield as well as in their conditional covariance. The latter corroborates the correct identification of the shock. The impacts are statistically significant and appear to be very persistent. In contrast, the bankruptcy appears to have a short-lasting but level reducing impact on the conditional variance of the Finex U.S. Dollar Index as the VIRF dies out quickly to zero from below. In the conditional (co-) variances we observe an immediate decline in volatility as well. This reduction is most pronounced and persistent for the conditional covariance with the S&P 500 and extremely sharp and short-lived for the conditional covariance with the U.S. 10 year treasury yield. The covariance VIRF bounces back to statistical insignificance after one year but shows a persistent covariance level reduction in the long term again.

The structural shock of the FOMC announcement of the first round of quantitative easing (see Figure 4) causes a very strong and persistent increase in the conditional variance of the treasury yield. While the first impact on the conditional covariance with the S&P 500 and the variance of the S&P 500 is negative – the pledge initially has calming effects on the market, even though this reaction is not statistically different from zero at the 5% level – we observe a delayed increase in humped shape form demonstrating the increasing nervousness of the markets in the long term. Furthermore, we observe a comparably strong but very short lived increase in the level of the conditional variance of the Finex U.S. Dollar Index accompanied by an equally strong increase in its conditional covariance with the treasury yield which, however, subsides quickly. These results are backed by general findings of higher foreign exchange volatility upon FOMC news (Mueller et al., 2017).

On a general note, the VIRFs of our structural model show long-lasting impacts of the structural shocks to the system for the S&P 500 and the treasury yield, whereas the volatility impulse responses of the (co-)variances of the Finex U.S. Dollar Index are shorter-lived. This finding aligns well with empirical findings in the finance literature suggesting that volatility spillover effects between equity and foreign exchange markets are very small in normal market times and only show effects in time periods preceding crises, see, e.g., Grobys (2015) or Cenedese and Mallucci (2016).

3.2.2 Scenario VIRFs

Historical VIRFs are useful for understanding volatility events in hindsight, but beyond that are of limited value because a single event, such as the Lehman Brothers' default, will never occur again. With the help of a structural model, however, meaningful shock scenarios – in an out-of-sample context or as counterfactuals – can be considered. A portfolio manager might, for instance, be interested in understanding the instant volatility impact on her portfolio resulting from certain tail events of FOMC decision days. A magnifold of other scenarios are conceivable. For illustration, we here adopt a risk manager's perspective and investigate the VIRFs of two scenario families associated with the 1% tail events in the equity and monetary policy shocks on the out-of-sample date 01/02/2015. The median VIRFs and 25 and 75% quantile VIRFs of the corresponding scenario families are plotted in Figures 5 and 6.

Figure 5 shows a pronounced positive median impact of the historical 1% quantile equity shocks on the conditional (co-)variance of the S&P 500 and the treasury yield on the first trading day of 2015. The increase in conditional (co-)variances is persistent and decreases only slowly over time. The covariance impulse responses of the Finex U.S. Dollar index to the equity tail event scenario family do not exhibit a clear sign and magnitude and indicate a wide spectrum of possible outcome paths. The median variance impulse of the index itself, however, shows a slight positive impact indicating a volatility increase over the medium term. In contrast, the median impact of the historical 1% quantile monetary policy shocks in Figure 6 has a small effect on the conditional covariance of the treasury yield and the S&P 500 and almost none on the conditional variance of the S&P 500. Remarkably, there is a strong positive effect on the conditional covariance of the Finex U.S. Dollar index and the treasury yield which displays a long persistence. Figures 5 and 6 hence give, dependent on the scenario materializing on the upcoming day, very different indications which (co-)variance levels in the asset return system can be expected to rise and to which extent – thus allowing for different risk managerial precautions.

4 Conclusion

In this paper, we have taken a new look at the volatility impulse response function (VIRF) developed by Hafner and Herwartz (2006) which is a handy device to analyze the impact of shocks on conditional variance matrices of MGARCH models. By deriving the asymptotic law of the VIRF in the BEKK model, we top up VIRF plots with analytical confidence intervals to assess the statistical significance of responses of volatilities to shocks. We show that the asymptotic variance matrix can, like the VIRF, be written as a function of the forecast horizon in a compact recursive form, which allows for an efficient numerical evaluation of confidence intervals. Building on recent advances for identification in MGARCH models, we extend the VIRF to benefit from the advantages of structural volatility models: interpretable, labeled shocks and specified structural propagation channels allow us to broaden the use case of the VIRF, to date limited to historical analyses, to counterfactual and out-of-sample scenario analyses. Moving even beyond a structural interpretation, we show how to endow the VIRF with a causal interpretation which allows one to use the microeconometricians' notion of causality when analyzing the impact of well-defined shock scenarios. In an empirical illustration to an identified system of equity, government bond and foreign exchange returns we demonstrate the abundance of use cases of the structural VIRF in historical and scenario analyses. For example, we illustrate how structural VIRFs can visualize the impact and persistence of structural tail event scenarios on forecasted out-of-sample (co-)variances.

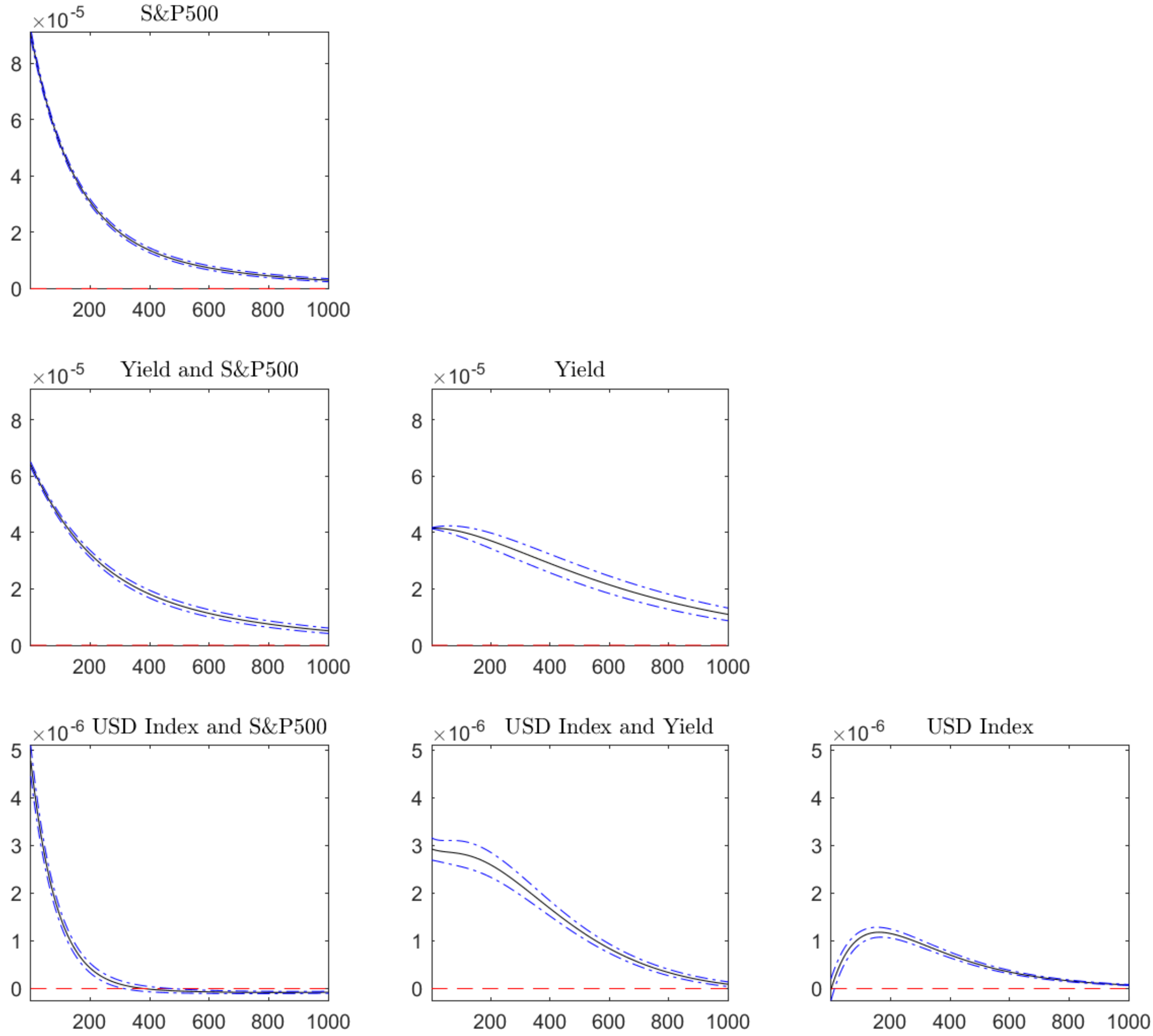

Figure 2: Plots of the estimated VIRFs with 5% confidence intervals (in blue) under the proxy-based identification scheme of the return system of the S&P 500 Composite Index (S&P500), the yield of the U.S. constant maturity 10 year treasury note (Yield) and the Finex U.S. Dollar Index (USD Index) in response to the structural shock on 02/27/2007 (Freddie Mac announcement).

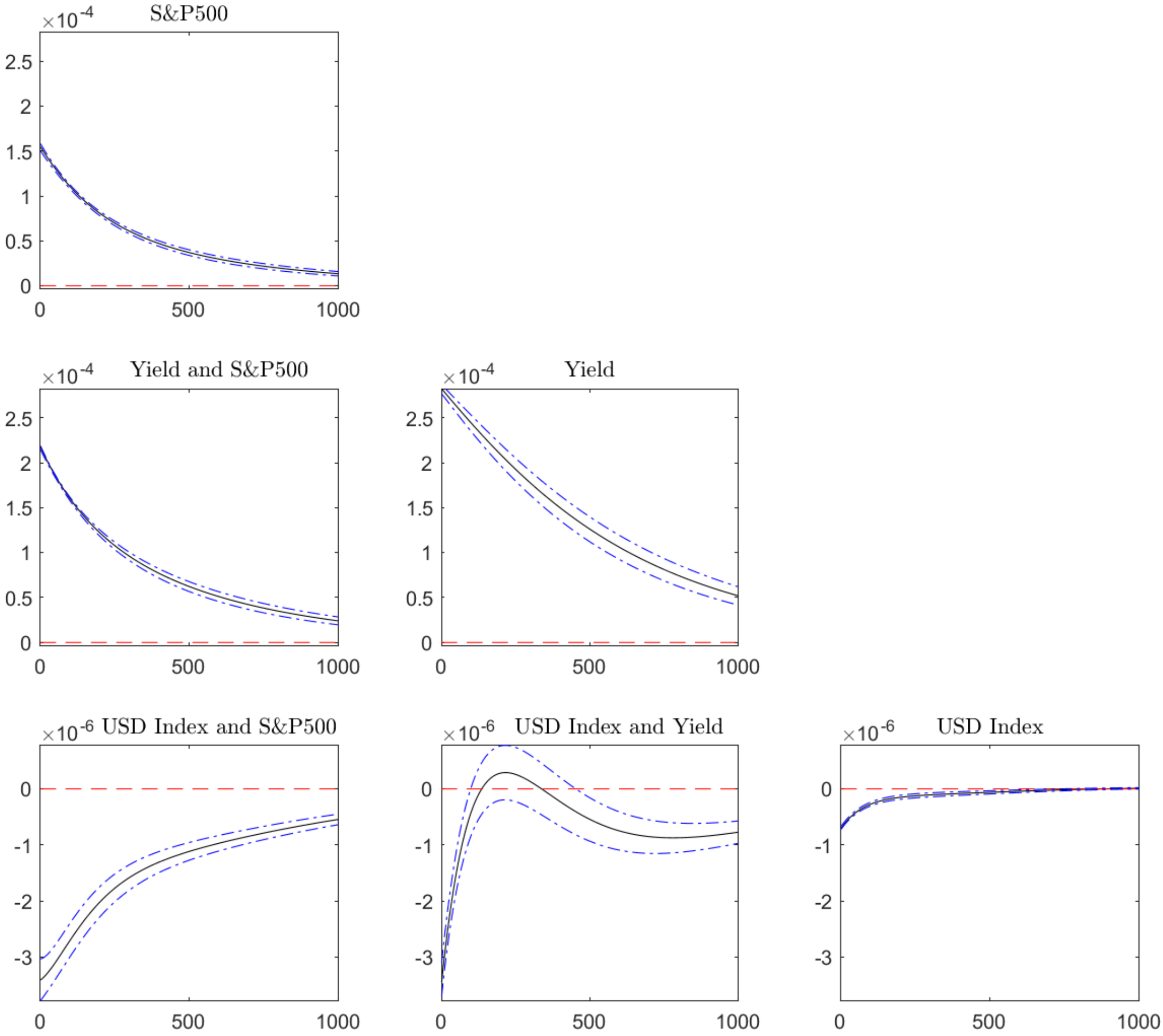

Figure 3: Plots of the estimated VIRFs with 5% confidence intervals (in blue) under the proxy-based identification scheme of the return system of the S&P 500 Composite Index (S&P500), the yield of the U.S. constant maturity 10 year treasury note (Yield) and the Finex U.S. Dollar Index (USD Index) in response to the structural shock on 09/15/2008 (Lehman Brothers bankruptcy)

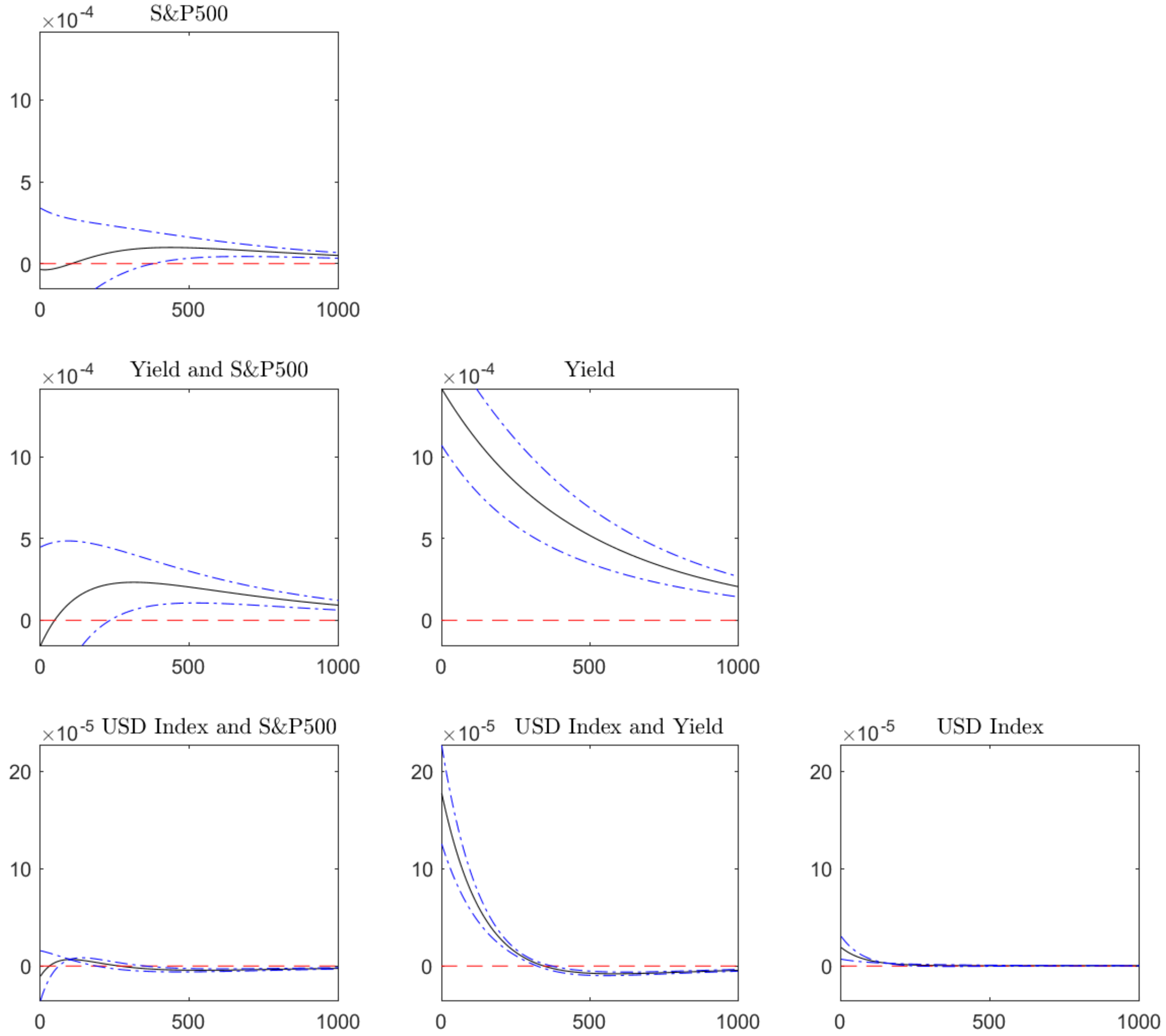

Figure 4: Plots of the estimated VIRFs with 5% confidence intervals (in blue) under the proxy-based identification scheme of the return system of the S&P 500 Composite Index (S&P500), the yield of the U.S. constant maturity 10 year treasury note (Yield) and the Finex U.S. Dollar Index (USD Index) in response to the structural shock on 03/18/2009 (FOMC announcement).

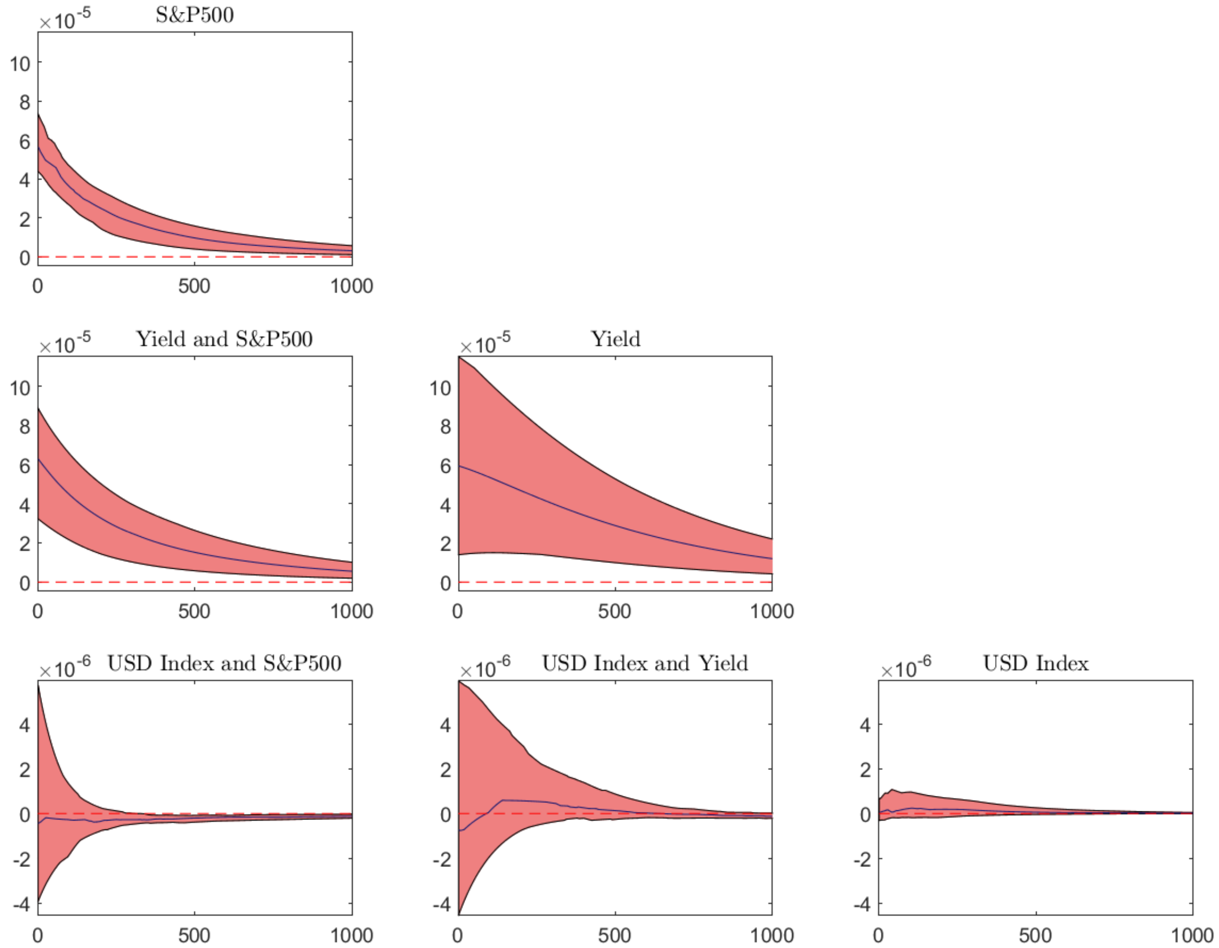

Figure 5: Plots of the median scenario VIRF and its 25 and 75% quantiles under the proxy-based identification scheme of the return system of the S&P 500 Composite Index (S&P500), the yield of the U.S. constant maturity 10 year treasury note (Yield) and the Finex U.S. Dollar Index (USD Index) from 1/1/1998 to 12/31/2014 in response to a scenario family of 1% structural equity tail event shocks on the out-of-sample date 01/02/2015.

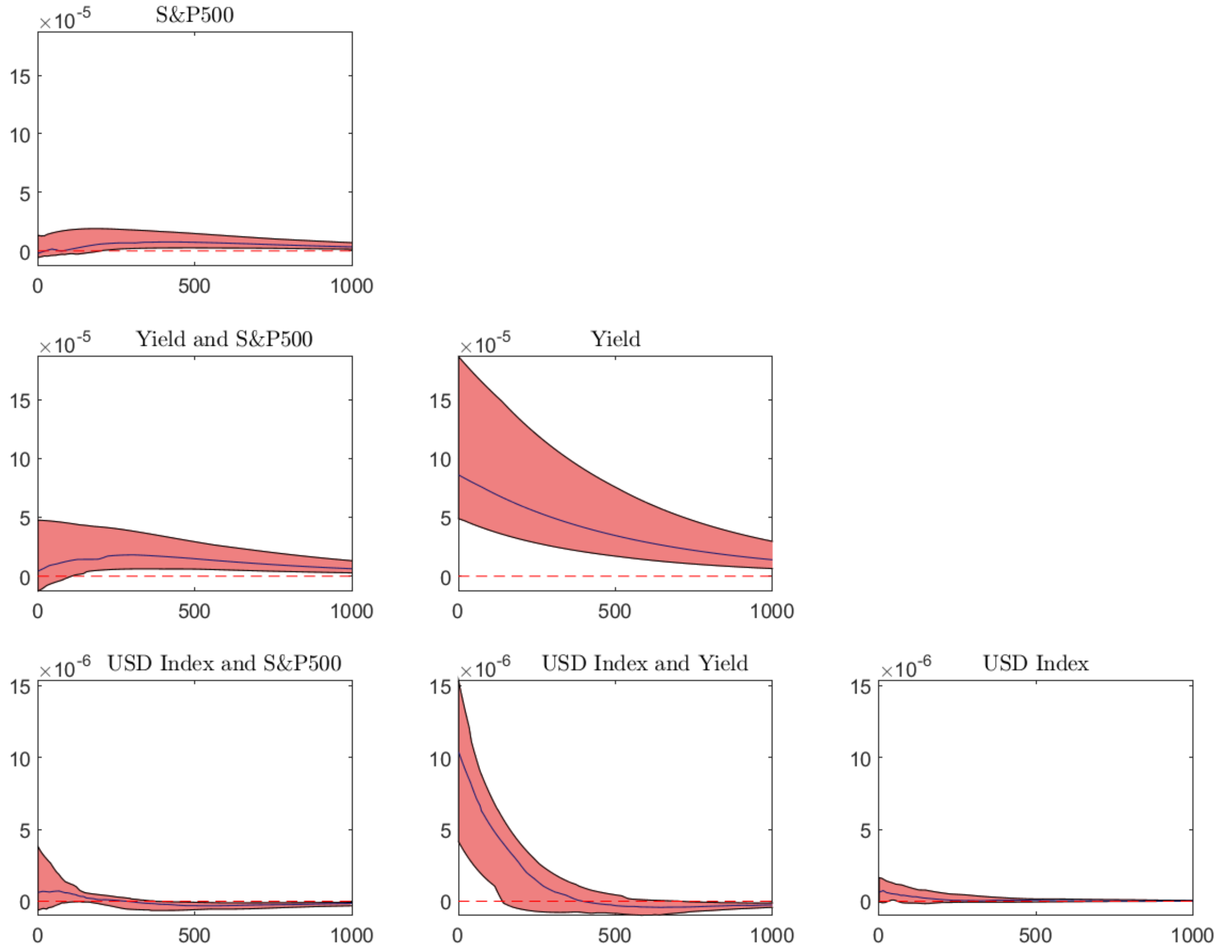

Figure 6: Plots of the median scenario VIRF and its 25 and 75% quantiles under the proxy-based identification scheme of the return system of the S&P 500 Composite Index (S&P500), the yield of the U.S. constant maturity 10 year treasury note (Yield) and the Finex U.S. Dollar Index (USD Index) from 1/1/1998 to 12/31/2014 in response to a scenario family of 1% structural bond market tail event shocks on the out-of-sample date 01/02/2015.

References

Amisano, G. and Giannini, C. (2012). *Topics in structural VAR econometrics*, Springer Science & Business Media.

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

Boussama, F., Fuchs, F. and Stelzer, R. (2011). Stationarity and geometric ergodicity of BEKK multivariate GARCH models, *Stochastic Processes and their Applications* **121**(10): 2331–2360.

Cenedese, G. and Mallucci, E. (2016). What moves international stock and bond markets?, *Journal of International Money and Finance* **60**: 94–113.

Cox, D. R. (1958). *Planning of experiments*, Wiley.

Davidson, J. (1994). *Stochastic Limit Theory: An Introduction for Econometricians*, Oxford University Press.

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

Fengler, M. R. and Polivka, J. (2021). Identifying structural shocks to volatility through a proxy-MGARCH model, *Technical report*, University of St.Gallen (HSG).

Francq, C. and Zakoïan, J. (2010). *GARCH Models: Structure, Statistical Inference and Financial Applications*, John Wiley & Sons Ltd.

Gallant, A., Rossi, P. and Tauchen, G. (1993). Nonlinear dynamic structures, *Econometrica* **61**(4): 871–907.

Grobys, K. (2015). Are volatility spillovers between currency and equity market driven by economic states? Evidence from the US economy, *Economics Letters* **127**: 72–75.

Hafner, C. and Herwartz, H. (2008). Analytical quasi maximum likelihood inference in multivariate volatility models, *Metrika: International Journal for Theoretical and Applied Statistics* **67**(2): 219–239.

Hafner, C. M. and Herwartz, H. (2006). Volatility impulse responses for multivariate GARCH models: An exchange rate illustration, *Journal of International Money and Finance* **25**(5): 719–740.

Hafner, C. M., Herwartz, H. and Maxand, S. (2020). Identification of structural multivariate GARCH models, *Journal of Econometrics.* (in press).

Hafner, C. and Preminger, A. (2009). On asymptotic theory for multivariate GARCH models, *Journal of Multivariate Analysis* **100**(9): 2044–2054.

Horn, R. A. and Johnson, C. R. (2012). *Matrix Analysis, 2nd Ed*, Cambridge University Press.

Inoue, A. and Kilian, L. (2013). Inference on impulse response functions in structural var models, *Journal of Econometrics* **177**(1): 1–13.

Jin, X., Lin, S. X. and Tamvakis, M. (2012). Volatility transmission and volatility impulse response functions in crude oil markets, *Energy economics* **34**(6): 2125–2134.

Kilian, L. (2013). Structural vector autoregressions, *Handbook of research methods and applications in empirical macroeconomics*, Edward Elgar Publishing.

Koop, G., Pesaran, M. H. and Potter, S. M. (1996). Impulse response analysis in nonlinear multivariate models, *Journal of Econometrics* **74**(1): 119–147.

Kremer, M., Lo Duca, M. and Holló, D. (2012). CISS - a composite indicator of systemic stress in the financial system, *Working Paper Series 1426*, European Central Bank.
URL: *https://ideas.repec.org/p/ecb/ecbwps/20121426.html*

Lin, W.-L. (1997). Impulse response function for conditional volatility in GARCH models, *Journal of Business & Economic Statistics* **15**(1): 15–25.

Liu, X. (2018). Structural volatility impulse response function and asymptotic inference, *Journal of Financial Econometrics* **16**(2): 316–339.

Longin, F. and Solnik, B. (1995). Is the correlation in international equity returns constant: 1960-1990?, *Journal of International Money and Finance* **14**(1): 3–26.

Lütkepohl, H., Staszewska-Bystrova, A. and Winker, P. (2015). Confidence bands for impulse responses: Bonferroni vs. Wald, *Oxford Bulletin of Economics and Statistics* **77**(6): 800–821.

Lütkepohl, H. (2010). Impulse response function, *Macroeconometrics and time series analysis*, Springer, pp. 145–150.

Magnus, J. R. and Neudecker, H. (1988). *Matrix Differential Calculus with Applications in Statistics and Econometrics*, John Wiley & Sons Ltd.

Mertens, K. and Ravn, M. O. (2013). The dynamic effects of personal and corporate income tax changes in the united states, *American Economic Review* **103**(4): 1212–1247.

Mueller, P., Tahbaz-Salehi, A. and Vedolin, A. (2017). Exchange Rates and Monetary Policy Uncertainty, *Journal of Finance* **72**(3): 1213–1252.

Olson, E., Vivian, A. J. and Wohar, M. E. (2014). The relationship between energy and equity markets: Evidence from volatility impulse response functions, *Energy Economics* **43**: 297–305.

Rambachan, A. and Shephard, N. (2020). Econometric analysis of potential outcomes time series: instruments, shocks, linearity and the causal response function. arXiv:1903.01637.
URL: *https://ideas.repec.org/p/arx/papers/1903.01637.html*

Rambachan, A. and Shephard, N. (2021). When do common time series estimands have nonparametric causal meaning?

Rubin, D. B. (1980). Randomization analysis of experimental data: The fisher randomization test comment, *Journal of the American Statistical Association* **75**: 591–593.

Sims, C. A. and Zha, T. (1999). Error bands for impulse responses, *Econometrica* **67**(5): 1113–1155.

Stock, J. H. and Watson, M. (2012). Disentangling the channels of the 2007-09 recession, *Brookings Papers on Economic Activity* **43**(1): 81–156.

A Proofs

A.1 VMA(∞) representation of the BEKK model

Proposition A.1. *The VMA(∞) representation of the BEKK(p, q) model (4) is given by*

$$X_t = \text{vech}(H) + \sum_{i=0}^{\infty} \Psi_i Y_{t-i} \tag{17}$$

where $\text{vech}(H) = \Phi(1)^{-1}c$ *denotes the long-run covariance matrix with* $c = \text{vech}(CC^\top)$ *and the* $(n^* \times n^*)$ *coefficient matrices* Ψ_i *are given by* $\Psi_0 = I_{n^*}$ *and* $\Psi_i = -\tilde{B}_i + \sum_{j=1}^{i} \left(\tilde{A}_j + \tilde{B}_j\right) \Psi_{i-j}$ *(*$i = 1, 2, \ldots$*) where* $\tilde{A}_j = D_n^+ \left(A_i \otimes A_i\right)^\top D_n$ *and* $\tilde{B}_j = D_n^+ \left(B_j \otimes B_j\right)^\top D_n$*.*

For the BEKK($1, 1$*) model, these expressions simplify to:* $\Psi_0 = I_{n^*}$*,* $\Psi_1 = \tilde{A}_1$ *and* $\Psi_i = \left(\tilde{A}_1 + \tilde{B}_1\right) \Psi_{i-1} = \left(\tilde{A}_1 + \tilde{B}_1\right)^{i-1} \tilde{A}_1$ *(*$i \geq 2$*).*

Proof. Transforming (4) to its equivalent VEC representation by applying the vec($\cdot$) operator

$$\text{vec}(H_t) = \text{vec}(CC^\top) + \sum_{i=1}^{p} \text{vec}(A_i^\top \varepsilon_{t-i}\varepsilon_{t-i}^\top A_i) + \sum_{j=1}^{q} \text{vec}(B_j^\top H_{t-j} B_j) \tag{18}$$

and using

$$\text{vec}(ABC) = \left(C^\top \otimes A\right) \text{vec}(B) \tag{19}$$

for appropriately defined matrices A, B and C and exploiting

$$\text{vec}(M) = D_n \, \text{vech}(M) \tag{20}$$

for any symmetric $(n \times n)$ matrix M, where D_n denotes the unique $(n^2 \times \frac{n(n+1)}{2})$ duplication matrix and noting that $M^\top \otimes M^\top = (M \otimes M)^\top$ for any matrix M, we get:

$$\begin{aligned} D_n \, \text{vech}(H_t) = D_n \, \text{vech}(CC^\top) &+ \sum_{i=1}^{p} \left(A_i \otimes A_i\right)^\top D_n \, \text{vech}(\varepsilon_{t-i}\varepsilon_{t-i}^\top) \\ &+ \sum_{j=1}^{q} \left(B_j \otimes B_j\right)^\top D_n \, \text{vech}(H_{t-j}). \end{aligned} \tag{21}$$

Multiplication with the Moore-Penrose inverse D_n^+ of the duplication matrix results in:

$$\begin{aligned} \text{vech}(H_t) = \underbrace{\text{vech}(CC^\top)}_{=:c} &+ \sum_{i=1}^{p} \underbrace{D_n^+ \left(A_i \otimes A_i\right)^\top D_n}_{=:\tilde{A}_i} \text{vech}(\varepsilon_{t-i}\varepsilon_{t-i}^\top) \\ &+ \sum_{j=1}^{q} \underbrace{D_n^+ \left(B_j \otimes B_j\right)^\top D_n}_{=:\tilde{B}_j} \text{vech}(H_{t-j}) \end{aligned} \tag{22}$$

To derive the equivalent VARMA(max$(p,q),q$) representation as in Hafner and Herwartz (2006), set $X_t := \text{vech}(\varepsilon_t\varepsilon_t^\top)$ and $Y_t := X_t - \text{vech}(H_t)$. Y_t is a weak white noise process with $\text{E}[Y_t] = 0$, $\text{Var}(Y_t) = H_Y$ and $\text{E}[Y_t Y_s^\top] = 0$ $(t \neq s)$. Rearranging (22) yields:

$$X_t = c + \sum_{i=1}^{\max(p,q)} \left(\tilde{A}_i + \tilde{B}_i\right) X_{t-i} - \sum_{j=1}^{q} \tilde{B}_j Y_{t-j} + Y_t \tag{23}$$

where $\tilde{A}_i = 0$ for $i > p$ and $\tilde{B}_i = 0$ for $i > q$. Given stationarity of (4), this can be rewritten in VMA(∞) form using the lag operator L:

$$\begin{aligned} \underbrace{\left(I_{n^*} - \sum_{i=1}^{\max(p,q)} \left(\tilde{A}_i + \tilde{B}_i\right) L^i\right)}_{=:\Phi(L)} X_t &= c + \underbrace{\left(I_{n^*} - \sum_{j=1}^{q} \tilde{B}_j L^j\right)}_{=:\Theta(L)} Y_t \\ \Leftrightarrow X_t &= \Phi(1)^{-1}c + \underbrace{\Phi(L)^{-1}\Theta(L)}_{=:\Psi(L)} Y_t \end{aligned} \tag{24}$$

such that X_t admits the VMA(∞) representation

$$X_t = \text{vech}(H) + \sum_{i=0}^{\infty} \Psi_i Y_{t-i} \tag{25}$$

where H satisfying $\text{vech}(H) = \Phi(1)^{-1}c$ denotes the long-run covariance matrix. The $(n^* \times n^*)$ coefficient matrices Ψ_i can be determined recursively by coefficient matching which yields the claim. □

A.2 Mathematical prerequisites for the VIRF

The following statements provide justification for interchanging infinite summation and expectation operators when exploiting vector moving average representations. Let $(\Omega, \mathcal{F}, \text{P})$ denote the probability space and let $\|\cdot\|$ be a matrix norm on $\mathbb{R}^{n\times n}$.

Definition A.1. A sequence of matrices $(\Psi_j)_{(j\in\mathbb{Z})} \subset \mathbb{R}^{k\times k}$ is called absolutely summable if $\sum_{j\in\mathbb{Z}} \|\Psi_j\| < \infty$. Similarly, the sequence $(\Psi_j)_{(j\in\mathbb{Z})}$ is called quadratically summable if $\sum_{j\in\mathbb{Z}} \|\Psi_j\|^2 < \infty$.

Corollary 1.2. *A sequence of matrices $(\Psi_j)_{(j\in\mathbb{Z})}$ in $\mathbb{R}^{k\times k}$ is absolutely summable if and only if $\sum_{j\in\mathbb{Z}} |\Psi_{mn,j}| < \infty$, where $\psi_{mn,j}$ denotes the (m,n)-th entry of Ψ_j.*

Proof. "$\Rightarrow$": Let $(\Psi_j)_{(j\in\mathbb{Z})}$ in $\mathbb{R}^{k\times k}$ be an absolutely summable sequence of matrices. Then $\sum_{j\in\mathbb{Z}}|\psi_{mn,j}| \leq \sum_{j\in\mathbb{Z}}\sum_{m=1}^{k}\sum_{n=1}^{k}|\psi_{mn,j}| = \sum_{j\in\mathbb{Z}}\|\Psi_j\|_1$, where $\|\cdot\|_1$ refers to the L^1-norm. Invoking the equivalence of matrix norms and the absolute summability of $(\Psi_j)_{(j\in\mathbb{Z})}$ the claim follows immediately.

"$\Leftarrow$": Let $(\psi_{mn,j})_{(j\in\mathbb{Z})}$, for all $m,n = 1,\dots,k$, be absolutely summable scalar sequences. By positivity and equivalence of matrix norms choosing the Frobenius norm: $0 \leq \|\Psi_j\| = \sqrt{\sum_{m=1}^{k}\sum_{n=1}^{k}|\psi_{mn,j}|^2} \leq \sum_{m=1}^{k}\sum_{n=1}^{k}\sqrt{|\psi_{mn,j}|^2} = \sum_{m=1}^{k}\sum_{n=1}^{k}|\psi_{mn,j}|$ such that $\sum_{j\in\mathbb{Z}}\|\Psi_j\| \leq \sum_{j\in\mathbb{Z}}\sum_{m=1}^{k}\sum_{n=1}^{k}|\psi_{mn,j}| = \sum_{m=1}^{k}\sum_{n=1}^{k}\left(\sum_{j\in\mathbb{Z}}|\psi_{mn,j}|\right)$. As the scalar sequences are absolutely summable by assumption and the outer sums are finite the claim follows immediately. □

Definition A.2. A random vector $Z \in \mathbb{R}^k$ is called (square-) integrable if all its components $Z_1,\dots,Z_k$ are (square-) integrable.

Corollary 1.3. *If the random vector $Z \in \mathbb{R}^k$ is square integrable, then $\operatorname{vech}(ZZ^\top)$ is integrable.*

Proof. Let $Z \in \mathbb{R}^k$ with $\mathrm{E}[Z_j^2] < \infty$ $(j = 1,\dots,k)$. Then by the Cauchy-Schwarz inequality $|\mathrm{E}[Z_m Z_n]| \leq \sqrt{\mathrm{E}[Z_m^2]\,\mathrm{E}[Z_n^2]} < \infty$ as each component is square integrable. □

Corollary 1.4. *Let $(Z_t)_{(t\in\mathbb{Z})}$ denote a random sequence on the probability space $(\Omega,\mathcal{F},\mathrm{P})$ with values in $\mathbb{R}^k$ and $\mathrm{E}[Z_t Z_t^\top] = H_z < \infty$. Then $\exists M \in \mathbb{R}$ such that for all $j = 1,\dots,k$: $\mathrm{E}[|Z_{j,t}|] < M < \infty$.*

Proof. The remark follows immediately by an application of Hölder's inequality and defining $M := \sqrt{\|H_z\|_{\max}} < \infty$ where $\|\cdot\|$ denotes the matrix maximum norm. □

Based on the previous auxiliary results, we can now prove the following:

Theorem 2. *Let $(Y_t)_{(t\in\mathbb{N}_0)}$ be a sequence of integrable random vectors in $\mathbb{R}^k$ defined on a probability space $(\Omega,\mathcal{F},\mathrm{P})$ and let $(\phi_i)_{(i\in\mathbb{N}_0)}$ be a sequence of absolutely summable matrices in $\mathbb{R}^{k\times k}$. If for all $m, j = 1,\dots,k$ the series $\sum_{i=0}^{\infty}|\phi_{mj,i}|\,\mathrm{E}[|Y_{j,h-i}|]$ converges then $\sum_{i=0}^{\infty}\phi_i Y_{h-i}(\omega)$ $(\omega \in \Omega)$ converges to a limiting random variable Y almost everywhere in Ω and $\mathrm{E}[Y] = \sum_{i=0}^{\infty}\phi_i\,\mathrm{E}[Y_{i-h}]$.*

Proof. Owing to Corollary 1.2 it is sufficient to consider sums of the form $\sum_{i=0}^{\infty}\phi_{mj,i}Y_{j,h-i}$ where $m, j = 1,\dots,k$ are arbitrary but fixed. An application of the monotone convergence theorem (Davidson, 1994, Theorem 4.5) then yields:

$$\mathrm{E}\left[\sum_{i=0}^{\infty}|\phi_{mj,i}Y_{j,h-i}|\right] = \sum_{i=0}^{\infty}\mathrm{E}\left[|\phi_{mj,i}Y_{j,h-i}|\right] = \sum_{i=0}^{\infty}|\phi_{mj,i}|\,\mathrm{E}\left[|Y_{j,h-i}|\right] < \infty,$$

where the right hand side is finite by assumption. Hence, $\sum_{i=0}^{\infty} \phi_{mj,i} Y_{j,h-i}$ exists. Hence, by the dominated convergence theorem (Davidson, 1994, Theorem 4.12): $\mathrm{E}\left[\sum_{i=0}^{\infty} \phi_{mj,i} Y_{j,h-i}\right] = \sum_{i=0}^{\infty} \mathrm{E}\left[\phi_{mj,i} Y_{j,h-i}\right]$. Using Corollary 1.2 this translates by means of finite summation into the claim: $\mathrm{E}[Y] = \sum_{i=0}^{\infty} \phi_i \, \mathrm{E}[Y_{i-h}]$. □

In order to apply the above theorem to the VIRF framework, we need to extend the previous result to conditional expectations. This generalization is based on the conditional versions of the monotone and dominated convergence theorem:

Corollary 2.1. *Let $(X_n)_{n\in\mathbb{N}}$ denote an increasing sequence of non-negative random variables on the probability space $(\Omega, \mathcal{F}, \mathrm{P})$ with almost sure limit X on Ω and let $\bar{\mathcal{F}}$ be a σ-algebra on Ω with $\bar{\mathcal{F}} \subset \mathcal{F}$. Then $\mathrm{E}[X|\bar{\mathcal{F}}] = \lim_{n\to\infty} \mathrm{E}[X_n|\bar{\mathcal{F}}]$ almost surely.*

Proof. For $F \in \bar{\mathcal{F}}$: By means of the tower property and the monotone convergence theorem it holds that $\mathrm{E}[\mathrm{E}[X|\bar{\mathcal{F}}]\mathbb{1}_F] = \mathrm{E}[X\mathbb{1}_F] = \lim_{n\to\infty} \mathrm{E}[X_n\mathbb{1}_F]$. By another application of the tower property and the monotone convergence theorem: $\lim_{n\to\infty} \mathrm{E}[X_n\mathbb{1}_F] = \lim_{n\to\infty} \mathrm{E}[\mathrm{E}[X_n\mathbb{1}_F|\bar{\mathcal{F}}]] = \lim_{n\to\infty} \mathrm{E}[\mathrm{E}[X_n|\bar{\mathcal{F}}]\mathbb{1}_F] = \mathrm{E}[\lim_{n\to\infty} \mathrm{E}[X_n|\bar{\mathcal{F}}]\mathbb{1}_F]$ which completes the proof. □

Corollary 2.2. *Let $(X_n)_{n\in\mathbb{N}}$ denote a sequence of random variables on the probability space $(\Omega, \mathcal{F}, \mathrm{P})$ and suppose $(X_n)_{n\in\mathbb{N}}$ converges almost surely to a random variable X. Let furthermore $\bar{\mathcal{F}}$ be a σ-algebra on Ω with $\bar{\mathcal{F}} \subset \mathcal{F}$. If there exists an integrable random variable Y such that for all n: $|X_n| \leq Y$ almost surely then we have: $\lim_{n\to\infty} \mathrm{E}[X_n|\bar{\mathcal{F}}] = \mathrm{E}[X|\bar{\mathcal{F}}]$ almost surely.*

Proof. For all $F \in \bar{\mathcal{F}}$: By means of the tower property and the dominated convergence theorem it holds that $\mathrm{E}[\mathrm{E}[X|\bar{\mathcal{F}}]\mathbb{1}_F] = \mathrm{E}[X\mathbb{1}_F] = \lim_{n\to\infty} \mathrm{E}[X_n\mathbb{1}_F]$ where $|X\mathbb{1}_F| \leq Y \in L^1$ and $|X_n\mathbb{1}_F| \leq Y \in L^1$. By another application of the tower property this equals $\lim_{n\to\infty} \mathrm{E}[\mathrm{E}[X_n|\bar{\mathcal{F}}]\mathbb{1}_F]$ which completes the proof. □

With these extensions at hand we can now formulate the following claim:

Corollary 2.3. *Let $(Y_t)_{(t\in\mathbb{N}_0)}$ be a sequence of integrable random vectors in $\mathbb{R}^k$ defined on a probability space $(\Omega, \mathcal{F}, \mathrm{P})$, let $\tilde{\mathcal{F}}$ be a σ-algebra on Ω with $\tilde{\mathcal{F}} \subset \mathcal{F}$ and let $(\phi_i)_{(i\in\mathbb{N}_0)}$ be a sequence of absolutely summable matrices in $\mathbb{R}^{k\times k}$. If for all $m, j = 1, \ldots, k$ it holds $\sum_{i=0}^{\infty} |\phi_{mj,i}| \, \mathrm{E}[|Y_{j,h-i}|] < \infty$ then $\sum_{i=0}^{\infty} \phi_i Y_{h-i}(\omega)$ exists and $\mathrm{E}[\sum_{i=0}^{\infty} \phi_i Y_{i-h}|\tilde{\mathcal{F}}] = \sum_{i=0}^{\infty} \phi_i \, \mathrm{E}[Y_{i-h}|\tilde{\mathcal{F}}]$. Thus we can interchange infinite summation and conditional expectation.*

Proof. The proof of Theorem 2 translates immediately into the proof for Corollary 2.3 by replacement of the monotone convergence theorem and the dominated convergence theorem with their conditional counterparts (see Corollary 2.1 and Corollary 2.2). □

Proof of Proposition 2.1. Let $X_t = \text{vech}(\varepsilon_t\varepsilon_t^\top)$ and $Y_t = X_t - \text{vech}(H_t)$ and $h \geq 1$. Exploiting the tower property of the conditional expectation yielding

$$\text{E}[\text{vech}(H_{t+h})|\mathcal{F}_{t-1}] = \text{E}[\text{E}[X_{t+h}|\mathcal{F}_{t+h-1}]|\mathcal{F}_{t-1}] = \text{E}[X_{t+h}|\mathcal{F}_{t-1}] \tag{26}$$

and inserting the VMA(∞) representation in (3) we obtain:

$$\text{V}_{t+h}(\xi_t;\eta) = \text{E}\left[\sum_{i=0}^{\infty}\Psi_i Y_{t+h-i}|\tilde{\mathcal{F}}_t\right] - \text{E}\left[\sum_{i=0}^{\infty}\Psi_i Y_{t+h-i}|\mathcal{F}_{t-1}\right]. \tag{27}$$

By square integrability of $(\varepsilon_t)_{t\in\mathbb{Z}}$, $(X_t)_{t\in\mathbb{Z}}$ is a sequence of integrable random variables and as the series $\text{vech}(H_t)_{t\in\mathbb{Z}}$ is integrable by the square integrability of ε_t (see Corollary 1.3), so is their difference $(Y_t)_{t\in\mathbb{Z}}$. As $\text{Var}(Y_t) = H_Y < \infty$, the absolute moments of Y_t are uniformly bounded (see Corollary 1.4). Hence, by Corollary 2.3, we can interchange infinite summation and conditional expectation yielding:

$$\text{V}_{t+h}(\xi_t;\eta) = \sum_{i=0}^{\infty}\Psi_i\left(\text{E}\left[Y_{t+h-i}|\tilde{\mathcal{F}}_t\right] - \text{E}[Y_{t+h-i}|\mathcal{F}_{t-1}]\right). \tag{28}$$

In (28), $\text{E}\left[Y_{t+h-i}|\tilde{\mathcal{F}}_t\right] - \text{E}[Y_{t+h-i}|\mathcal{F}_{t-1}] = 0$ for all $(t+h-i) \leq t-1$ due to measurability given $\mathcal{F}_{t-1}$. Secondly, $\text{E}\left[Y_{t+h-i}|\tilde{\mathcal{F}}_t\right] - \text{E}[Y_{t+h-i}|\mathcal{F}_{t-1}] = 0$ for all $(t+h-i) \geq t+1$ by the tower property. Similarly, using the tower property and the independence of ξ_{t+h-i} of $\mathcal{F}_{t-1}$ and of ξ_t we obtain $\text{E}[Y_{t+h-i}|\tilde{\mathcal{F}}_t] = 0$. Thus, (28) reduces to

$$\text{V}_{t+h}(\xi_t,\eta) = \Psi_h\left(\text{E}\left[Y_t|\tilde{\mathcal{F}}_t\right] - \text{E}[Y_t|\mathcal{F}_{t-1}]\right). \tag{29}$$

Using predictability[5] and tower property arguments and inserting the structural model (2) we obtain:

$$\begin{aligned}
\text{V}_{t+h}(\xi_t;\eta) &= \Psi_h\left(\text{E}\left[X_t - \text{vech}(H_t)|\tilde{\mathcal{F}}_t\right] - \text{E}[X_t - \text{vech}(H_t)|\mathcal{F}_{t-1}]\right)\\
&= \Psi_h\left(\text{E}\left[\text{vech}(\varepsilon_t\varepsilon_t^\top)|\tilde{\mathcal{F}}_t\right] - \text{E}\left[\text{vech}(\varepsilon_t\varepsilon_t^\top)|\mathcal{F}_{t-1}\right]\right)\\
&= \Psi_h\left(\text{E}\left[\text{vech}\left(H_t^{1/2}\tilde{R}\xi_t\xi_t^\top\tilde{R}^\top H_t^{1/2^\top}\right)|\tilde{\mathcal{F}}_t\right] - \text{E}\left[\text{vech}(H_t)|\mathcal{F}_{t-1}\right]\right)\\
&= \Psi_h\left(\text{vech}\left(H_t^{1/2}\tilde{R}\xi_t\xi_t^\top\tilde{R}^\top H_t^{1/2^\top}\right) - \text{vech}\left(H_t^{1/2}\tilde{R}\tilde{R}^\top H_t^{1/2^\top}\right)\right)\\
&= \Psi_h\left(\text{vech}\left(H_t^{1/2}(\tilde{R}\xi_t\xi_t^\top\tilde{R}^\top - I_n)H_t^{1/2^\top}\right)\right).
\end{aligned} \tag{30}$$

By the symmetry of $(\tilde{R}\xi_t\xi_t^\top\tilde{R}^\top - I_n)$ and (19):

$$\text{V}_{t+h}(\xi_t;\eta) = \Psi_h D_n^+\left(H_t^{1/2}\otimes H_t^{1/2}\right)D_n\left(\text{vech}(\tilde{R}\xi_t\xi_t^\top\tilde{R}^\top - I_n)\right). \tag{31}$$

□

[5] Note that $H_t^{1/2}$ is $\mathcal{F}_{t-1}$-measurable. The measurability follows from the $\mathcal{F}_{t-1}$-measurability of H_t and because the principal square root is a (uniformly) continuous operator in the space of positive definite matrices. Matrix multiplication with $\tilde{R}$ preserves measurability.

A.3 Asymptotic theory for VIRFs

Proof of Proposition 2.4. Let $\eta \in \mathbb{R}^m$ denote the vector of stacked parameters of the BEKK(p,q) model: $\eta = \left(\mathrm{vec}(C)^\top, \mathrm{vec}(A_1)^\top, \ldots, \mathrm{vec}(A_p)^\top, \mathrm{vec}(B_1)^\top, \ldots, \mathrm{vec}(B_q)^\top\right)^\top$. Starting with the definition of the VIRF for BEKK(p,q) models we have:

$$\mathrm{V}_{t+h}(\xi_t;\eta) = \Psi_h D^+ \left(\mathrm{vec}(H_t^{1/2}\tilde{R}\xi_t\xi_t^\top \tilde{R}^\top (H_t^{1/2})^\top) - \mathrm{vec}(H_t)\right) \tag{32}$$

where $(\Psi_h)_{h\in\mathbb{N}}$ are given in Proposition A.1. To calculate the derivative of the VIRF with respect to η we make use of the following results. Firstly, we use that for $n \times n$ matrices X and Z, Z symmetric, it holds:

$$\frac{\partial\, \mathrm{vec}(XZX)}{\partial\, \mathrm{vec}(X)^\top} = (XZ \otimes I_n) + (I_n \otimes XZ) \tag{33}$$

by an application of the chain rule (see Magnus and Neudecker, 1988, Theorem 5.12) in conjunction with two applications of (19). Secondly,

$$\frac{\partial\, \mathrm{vec}(H_t^{1/2})}{\partial\, \mathrm{vec}(H_t)^\top} = \left[\left(I_n \otimes H_t^{1/2}\right) + (H_t^{1/2} \otimes I_n)\right]^{-1} \tag{34}$$

which can be derived by solving a Sylvester type equation for the differential. Moreover, the derivatives of the BEKK model with regard to the parameter matrices are available in closed form (Hafner and Herwartz, 2008), such that $\frac{\partial\, \mathrm{vec}(H_t)}{\partial \eta^\top}$ is known. Then, for the VIRF at time $h = 0$ ($\Psi_0 = I_{n^*}$):

$$\begin{aligned}
\frac{\partial \mathrm{V}_t(\xi_t;\eta)}{\partial \eta^\top} &= D^+ \left[\frac{\partial\, \mathrm{vec}\left(H_t^{1/2}\tilde{R}\xi_t\xi_t^\top\tilde{R}^\top H_t^{1/2}\right)}{\partial\, \mathrm{vec}\left(H_t^{1/2}\right)^\top} \frac{\partial\, \mathrm{vec}\left(H_t^{1/2}\right)}{\partial\, \mathrm{vec}\left(H_t\right)^\top} \frac{\partial\, \mathrm{vec}(H_t)}{\partial \eta^\top} - \frac{\partial\, \mathrm{vec}(H_t)}{\partial \eta^\top}\right] \\
&\quad \times \left[\left(H_t^{1/2} \otimes I_n\right) + \left(I_n \otimes H_t^{1/2}\right)\right]^{-1} \frac{\partial\, \mathrm{vec}(H_t)}{\partial \eta^\top} - \frac{\partial\, \mathrm{vec}(H_t)}{\partial \eta^\top}\Bigg] \\
&= D^+ \Bigg[\left(\left(H_t^{1/2}\tilde{R}\xi_t\xi_t^\top\tilde{R}^\top \otimes I_n\right) + \left(I_n \otimes H_t^{1/2}\tilde{R}\xi_t\xi_t^\top\tilde{R}^\top\right)\right) \\
&\quad \times \left[\left(H_t^{1/2} \otimes I_n\right) + \left(I_n \otimes H_t^{1/2}\right)\right]^{-1} - I_{n^2}\Bigg] \frac{\partial\, \mathrm{vec}(H_t)}{\partial \eta^\top}.
\end{aligned}$$

Now let $h \in \mathbb{N}$. Exploiting the fact that the VMA coefficients Ψ_i, $i = 1, \ldots, h$ are recursively defined, the general derivative of the BEKK(p,q) VIRF can be derived building on the product rule (see Magnus and Neudecker, 1988, Theorem 5.12):

$$\frac{\partial \mathrm{V}_{t+h}(\xi_t;\eta)}{\partial \eta^\top} = \left(\mathrm{V}_t^\top \otimes I_{\frac{n(n+1)}{2}}\right) \frac{\partial\, \mathrm{vec}(\Psi_h)}{\partial \eta^\top} + \left(I_{\frac{n(n+1)}{2}} \otimes \Psi_h\right) \frac{\partial \mathrm{V}_t(\xi_t;\eta)}{\partial \eta^\top}$$

Moreover, we can establish the recursion

$$\frac{\partial \operatorname{vec}(\Psi_h)}{\partial \eta^\top} = \frac{\partial \operatorname{vec}(-\tilde{B}_h)}{\partial \eta^\top} + \sum_{j=1}^{h} \frac{\partial \operatorname{vec}\left((\tilde{A}_j + \tilde{B}_j)\, \Psi_{h-j}\right)}{\partial \eta^\top}$$

$$= \frac{\partial \operatorname{vec}(-\tilde{B}_h)}{\partial \eta^\top} + \sum_{j=1}^{h} \left(\Psi_{h-j}^\top \otimes I_{n^*}\right) \frac{\partial \operatorname{vec}\left(\tilde{A}_j + \tilde{B}_j\right)}{\partial \eta^\top} + \left(I_{n^*} \otimes (\tilde{A}_j + \tilde{B}_j)\right) \frac{\partial \operatorname{vec}\left(\Psi_{h-j}\right)}{\partial \eta^\top}.$$

To evaluate this expression, we derive $\frac{\partial \operatorname{vec}(\tilde{A}_j)}{\partial \eta^\top}$, $(j = 1, \ldots, h)$. The derivations for $\tilde{B}_j$ follow analogously and the vec operator is a linear operator. To this end, we make use of the following relations: For some $(n \times n)$ matrices M, P it holds (see Magnus and Neudecker (1988, Theorem 3.10))

$$\operatorname{vec}(M \otimes P) = (I_n \otimes K_n \otimes I_n)\left[\operatorname{vec}(M) \otimes \operatorname{vec}(P)\right] \tag{35}$$

where K_n denotes the $(n^2 \times n^2)$ commutation matrix which satisfies $K_n \operatorname{vec}(M) = \operatorname{vec}(M^\top)$. Furthermore, it holds that $\frac{\partial \operatorname{vec}(M) \otimes \operatorname{vec}(P)}{\partial \operatorname{vec}(M)^\top} = I_n \otimes \operatorname{vec}(P)$ such that by an application of the product rule $\frac{\partial(\operatorname{vec}(M) \otimes \operatorname{vec}(M))}{\partial \operatorname{vec}(M)^\top} = (I_n \otimes \operatorname{vec}(M)) + (\operatorname{vec}(M) \otimes I_n)$. Finally, $\frac{\partial \operatorname{vec}(M^\top)}{\partial \operatorname{vec}(M)^\top} = K_n$. Thus it holds by inserting the definition of $\tilde{A}_j$, using (19) and the aforementioned formulae:

$$\frac{\partial \operatorname{vec}\left(\tilde{A}_j\right)}{\eta^\top} = \frac{\partial \operatorname{vec}\left(D_n^+ \left(A_j \otimes A_j\right)^\top D_n\right)}{\partial \eta^\top}$$

$$= \frac{(D_n^\top \otimes D_n^+) \partial \operatorname{vec}\left(A_j \otimes A_j\right)^\top}{\partial \eta^\top}$$

$$= \frac{(D_n^\top \otimes D_n^+) \partial \operatorname{vec}\left(A_j^\top \otimes A_j^\top\right)}{\partial \eta^\top}$$

$$= (D_n^\top \otimes D_n^+)\left(I_n \otimes K_n \otimes I_n\right) \frac{\partial\left(\operatorname{vec}(A_j^\top) \otimes \operatorname{vec}(A_j^\top)\right)}{\partial \eta^\top}.$$

Thus, the formula for the derivative of the BEKK(p,q) VIRF can be implemented recursively.

For the BEKK(1,1) model, we can derive a more compact recursive formula for $\frac{\partial \mathrm{V}_{t+h}(\eta)}{\partial \eta^\top}$ based on (6). Let $h \in \mathbb{N}$. By an application of the product rule it holds:

$$\frac{\partial \mathrm{V}_{t+h}(\xi_t; \eta)}{\partial \eta^\top} = \frac{\partial \operatorname{vec}\left(\left(\tilde{A} + \tilde{B}\mathbb{1}_{\{h>1\}}\right) \mathrm{V}_{t+h-1}(\xi_t; \eta)\right)}{\partial \eta^\top}$$

$$= \left(\mathrm{V}_{t+h-1}(\xi_t; \eta)^\top \otimes I_{\frac{n(n+1)}{2}}\right) \frac{\partial \operatorname{vec}\left(\tilde{A} + \tilde{B}\mathbb{1}_{\{h>1\}}\right)}{\eta^\top} + \left(\tilde{A} + \tilde{B}\mathbb{1}_{\{h>1\}}\right) \frac{\partial \mathrm{V}_{t+h-1}(\xi_t; \eta)}{\partial \eta^\top}.$$

Hence, in this case we have to calculate $\frac{\partial \operatorname{vec}\left(\tilde{A}+\tilde{B}\mathbb{1}_{\{h>1\}}\right)}{\eta^\top}$ only once to establish the recursion for $\frac{\partial \mathrm{V}_{t+h}(\xi_t;\eta)}{\partial \eta^\top}$. This completes the proof. □