Structural Volatility Impulse Response Analysis*

Matthias Fengler‡ Jeannine Polivka§

November 6, 2022

Go to most recent version.

Abstract

In this paper, we make three contributions to the volatility impulse response function (VIRF) developed by Hafner and Herwartz (2006), the most widely applied impulse response function in the context of multivariate volatility models. First, we derive its law for multivariate generalized autoregressive conditional heteroscedasticity (MGARCH) models of the BEKK type. Second, we present a structural embedding of the VIRF by relying on recent developments concerning identification of MGARCH models. This broadens the use cases of the VIRF, which has previously been limited to historical analyses, by allowing for counterfactual and out-of-sample scenario analyses of volatility responses. Third, 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 shock events as well as the consequences of well-defined future shock scenarios on the U.S. equity, government bond and foreign exchange markets. Our findings suggest that it is vital to be able to assess the statistical significance of volatility impulse responses.

Keywords: causality in volatility, multivariate GARCH models, proxy identification, 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").

‡SFI, 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 for analyzing how dynamic multivariate systems respond to shocks. It shows how an unanticipated perturbation impacts the modeled variables and exhibits how the 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) can afford the relevant insights.

The VIRF, as first conceptualized by Hafner and Herwartz (2006), extends the ideas of the generalized IRF (GIRF) developed by 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 the extant VIRF specifications, the VIRF of Hafner and Herwartz (2006) is especially attractive because it allows a closed-form expression for one of the most general multivariate generalized autoregressive conditional heteroskedasticity (MGARCH) models, the Baba, Engle, Kraft and Kroner or BEKK(p, q) model (Engle and Kroner, 1995), and is also available for the Markov-switching MGARCH (Cavicchioli, 2019). In this work, we take a fresh look at the VIRF and add to the extant literature in three ways: First, we derive the asymptotic distribution of the VIRF in the BEKK model, which is the most frequently applied form 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 until now could only be obtained by time-consuming simulation techniques, such as the bootstrap method. Importantly, our empirical analysis suggests that the interpretation of VIRFs without taking into account statistical significance may lead researchers astray because seemingly long-term impacts can be of much shorter-lived significance than one may be tempted to expect.

As a second contribution, we endow the VIRF with a structural interpretation by relying on recent advances regarding identification in MGARCH models (Fengler and Polivka, 2021; Hafner et al., 2022). This new interpretation materially broadens the use cases of the VIRF, which has previously been limited to historical analyses. A structural volatility model featuring identified and labeled structural shocks raises the curtain for proper scenario analyses, as is common in structural vector autoregressive (VAR) analyses of the mean equation (Amisano and Giannini, 2012; Kilian, 2013). It allows one, e.g., to define counterfactual scenarios and 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 designed for asset price dynamics, risk management and portfolio optimization, and has been investigated by, among others, Gallant et al. (1993), Lin (1997), Hafner and Herwartz (2006) and more recently by Liu (2018) and Cavicchioli (2019). However, despite this long

history, structural advances in 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 useful baseline for setting the conditional volatility profile into context, because it defines the news to appear in the conditionally heteroskedastic errors. Lin (1997) bases his univariate volatility impulse response analysis on reduced-form GARCH models with the result that the shocks lack a 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 devised by 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, which does not offer 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 the statistical inference of their VIRFs. As an initial step toward inferential statistics, Liu (2018) derives confidence intervals for an alternative VIRF definition based on the dynamic conditional correlation (DCC) model. However, he applies his VIRF to shocks which are identified statistically via time-varying heteroskedasticity and thus still do not provide a direct economic interpretation.

The remainder of this paper is structured as follows. Section 2 presents the concept of the VIRF, including rigorous mathematical derivations, derives its asymptotic distribution and establishes the connections to structural and causal modeling. Section 3 briefly presents the structural model we use for the 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 this work. Appendix A provides an overview of the definitions and mathematical prerequisites as well as featuring all the proofs.

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 $\{\mathcal{F}_t\}$. 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 (see Definition D.1) of H_t which exists for all t and is uniquely defined. Furthermore, we define $\tilde{R}$ to be a structural rotation matrix (see Definition D.2) and $\xi = (\xi_t)_{t\in\mathbb{Z}}$ is 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 (Fengler and Polivka, 2021; Hafner et al., 2022).

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 $\tilde{R}$ as 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 economic grounds for asset return systems. Potential avenues for the identification of $\tilde{R}$ include the identification by non-Gaussianity developed by Hafner et al. (2022) and the proxy-based scheme devised by Fengler and Polivka (2021) which may additionally allow for shock labeling. We sketch the latter identification scheme in Section 3.1.

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 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 GIRF developed by Koop et al. (1996) to address the problems of history, shock and compositional dependence of impulse responses in nonlinear multivariate models (see the discussion at the end of this section).

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

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

where the vech operator is defined in Definition D.4. Thus, $\mathrm{V}_{t+h}(\xi_t)$ is an n^*-dimensional vector of impulse responses of the conditional (co-)variances with $n^* = \frac{n(n+1)}{2}$. Note that despite referring 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 IRFs for conditional means, there is no natural baseline for volatility based on a fixed return vector.

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 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 a quadratic form is independent of the underlying structural model.[1] This invariance allows one to infer past shocks quickly and efficiently and to calculate corresponding historical VIRFs by means of (2) and (3) using $\tilde{R} = I_n$; this makes any additional computational effort related to structural modeling redundant in the historical set-up. On the other hand, however, in a 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): first, 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; second, it is unrealistic to observe perturbations in solely one shock while keeping the others fixed, because impulse responses generally depend on compositions of shocks. A structural volatility model, in a similar manner 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

[1] 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 (compare Section 2.5).

one to investigate the impacts of specific future market scenarios defined by the labeled structural shocks. Such 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 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 the 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 the consistency as well as asymptotic normality of the quasi-maximum likelihood (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 the sign. For the BEKK(1, 1), the model is uniquely identified if one assumes 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, a closed-form expression exists for the h-step ahead VIRF (Hafner and Herwartz, 2006). For the sake of completeness, we provide a rigorous derivation in the appendix, which we adapted to the case of a structural model of type (2) and which will be of further value in Section 2.4.

Proposition 2.1. *Assume H_t has the BEKK(p,q) representation given in (2) with parameter vector $\eta = \left(\mathrm{vec}(C)^\top, \mathrm{vec}(A_i)^\top, \mathrm{vec}(B_j)^\top\right)^\top$, $(i = 1,\dots,p;\ j = 1,\dots,q;\ p,q \in \mathbb{N})$, see D.3 for a definition of $\mathrm{vec}(\cdot)$, and with a VMA(∞) representation with coefficients $(\Psi_i)_{i\in\mathbb{N}}$ as provided in Proposition A.1. Then the h-step ahead VIRF given a structural shock ξ_t is*

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

where D_n denotes the duplication matrix and D_n^+ its Moore-Penrose inverse (see Definition D.6).

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

Notably, the IRF 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.

Remark. 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 (Hafner and Herwartz, 2006):

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

This can be shown by deriving the vech form of the BEKK(1, 1) model, as in Proposition A.1, Equation (35), and inserting the VMA(∞) coefficients in (5).

2.4 Asymptotic law of the VIRF

As our first central contribution, we now complete the notion of the VIRF by providing the asymptotic theory.

2.4.1 Consistency of the VIRF

Proposition 2.2. *Let $h \in \mathbb{N}$ and $\xi_t \in \mathbb{R}^n$ arbitrary but fixed. 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, $\mathrm{V}_{t+h}(\xi_t, \eta)$ can be estimated consistently by replacing η with its QMLE $\hat{\eta}$.*

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

Proof. By continuity and the continuous mapping theorem. □

Remark. The BEKK(p, q) model is continuous in η and therefore the BEKK-VIRF.

2.4.2 Asymptotics of the BEKK-VIRF

Theorem 1. *Let $h \in \mathbb{N}$. Under the regularity conditions of Hafner and Preminger (2009)[2] and given continuous differentiability of the BEKK-VIRF as function of $\eta \in \mathbb{R}^m$, we have:*

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

[2] In particular, assuming that the true parameter η lies inside the compact parameter space $\mathcal{E}$.

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*;* $\boldsymbol{\mathcal{H}}(\eta) = \frac{\partial^2 \log(l_t(\eta))}{\partial \eta \partial \eta^\top}$ *is the Hessian matrix of the log-likelihood contribution* $\log(l_t)$ *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. The* $(n^* \times m)$ *Jacobian of the VIRF with respect to* $\eta \in \mathbb{R}^m$ *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 which is equal to one if the condition in the subscript is satisfied and zero otherwise.*

Proof. See Appendix A.3. □

The asymptotic distribution of the structural VIRF estimator allows us to construct simultaneous confidence intervals for the impulse responses of individual asset returns to structural shocks. For our application, we choose the χ^2-approximation to Hotelling's T^2 statistic to compute simultaneous confidence intervals for all components of the VIRF for a given significance level α (see, e.g., Sims and Zha, 1999, and Lütkepohl et al., 2015, for discussions regarding confidence interval construction for impulse responses). This enables 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 past financial, economic, or political events. We accordingly refer to this approach as historical VIRF analysis. By taking the estimated residual $\hat{\varepsilon}_t$ and the volatility state $\hat{H}_t$ at the time of the shock occurrence, Hafner and Herwartz (2006) define their structural shocks $\hat{\xi}_t = \hat{H}_t^{-1/2}\hat{\varepsilon}_t$ using the principal square root, an approach which to the best of our knowledge has not been questioned so far. Hafner et al. (2022) and Fengler and Polivka (2021), however, 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:

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

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

However, structural volatility models stimulate new usages of VIRFs. First, in out of sample settings, structural decompositions allow one to process responses to synthetic but economically meaningful structural shocks. Second, structural VIRFs allow for conducting counterfactual analyses by defining specific shock scenarios. We coin the term "scenario VIRF" for this new application framework and illustrate its merits in Section 3.

2.6 Causality

Our volatility impulse response analysis connects well to recent advances on causality in times series where changes in model outcomes are interpreted as the result of changes in structural shocks, which are regarded as treatment variables. This notion of causality (see Lechner (2010) for an overview of the different causality concepts in econometrics) differs from the one usually employed in time series analysis, the Granger-Sims causality (Granger, 1969; Sims, 1972). The absence of Granger causality means non-predictability, whereas here we compare different states of the world, which are identified with different values of the causing variable. Therefore, in absence of causation, the outcomes would be the same whatever the state of the causing variable. This interpretation requires additional identifying assumptions that allow one to relate the observed data to the distribution of the potential outcome variables allowing the researcher to infer the causal effect from observed data. In their recent advances, Rambachan and Shephard (2020, 2021) provide conditions how the GIRF of Koop et al. (1996) can be endowed with this causal interpretation. In contrast to the GIRF, the VIRF does not trace the dynamic effects of a shock to the conditional mean but to the conditional covariance matrix. It does not have a causal content a priori, as it is the difference of two conditional expectations. In contrast, a causal effect traces the effects of changes in treatments. To show that the VIRF can be endowed with a causal interpretation as well, we adapt the assumptions of Rambachan and Shephard (2020, 2021) to model (2). We start by rephrasing our variables of interest with terms stemming from causal inference:

Definition 2.3. Let $\xi := (\xi_t)_{t\in I}$, $(I = \{1, \dots, T\})$, denote a 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 $\mathcal{G}_t := \sigma(\{X_s, \xi_s \ (s \leq t)\})$.

The demeaned return vectors, and their respective outer products, are the observable, continuously valued, multidimensional outcomes, for which the VMA(∞) representation of the BEKK model is available. Thus, 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 $\mathcal{G}_t \subseteq \sigma(\varepsilon_s, \xi_s, s \leq t\})$ 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 treated as deterministic, it even holds that $\mathcal{G}_t = \tilde{\mathcal{F}}_t$ (see Definition 2.2). To define subsequences of entire time series, we use the notation $\xi_{s:t}$ to indicate the start s and end point t of the subseries in time. Furthermore, let 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 such 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 the stable unit treatment value assumption (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 $\mathcal{G}_{t-1}$-measurable, and the fact that the sequence of structural shocks ξ is assumed to be strict white noise.

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

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

Assumption 2.2 defines non-anticipating treatment paths conditional on the information available up to time $t-1$. With ξ_t being independent and identically distributed, As-

sumption 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. Indeed, going even further than non-anticipation, Francq and Zakoïan (2010, Thm 11.5) provide the conditions for the MGARCH model to admit a strictly stationary and non-anticipative solution.[3]

Assumption 2.3 (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 proving Corollary 1.1. As we are handling a continuous treatment variable, the overlap assumption is stated for Borel sets which can be chosen as ϵ-neighborhoods of the structural shock of interest $\bar{\xi}_t$: $\mathcal{T}_t = \{\tilde{\xi} \in \mathcal{W} : d(\bar{\xi}_t, \tilde{\xi}) < \epsilon\}$ 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 where a wide range of potential scenario settings is considered. 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 connected Borel sets of conditional covariance matrices $H_{t+h}(\mathcal{T}_t) = \mathcal{H}_{t+h}$, outer return products $X_{t+h}(\mathcal{T}_t) = \mathcal{X}_{t+h}$ and VIRF vectors $V_{t+h}(\mathcal{T}_t)$.

Based on these assumptions, we can show that the VIRF applied to an ϵ-neighborhood $\mathcal{T}_t$ of a structural shock of interest $\bar{\xi}_t$ can be decomposed into a filtered treatment effect,

[3] 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 $\varepsilon_t \perp\!\!\!\perp \sigma(\xi_{t+h}, h > 0)$.

i.e., a filtered causal volatility impulse response, and a selection bias term which vanishes under time series unconfoundedness.

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 $\tilde{\xi}_t \in \mathcal{T}_t = \{\tilde{\xi} \in \mathcal{W} : d(\bar{\xi}_t, \tilde{\xi}) < \epsilon\}$ 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 for the resulting set of VIRFs:*

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

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

Proof. The proof is given in (1.1). □

Remark. The filtered treatment effect is closely related to 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 occurrence 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 artificial base shock $\bar{\xi}_t^*$.[4] By integrating out the effect of the shock $\bar{\xi}_t^*$, where the expectation is calculated from the law of the structural shocks, the integrated causal response function coincides with the VIRF. This gives a causal meaning to the VIRF as a vectorized integrated causal response function applied to the outer product of returns.

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

[4]The VIRF also allows one to compare the causal impacts of two specific shock scenarios. This amounts to calculating the difference of the two respective VIRFs.

3 Empirical Application

We illustrate the merits of the structural VIRF approach by analyzing a system of daily asset returns. It comprises three asset classes which are among the most important for portfolio optimization as well as key ingredients in financial stress scenarios mandated by the European Central Bank (Kremer et al., 2012): equity, fixed income and the foreign exchange markets. To obtain an economically interpretable structural model, we employ the structural proxy-MGARCH approach of Fengler and Polivka (2021).

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 $i = 1, \ldots, n-1$,

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

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

where the product process $(\xi_t Z_{it})_{(t=1,\ldots,T)}$ 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

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

where $u_t = H_t^{-1/2} \varepsilon_t$ denotes the standardized residual, i.e., the return standardized with the principal matrix square root of H_t. One can estimate the full rotation matrix by means of a proxy-guided Givens orthogonalization scheme (Fengler and Polivka, 2021).

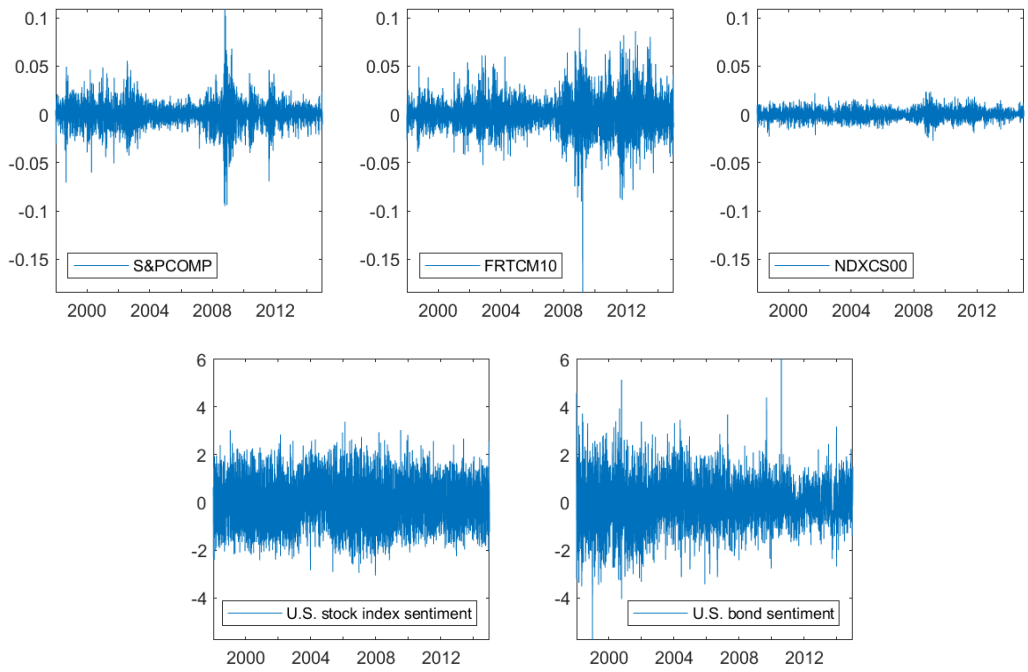

Figure 1: Demeaned daily log returns of the S&P 500 Composite Index (SP500), the yield of the US constant maturity 10-year treasury note (FRTCM10) and the Finex US Dollar Index (NDXCS00) from 1/1/1998 to 12/31/2014.

Consistent estimation of the rotation matrix and asymptotic inference is possible by replacing expectations with their sample mean analogues (Fengler and Polivka, 2021).

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. The three assets consist of the Standard and Poor's (S&P) 500 Composite Index (SP500), the yield of the US constant maturity 10-year treasury note (FRTCM10) and the Finex US Dollar Index (NDXCS00). The daily log returns r_t for each asset are displayed in Figure 1.

With the aim of identifying an equity market and a bond market shock, we use two series of news data taken from Thomson Reuters MarketPsych Indices (TRMI) to proxy for the underlying structural shocks: the US stock index news sentiment and the US bond news

sentiment. The data are provided on a daily frequency, and we fit flexible autoregressive moving average (ARMA) models to them to distill the unexpected innovations. Our choice of these proxy variables is motivated by the well-documented link between news and salient intraday stock returns (see, e.g., Jeon et al., 2021).

Following Fengler and Polivka (2021), we estimate the structural model using a BEKK(1, 1) specification. Table 1 shows that the rotation differs from the identity matrix. 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 second section of Table 1 shows that the stock market index Z_1 and bond market sentiment Z_2 are relevant instruments. Fengler and Polivka (2021, Tables 6-8) conduct a narrative corroboration of the first percentiles of the structural shocks by studying the extent to which large shocks can be related to major financial market news referring to the day of the shock occurrence. Because they can connect each structural shock to specific economic and financial turmoil events, they identify ξ_1 as equity shock, ξ_2 as bond market shock and refer to ξ_3 as currency shock.

		proxy-MGARCH model		
$\hat{\hat{R}}$		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

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 US constant maturity 10-year treasury note (FRTCM10) and the Finex US 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.

3.2 Structural volatility impulse response analysis

One of the most important applications of MGARCH models is the analysis of responses of volatility to shocks (Bauwens et al., 2006). 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 historical key events in our sample: First, as an example of a structural 1% marginal equity tail shock, we consider the NASDAQ crash on April 14, 2000, which accompanied the burst of the Dotcom bubble. Second, we consider the semiannual monetary policy report of the chairman of the Federal Reserve Board, Alan Greenspan, given to the US Senate on March 7, 2002. This is an example of a 1% marginal currency tail and

	Event		
Shock vector component	Dotcom crisis 04/14/2000	Greenspan speech 03/07/2002	EU debt crisis 08/04/2011
Equity	-4.4173	0.7168	-4.3829
Bond	0.0050	-2.5834	0.6824
Currency	1.5779	3.3575	-0.7705
Return vector	Dotcom crisis	Greenspan speech	EU debt crisis
S&P 500	-0.0603	-0.0048	-0.0493
Yield	-0.015	0.0314	-0.0663
USD Index	-0.0044	-0.0101	0.0155

Table 2: Structural shock vectors selected for the historical VIRF analysis with corresponding returns.

bond market tail shock. Third, we examine the concerns in the US equity markets over the European debt crisis on August 4, 2011, as another example of an equity shock. The corresponding shock vectors are documented in Table 2 and the resulting VIRFs are displayed in Figures 2 to 4. The 5% critical values for the simultaneous confidence intervals are obtained from the $\chi^2(3)$.

The structural shock of the NASDAQ crash on 04/14/2000 leads to a strong positive response in the predicted variance of the S&P 500 which is statistically significant at the 5% level (see Figure 2). The increase in the level of volatility of the S&P 500 is significant up to a duration of approximately 75 days. Although there is a strong economic reaction in the covariance of the S&P 500 and the 10-year treasury yield as well, this impulse response appears to be insignificant. Overall, the structural model indicates that the NASDAQ crash on 04/14/2000 has a strong medium-term impact on the level of volatility of the equity market with other (co-)volatilities being largely unaffected. Indeed, our MGARCH model displays elevated equity volatility levels after the NASDAQ crash, as also found

by Sornette et al. (2018).

On 03/07/2002, the chair of the Federal Reserve Board, Alan Greenspan, made an unusual appearance in Congress, which took markets by surprise: He offered a much more upbeat discussion of the economic outlook in his speech to the Senate than in his testimony to the House of Representatives only seven days before. The structural shock associated with this event exhibits a complex set of different impacts on the volatilities of the return system (see Figure 3). First of all, even though the equity component of the shock is close to zero, we observe a small but statistically significant decrease in the predicted variance of the S&P 500. This effect lasts for almost 60 days before turning insignificant and converging to zero. Thus, the news of the accelerated recovery of the US economy forecast a calming effect on equity markets according to the VIRF.

At the same time, the strong negative shock to the bond market[5] leads to a high-magnitude short-term increase in the predicted yield volatility, which stays significantly different from zero for almost 30 days. The negative bond market shock may reflect markets' anticipations of rate hikes. The reaction of treasury yields and their rising volatility in response to Federal Open Market Committee (FOMC) announcements and monetary policy reports has been studied extensively in the literature (see, e.g., Jones et al., 1998, Bomfim, 2003, Rosa, 2013, and the references therein). Remarkably, without taking into account the confidence intervals, the VIRF plot could give the misleading impression of a long-term positive effect on the yield variance.

The predicted covariance impulse responses of the S&P 500 with the yield and of the S&P 500 with the USD Index are both insignificant. In contrast, the structural shock has a significant impact on the predicted covariance of the USD Index with the yield and on

[5] A negative bond market shock leads to decreases in bond prices and increases in bond yields.

the predicted variance of the USD Index itself. In the identified system, positive currency shocks, as present in the structural shock vector of 03/07/2002, tend to be related to news of a weakening USD (Fengler and Polivka, 2021). Greenspan's announcement was no exception: On the day of his testimony, the USD sank against the Yen and the Euro (see Table 2). In contrast, the yield return spiked in response to the negative bond market shock. This negative association gives rise to an immediate and long-lasting decrease in the predicted covariance of the USD Index and the yield in our model. The decrease may seem puzzling at first sight because standard exchange rate models may suggest a positive association between yields and the domestic currency (Hofmann et al., 2020), but the VIRF does not allow for conclusions regarding the levels of (co-)volatilities. However, an inspection of the MGARCH model indicates that the estimated covariance of the yield and the USD index is indeed positive and above the long term mean on the preceding day as well as on the day of the shock occurrence. Thus, the structural shock merely lowers this level rather than strengthening the association further. Finally, we observe a strong positive response of the predicted variance of the USD index to the structural shock, which lasts over a horizon of almost 100 days. Mueller et al. (2017) provide empirical findings on higher foreign exchange volatility upon FOMC news in the short term. For this event, the VIRF signals an increased forex (FX) volatility in response to Greenspan's speech even over a long-term period.

Turning to the structural shock of the European debt crisis (see Figure 4), one observes an immediate, large positive impact on the predicted variance of the S&P 500 lasting for approximately 30 days. The predicted elevated level of volatility is consistent with heightened market fears of European debt defaults. With the exception of the equity market shock, the other components of the structural shock on 08/04/2011 are well within one standard deviation around the zero mean (see Table 2). However, the shock has strong, statistically significant effects on the predicted (co-)variances of all asset return compo-

nents: The historical VIRF of the 10-year treasury yield is significantly different from zero for a time period of approximately 45 days. The other VIRFs display medium- to long-term impacts of between 70 and 90 days. The sources for these large significant responses can be found in the data: On the day of the shock occurrence, the covariance of the yield with the S&P 500 and the variance of the yield itself are well above the long-term mean. The covariances of the USD Index with the other assets are well below their long-term means. Hence, the comparably small shock components are strongly magnified through the large entries in the structural transmission matrix $H_t^{1/2}\tilde{R}$ – resulting in significant volatility responses for all assets. The upward jump in the predicted variance of the yield and in its forecasted covariance with the S&P 500 may mirror movements in US treasuries due to flight-to-safety investments in view of the falling equity markets and the European debt crisis. Furthermore, the VIRF suggests a significant and strong decrease in the predicted covariance of the S&P 500 and the USD index. This finding is consistent with the literature documenting that volatility spillover effects between equity and foreign exchange markets are small in normal market times but strong in periods preceding crises (see, e.g., Grobys, 2015, or Cenedese and Mallucci, 2016). The even stronger drop in the predicted covariance of the USD Index and the 10-year treasury yield mirrors two effects: a surge of the USD relative to the Euro, which is indeed reflected in the positive USD Index return at the shock date (see Table 2), and a decrease in the yield, potentially through a higher demand for safe-haven investments. Finally, the predicted increase in the USD Index volatility in the wake of the European debt crisis is consistent with potentially higher levels of uncertainty in foreign exchange markets. This uncertainty manifests itself in our data through an increase in the conditional variance of the USD Index in the subsequent quarter. Overall, our historical VIRF analysis emphasizes the importance of considering confidence intervals to be able to appreciate volatility impulse responses. Moreover, our discussion highlights the benefits of using structural shocks for their interpretation. To the practitioner, employing structural VIRFs has immediate benefits in asset allocation

and risk management. While an asset manager may, e.g., use a reduced-form volatility model for portfolio diversification, a structural model allows for understanding whether and to what extent these assets were historically affected by the same economic shocks. Thus, the structural VIRF grants a deeper understanding of the risk to which a portfolio is exposed.

3.2.2 Causal scenario VIRFs

Historical VIRFs are useful for understanding volatility events in hindsight, but beyond that are of limited value because certain events, such as those singled out in Section 3.2, will never occur again. With the help of a structural model, however, the causal effects of 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 manifold of other scenarios is conceivable.

For illustration, here we adopt a risk manager's perspective and investigate the VIRFs of (a) a scenario associated with 1% marginal tail equity shocks; and (b) a scenario associated with marginal 1% tail bond shocks on the out-of-sample date 01/02/2015. Because our causality concept in Section 2.6 involves sets of structural shocks, we estimate the multivariate density of the structural shocks. This allows us to draw observations from the intervals of interest. To model multivariate dependence, we initially experimented with a t-copula, but found only very mild departures from independence. For the sake of simplicity, we therefore decided to proceed with the independence copula. The marginal distributions are fitted nonparametrically with Gaussian kernel density estimators. Subsequently, we independently draw for each scenario $10,000$ observations from the 1% fitted quantile of the shock component of interest and from the full distributional range

of the other components to create marginal tail risk scenarios. This means that while sampling from the estimated marginal quantile of one shock component we do not set the other structural shock components artificially to zero. To illustrate the responses to the set of structural shocks, we compute the pointwise median VIRFs and the pointwise 25% and 75% quantile VIRFs at every forecast horizon h. To visualize the parameter estimation uncertainty, we add the analytical confidence intervals of the median target VIRF which we calculate following Fry and Pagan (2011).

Figure 5 displays the VIRFs resulting from shock scenario (a). According to our analysis in Section 2.6, we can interpret these responses as the causal effects associated with the 1% marginal equity shock quantile. We detect a pronounced positive median impact on the forecasted conditional (co-)variance of the S&P 500 and the treasury yield and the variance of the USD Index. Judged by the impact of the median target VIRF, the effect on the S&P 500 is statistically significant at the 5% level for approximately 70 days, the effect on the yield and on its covariance with the S&P 500 is statistically significant for approximately 100 respectively 35 days. The covariance impulse responses of the Finex US Dollar Index to the equity tail events do not exhibit a clear sign, nor do they indicate a clear form of possible outcome paths. The confidence intervals of the median target VIRFs of the predicted covariance of the USD Index with the yield and the predicted variance of the USD Index itself, however, indicate a slight negative impact. To summarize, the shock causes an increase in uncertainty levels and in correlations between equity and fixed income markets over the upcoming medium term.

In contrast, in Figure 6, the sampled 1% quantile bond shocks of scenario (b) exhibit only a very small impact on the predicted conditional covariance of the treasury yield and the S&P 500 and almost none on the forecasted conditional variance of the S&P 500. Remarkably, there seems to be a strong positive effect on the predicted variance of the yield which

is, however, statistically not significant. In contrast, the effect on the predicted conditional covariance with the Finex US Dollar Index is statistically significant for the median target VIRF over a period of almost 25 days. Thus, the increase in covariance between fixed income and fx markets caused by the shock may be of primary concern for risk assessment. In summary, Figures 5 and 6 provide, depending on the shock scenario, very different predictions regarding (co-)variance levels in the asset return system can be expected to rise and to what extent – thus allowing for differential risk managerial precautions.

4 Conclusion

In this paper, we made a fresh examination of the VIRF of Hafner and Herwartz (2006), which is a valuable device for analyzing the impact of shocks on conditional variance matrices in MGARCH models. By deriving the asymptotic law of the VIRF in the BEKK model, we refine the VIRF analysis by providing asymptotic confidence intervals. 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. 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 counterfactual and out-of-sample scenario analyses. Moving 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 application to an identified system of equity, government bond and foreign exchange returns, we discuss two obvious use cases of the structural VIRF. All our applications demonstrate that it is vital to be able to assess the statistical significance of volatility impulse responses.

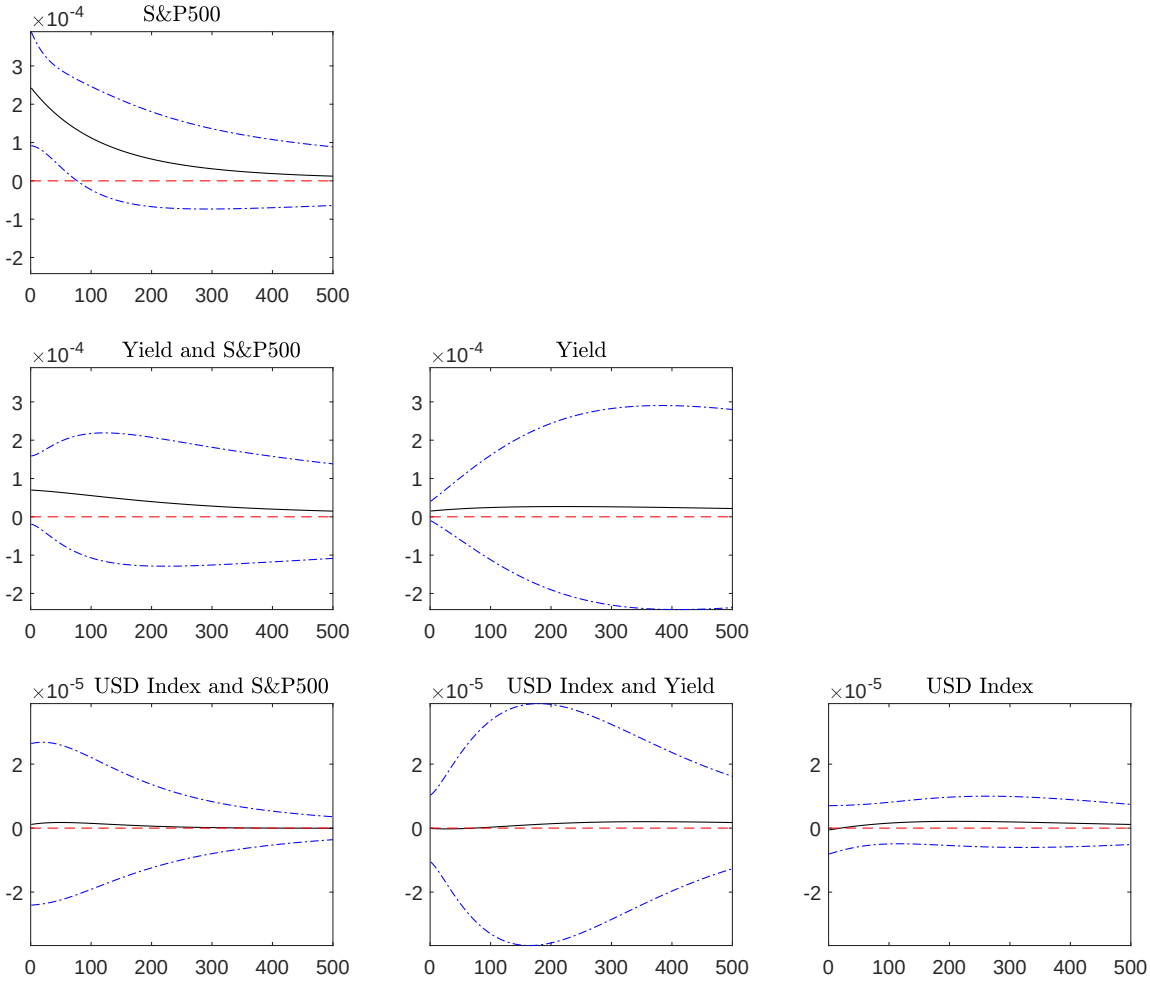

Figure 2: Historical VIRFs in response to the Dotcom bubble burst.

The predicted 500-step ahead VIRFs (black) with 95% confidence intervals (blue) are driven by the structural shock on 04/14/2000, see Table 2. The return system consists of the S&P 500 Composite Index (S&P500), the yield of the US constant maturity 10-year treasury note (Yield) and the Finex US Dollar Index (USD Index).

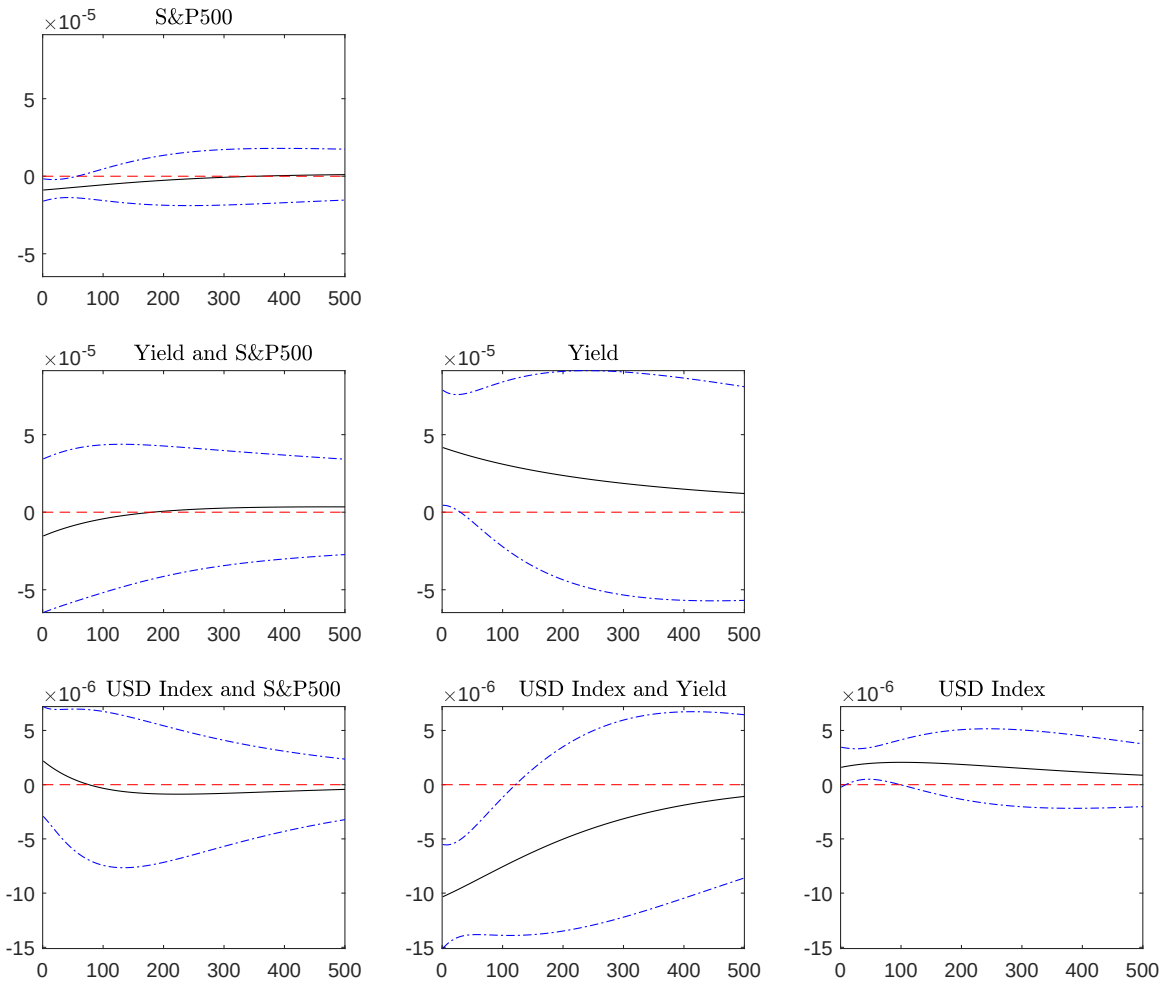

Figure 3: Historical VIRFs in response to the Greenspan testimony.

The predicted 500-step ahead VIRFs (black) with 95% confidence intervals (blue) are driven by the structural shock on 03/07/2002, see Table 2. The return system consists of the S&P 500 Composite Index (S&P500), the yield of the US constant maturity 10-year treasury note (Yield) and the Finex US Dollar Index (USD Index).

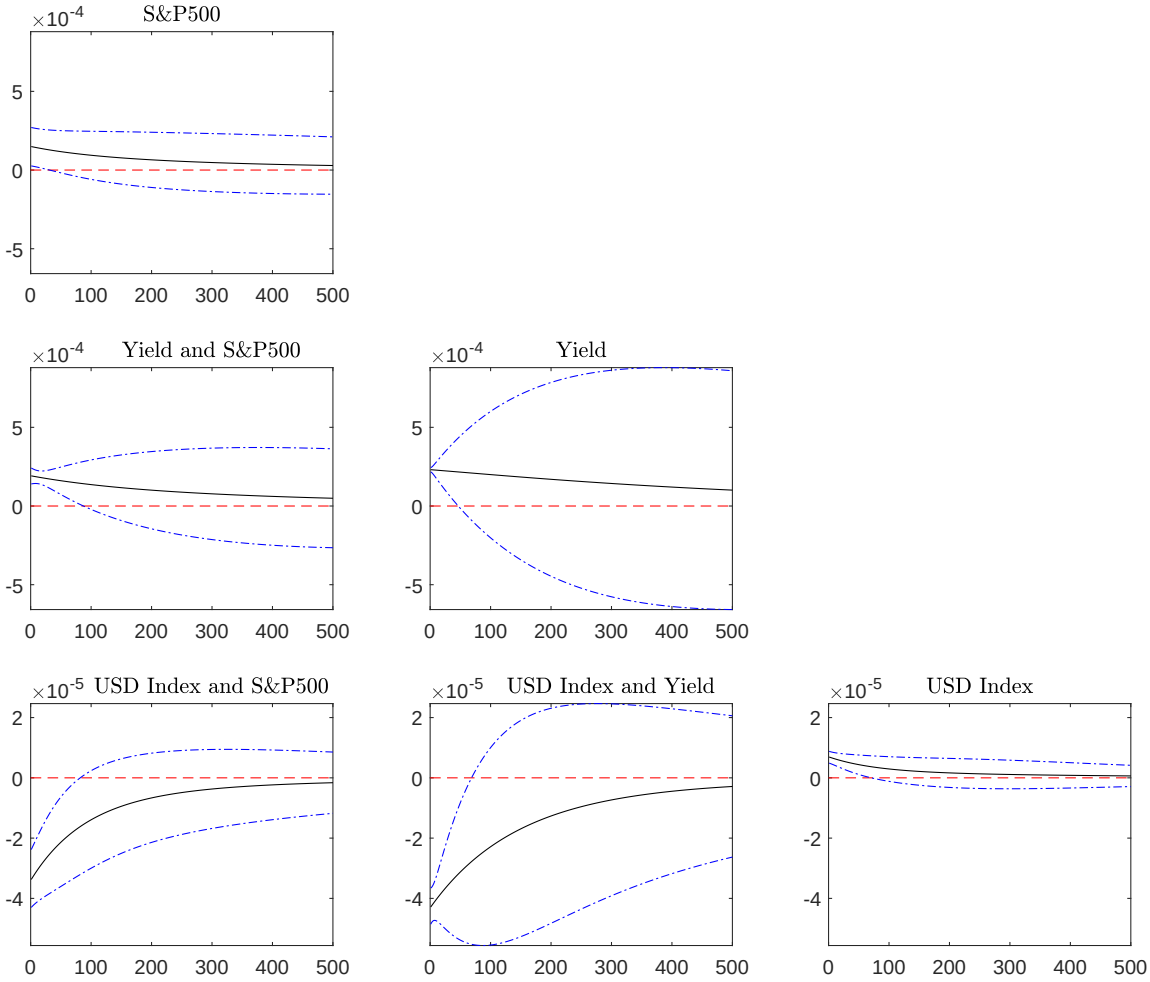

Figure 4: Historical VIRFs in response to the EU debt crisis.

The predicted 500-step ahead VIRFs (black) with 95% confidence intervals (blue) are driven by the structural shock on 08/04/2011, see Table 2. The return system consists of the S&P 500 Composite Index (S&P500), the yield of the US constant maturity 10-year treasury note (Yield) and the Finex US Dollar Index (USD Index).

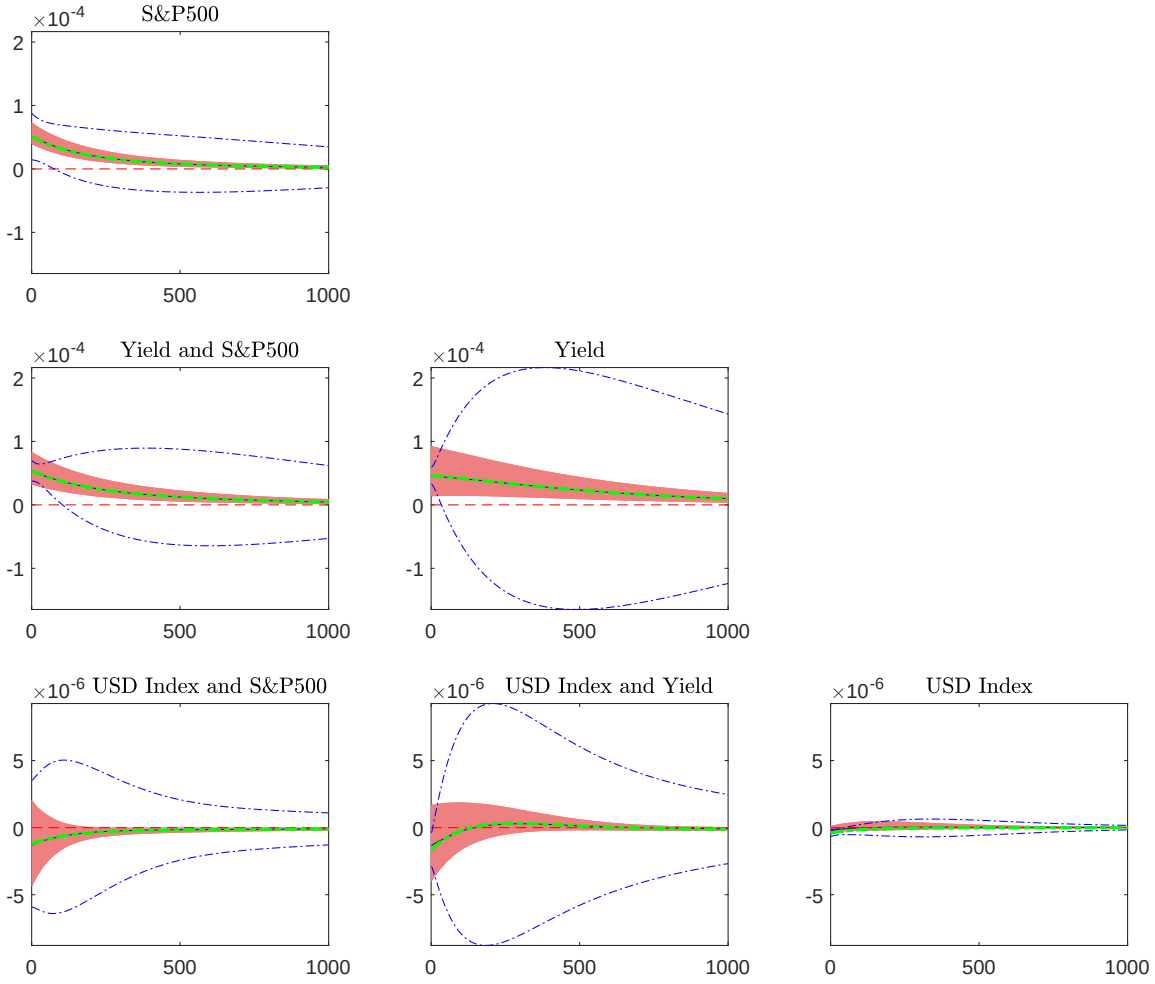

Figure 5: Scenario VIRFs in response to equity tail shocks.

Predicted 500-step ahead VIRFs in response to a scenario family of 1% structural equity tail shocks on the out-of-sample date 01/02/2015: median scenario VIRF (black) with pointwise 25% and 75% quantiles (salmon) and corresponding median target VIRF (green) with analytical confidence intervals (blue). The return system consists of the S&P 500 Composite Index (S&P500), the yield of the US constant maturity 10-year treasury note (Yield) and the Finex US Dollar Index (USD Index).

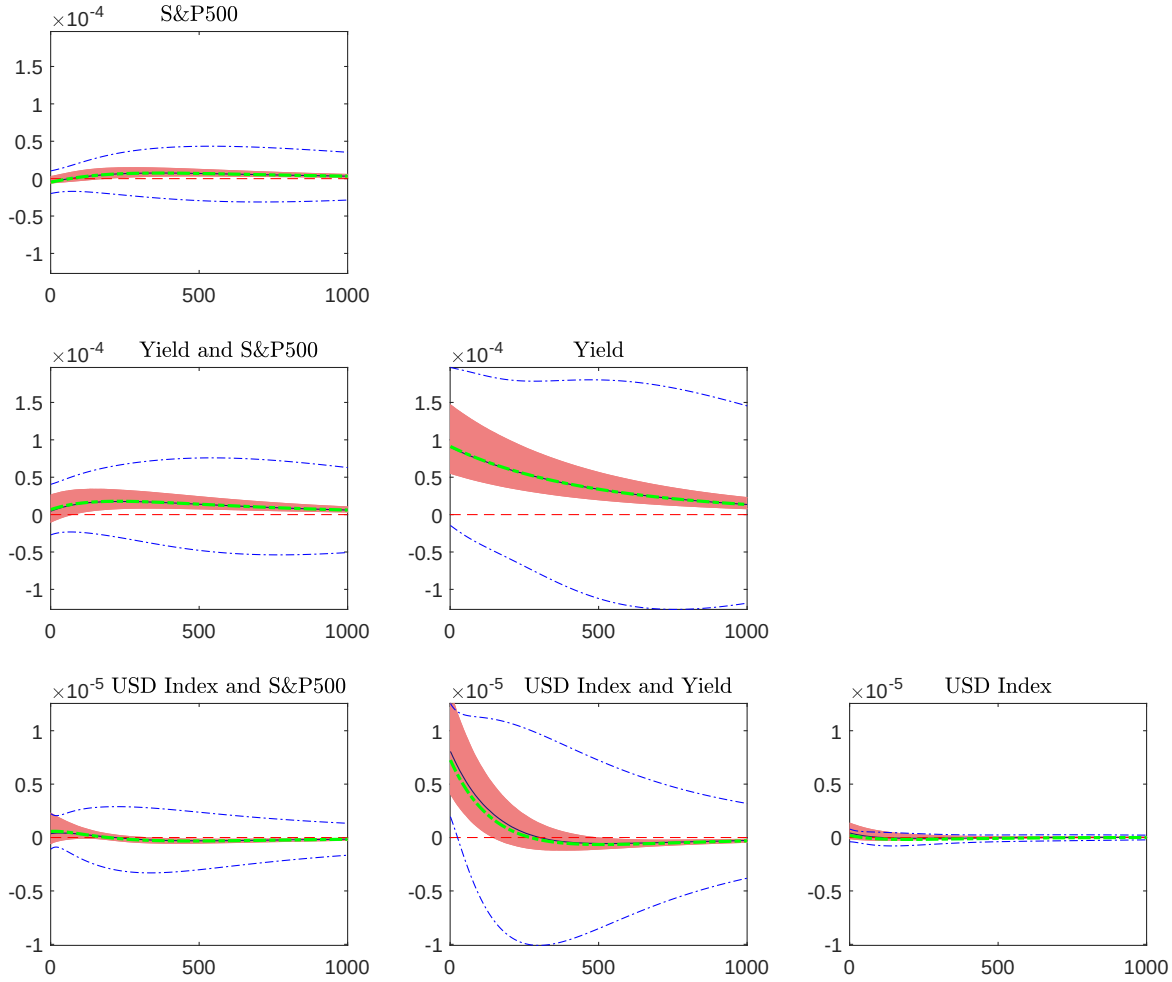

Figure 6: Scenario VIRFs in response to bond tail shocks.

Predicted 500-step ahead VIRFs in response to a scenario family of 1% structural bond tail event shocks on the out-of-sample date 01/02/2015: median scenario VIRF (black) with pointwise 25% and 75% quantiles (salmon) and corresponding median target VIRF (green) with analytical confidence intervals (blue). The return system consists of the S&P 500 Composite Index (S&P500), the yield of the US constant maturity 10-year treasury note (Yield) and the Finex US Dollar Index (USD Index).

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.

Bomfim, A. N. (2003). Pre-announcement effects, news effects, and volatility: Monetary policy and the stock market, *Journal of Banking and Finance* **27**(1): 133–151.

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.

Cavicchioli, M. (2019). Fourth moment structure of Markov switching multivariate GARCH models, *Journal of Financial Econometrics* **19**(4): 565–582.

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.

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.

Fry, R. and Pagan, A. (2011). Sign restrictions in structural vector autoregressions: A critical review, *Journal of Economic Literature* **49**(4): 938–960.

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

Granger, C. W. J. (1969). Investigating causal relations by econometric models and cross-spectral methods, *Econometrica* **37**(3): 424–438.

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. (2022). Identification of structural multivariate GARCH models, *Journal of Econometrics* **227**(1): 212–227. Annals Issue: Time Series Analysis of Higher Moments and Distributions of Financial Data.

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

Hofmann, B., Shim, I. and Shin, H. S. (2020). Bond risk premia and the exchange rate, *Journal of Money, Credit and Banking* **52**(S2): 497–520.

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.

Jeon, Y., McCurdy, T. H. and Zhao, X. (2021). News as sources of jumps in stock returns: Evidence from 21 million news articles for 9000 companies, *Journal of Financial Economics* **145**(2): 1–17.

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.

Jones, C. M., Lamont, O. and Lumsdaine, R. L. (1998). Macroeconomic news and bond market volatility, *Journal of Financial Economics* **47**(3): 315–337.

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*

Lechner, M. (2010). The relation of different concepts of causality used in time series and microeconometrics, *Econometric Reviews* **30**(1): 109–127.

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.

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

Lütkepohl, H. (2005). *New Introduction to Multiple Time Series Analysis*, Springer.

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.

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?
URL: *https://scholar.harvard.edu/shephard/publications/nonparametric-dynamic-causal-model-macroeconometrics*

Rosa, C. (2013). The financial market effect of FOMC minutes, *Economic Policy Review* (December Issue, Federal Reserve Bank of New York): 67–81.

Rubin, D. B. (1980). Randomization analysis of experimental data: The Fisher randomization test, *Journal of the American Statistical Association* **75**(371): 591–593.

Sims, C. A. (1972). Money, income, and causality, *American Economic Review* **62**: 540–552.

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

Sornette, D., Cauwels, P. and Smilyanov, G. (2018). Can we use volatility to diagnose financial bubbles? Lessons from 40 historical bubbles, *Quantitative Finance and Economics* **2**(1): 486–590.

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 Notation and results from matrix algebra

Definition D.1. Principal matrix square root. 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).

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

Definition D.3. $\text{vec}(\cdot)$ operator. The $\text{vec}(\cdot)$ operator stacks, starting with the first column, the columns of an $(n \times n)$ matrix M in an n^2-dimensional vector. It is a linear operator.

Definition D.4. vech$(\cdot)$ operator. The vech$(\cdot)$ operator stacks, starting with the first column, the lower triangular part of a (symmetric) $(n \times n)$ matrix M in an n^*-dimensional vector where $n^* = \frac{n(n+1)}{2}$.

Definition D.5. Moore-Penrose inverse. The Moore-Penrose inverse of an $(m \times n)$ matrix M with $M^\top M$ non-singular is defined as

$$M^+ = (M^\top M)^{-1} M^\top. \tag{16}$$

Definition D.6. Duplication matrix. For any symmetric $(n \times n)$ matrix M, the duplication matrix D_n denotes the unique $\left(n^2 \times \frac{n(n+1)}{2}\right)$ matrix such that

$$\text{vec}(M) = D_n \, \text{vech}(M). \tag{17}$$

The Moore-Penrose inverse of the duplication matrix is denoted by D_n^+.

Definition D.7. Commution matrix. For every $(m \times n)$ matrix M, the $(mn \times mn)$ commutation matrix K_{mn} is defined by

$$K \, \text{vec}(M) = \text{vec}(M^\top). \tag{18}$$

For $n = m$ we abbreviate $K_{nn} = K_n$.

Result R.1. vec$(\cdot)$ operations. For appropriately defined matrices A, B and C and for some $(n \times n)$ matrices M, P:

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

$$\text{vec}(A^\top \otimes A^\top) = \text{vec}((A \otimes A)^\top) \tag{20}$$

$$\text{vec}(M \otimes P) = (I_n \otimes K_n \otimes I_n) \left[\text{vec}(M) \otimes \text{vec}(P)\right] \tag{21}$$

see (Magnus and Neudecker, 1988, Theorem 3.10) for a proof of (21).

Result R.2. Matrix derivatives results.

1) For $n \times n$ matrices X and Z, Z symmetric, it holds:

$$\frac{\partial \, \text{vec}(XZX)}{\partial \, \text{vec}(X)^\top} = (XZ \otimes I_n) + (I_n \otimes XZ) \tag{22}$$

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

2) For any symmetric, positive semidefinite $n \times n$ matrix H with principal square root $H^{1/2}$, it holds:

$$\frac{\partial \, \text{vec}(H^{1/2})}{\partial \, \text{vec}(H)^\top} = \left[\left(I_n \otimes H^{1/2}\right) + (H^{1/2} \otimes I_n)\right]^{-1} \tag{23}$$

which can be derived by solving a Sylvester type equation for the differential.

3) For any $(n \times n)$ matrices M and P, it holds:

$$\frac{\partial \operatorname{vec}(M) \otimes \operatorname{vec}(P)}{\partial \operatorname{vec}(M)^\top} = I_n \otimes \operatorname{vec}(P). \tag{24}$$

4) For any $(n \times n)$ matrix M:

$$\frac{\partial \left(\operatorname{vec}(M) \otimes \operatorname{vec}(M)\right)}{\partial \operatorname{vec}(M)^\top} = \left(I_n \otimes \operatorname{vec}(M)\right) + \left(\operatorname{vec}(M) \otimes I_n\right) \tag{25}$$

by (24) and the product rule.

5) For any $(n \times n)$ matrix M:

$$\frac{\partial \operatorname{vec}(M^\top)}{\partial \operatorname{vec}(M)^\top} = K_n. \tag{26}$$

A.2 Mathematical prerequisites for the VIRF

Corollary 1.2. *Let $(Y_t)_{(t\in\mathbb{N}_0)}$ be a sequence of integrable random vectors in $\mathbb{R}^k$ defined on a probability space $(\Omega, \mathcal{F}, \mathrm{P})$, 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 $\mathbb{R}^{k\times k}$. Then $\sum_{i=0}^{\infty} \phi_i Y_{t-i}(\omega)$ exists and $\mathrm{E}[\sum_{i=0}^{\infty} \phi_i Y_{t-i}|\tilde{\mathcal{F}}] = \sum_{i=0}^{\infty} \phi_i \mathrm{E}[Y_{t-i}|\tilde{\mathcal{F}}]$. Thus we can interchange infinite summation and conditional expectation.*

Proof. The proof is available from the authors on request. □

Proof of Proposition 2.1. Let $X_t = \operatorname{vech}(\varepsilon_t \varepsilon_t^\top)$ and $Y_t = X_t - \operatorname{vech}(H_t)$ and $h \geq 1$. Then we have

$$\mathrm{E}[\operatorname{vech}(H_{t+h})|\mathcal{F}_{t-1}] = \mathrm{E}[\mathrm{E}[X_{t+h}|\mathcal{F}_{t+h-1}]|\mathcal{F}_{t-1}] = \mathrm{E}[X_{t+h}|\mathcal{F}_{t-1}]. \tag{27}$$

Together with the VMA(∞) representation we obtain for the VIRF in (3)

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

By square integrability of $(\varepsilon_t)_{t\in\mathbb{Z}}$, $(X_t)_{t\in\mathbb{Z}}$ is a sequence of integrable random variables, and because the series $\operatorname{vech}(H_t)_{t\in\mathbb{Z}}$ is integrable by the square integrability of ε_t, so is $(Y_t)_{t\in\mathbb{Z}}$. Because $\operatorname{Var}(Y_t) = H_Y < \infty$, the absolute moments of Y_t are uniformly bounded. Hence, by Corollary 1.2, we can interchange infinite summation and conditional expectation:

$$\begin{aligned} \mathrm{V}_{t+h}(\xi_t; \eta) &= \sum_{i=0}^{\infty} \Psi_i \left(\mathrm{E}\left[Y_{t+h-i}|\tilde{\mathcal{F}}_t\right] - \mathrm{E}[Y_{t+h-i}|\mathcal{F}_{t-1}]\right) \\ &= \Psi_h \left(\mathrm{E}\left[Y_t|\tilde{\mathcal{F}}_t\right] - \mathrm{E}[Y_t|\mathcal{F}_{t-1}]\right). \end{aligned} \tag{29}$$

This follows from $\mathrm{E}\left[Y_{t+h-i}|\tilde{\mathcal{F}}_t\right] - \mathrm{E}[Y_{t+h-i}|\mathcal{F}_{t-1}] = 0$ for all $(t+h-i) \leq t-1$ due to measurability given $\mathcal{F}_{t-1}$ and $\mathrm{E}\left[Y_{t+h-i}|\tilde{\mathcal{F}}_t\right] - \mathrm{E}[Y_{t+h-i}|\mathcal{F}_{t-1}] = 0$ 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 ξ_t, $\mathrm{E}[Y_{t+h-i}|\tilde{\mathcal{F}}_t] = 0$. Using predictability[6] and tower property arguments and inserting the structural model (2), we obtain:

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

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

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

□

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

$$X_t = \mathrm{vech}(H) + \sum_{i=0}^{\infty} \Psi_i Y_{t-i} \tag{32}$$

where $\mathrm{vech}(H) = \Phi(1)^{-1}c$ *denotes the long-run covariance matrix with* $c = \mathrm{vech}(CC^\top)$ *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. Transform (4) into its equivalent VEC representation:

$$\mathrm{vec}(H_t) = \mathrm{vec}(CC^\top) + \sum_{i=1}^{p}\mathrm{vec}(A_i^\top\varepsilon_{t-i}\varepsilon_{t-i}^\top A_i) + \sum_{j=1}^{q}\mathrm{vec}(B_j^\top H_{t-j}B_j). \tag{33}$$

[6]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.

Using (19) and exploiting (17) we get:

$$\begin{aligned} D_n \operatorname{vech}(H_t) = D_n \operatorname{vech}(CC^\top) &+ \sum_{i=1}^{p} \left(A_i \otimes A_i\right)^\top D_n \operatorname{vech}(\varepsilon_{t-i}\varepsilon_{t-i}^\top) \\ &+ \sum_{j=1}^{q} \left(B_j \otimes B_j\right)^\top D_n \operatorname{vech}(H_{t-j}). \end{aligned} \tag{34}$$

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

$$\begin{aligned} \operatorname{vech}(H_t) = \underbrace{\operatorname{vech}(CC^\top)}_{=:c} &+ \sum_{i=1}^{p} \underbrace{D_n^+ \left(A_i \otimes A_i\right)^\top D_n}_{=:\tilde{A}_i} \operatorname{vech}(\varepsilon_{t-i}\varepsilon_{t-i}^\top) \\ &+ \sum_{j=1}^{q} \underbrace{D_n^+ \left(B_j \otimes B_j\right)^\top D_n}_{=:\tilde{B}_j} \operatorname{vech}(H_{t-j}) \end{aligned} \tag{35}$$

To derive the equivalent VARMA($\max(p,q),q$) representation as in Hafner and Herwartz (2006), set $X_t := \operatorname{vech}(\varepsilon_t\varepsilon_t^\top)$ and $Y_t := X_t - \operatorname{vech}(H_t)$. Y_t is a weak white noise process with $\mathrm{E}[Y_t] = 0$, $\mathrm{Var}(Y_t) = H_Y$ and $\mathrm{E}[Y_t Y_s^\top] = 0$ $(t \neq s)$. Rearranging (35) yields:

$$X_t = c + \sum_{i=1}^{\max(p,q)} \left(\tilde{A}_i + \tilde{B}_i\right) X_{t-i} - \sum_{j=1}^{q} \tilde{B}_j Y_{t-j} + Y_t \tag{36}$$

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

$$\begin{aligned} \underbrace{\left(I_{n^*} - \sum_{i=1}^{\max(p,q)} \left(\tilde{A}_i + \tilde{B}_i\right) L^i\right)}_{=:\Phi(L)} X_t &= c + \underbrace{\left(I_{n^*} - \sum_{j=1}^{q} \tilde{B}_j L^j\right)}_{=:\Theta(L)} Y_t \\ \Leftrightarrow X_t &= \Phi(1)^{-1} c + \underbrace{\Phi(L)^{-1}\Theta(L)}_{=:\Psi(L)} Y_t \\ &= \operatorname{vech}(H) + \sum_{i=0}^{\infty} \Psi_i Y_{t-i}. \end{aligned} \tag{37}$$

The $(n^* \times n^*)$ coefficient matrices Ψ_i are determined recursively by coefficient matching (Lütkepohl, 2005). □

A.3 Asymptotic theory for VIRFs

Proof of Theorem 1. The result follows by an application of the Delta method together with the asymptotic normality of the QML estimator. The Jacobian can be derived as follows: Let $\eta \in \mathbb{R}^m$ denote the vector of stacked parameters of the BEKK(p,q) model:

$\eta = \left(\text{vec}(C)^\top, \text{vec}(A_1)^\top, \ldots, \text{vec}(A_p)^\top, \text{vec}(B_1)^\top, \ldots, \text{vec}(B_q)^\top\right)^\top$. Starting with the definition of the VIRF for BEKK(p,q) models, we have:

$$\text{V}_{t+h}(\xi_t;\eta) = \Psi_h D_n^+ \left(\text{vec}(H_t^{1/2}\tilde{R}\xi_t\xi_t^\top\tilde{R}^\top(H_t^{1/2})^\top) - \text{vec}(H_t)\right) \tag{38}$$

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 (22) and (23). Moreover, the derivatives of the BEKK model with regard to its parameters, i.e., $\frac{\partial \text{vec}(H_t)}{\partial \eta^\top}$ can be found in Hafner and Herwartz (2008), such that $\frac{\partial \text{vec}(H_t)}{\partial \eta^\top}$ is known. Then, for the VIRF at time $h = 0$:

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

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

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

Moreover, we can establish the recursion

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

To evaluate this expression, we derive $\frac{\partial \text{vec}\left(\tilde{A}_j\right)}{\partial \eta^\top}$, $(j = 1,\ldots,h)$. The derivations for $\tilde{B}_j$ follow analogously. To this end, we insert the definition of $\tilde{A}_j$ and make use of (19), (20)

and (21) to make the following transformations:

$$\begin{aligned}\frac{\partial \operatorname{vec}\left(\tilde{A}_j\right)}{\eta^\top} &= \frac{\partial \operatorname{vec}\left(D_n^+ \left(A_j \otimes A_j\right)^\top D_n\right)}{\partial \eta^\top} \\ &= \frac{(D_n^\top \otimes D_n^+)\partial \operatorname{vec}\left(A_j \otimes A_j\right)^\top}{\partial \eta^\top} \\ &= \frac{(D_n^\top \otimes D_n^+)\partial \operatorname{vec}\left(A_j^\top \otimes A_j^\top\right)}{\partial \eta^\top} \\ &= \left(D_n^\top \otimes D_n^+\right)\left(I_n \otimes K_n \otimes I_n\right) \frac{\partial \left(\operatorname{vec}(A_j^\top) \otimes \operatorname{vec}(A_j^\top)\right)}{\partial \eta^\top}.\end{aligned}$$

Finally, by an application of (25) and (26) it holds:

$$\frac{\partial \left(\operatorname{vec}(A_j^\top) \otimes \operatorname{vec}(A_j^\top)\right)}{\partial \eta^\top} = \left(K_n \otimes \operatorname{vec}(A^\top)\right) + \left(\operatorname{vec}(A^\top) \otimes K_n\right) \tag{39}$$

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:

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

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. □

A.4 Causality of the VIRF

Proof of Corollary 1.1.

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

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