Università Politecnica delle Marche
DIPARTIMENTO DI ECONOMIA

Does macroeconomics help in predicting stock markets volatility comovements? A nonlinear approach

Andrea Bucci Giulio Palomba
Eduardo Rossi

QUADERNI DI RICERCA n. 440

ISSN: 2279–9575

Ottobre 2019

Abstract

This paper addresses the question of the relevance of macroeconomic determinants in forecasting the evolution of stock markets volatilities and co-volatilities. Our approach combines the Cholesky decomposition of the covariance matrix with the use of the Vector Logistic Smooth Transition Autoregressive Model. The model includes predetermined variables and takes into account the asymmetries in volatility process. Structural breaks and nonlinearity tests are also implemented to determine the number of regimes and to identify the transition variables. The model is applied to realized volatility of stock indices of several countries in order to evaluate the role of economic variables in predicting the future evolution of conditional covariances. Our results show that the forecast accuracy of our model is significantly different from the accuracy of the forecasts obtained via other standard approaches.

JEL Class.: C32, C58, G11, G17.

Keywords: Multivariate realized volatility, Non-linear models, Smooth transition, Forecast evaluation, Portfolio optimization

Indirizzo: Andrea Bucci, Dipartimento di Scienze Economiche e Sociali, Università Politecnica delle Marche, Ancona, Italy. E-mail: a.bucci@univpm.it; Giulio Palomba. Corresponding author. Dipartimento di Scienze Economiche e Sociali, Università Politecnica delle Marche, Ancona, Italy. E-mail: g.palomba@univpm.it, Tel.: +39 071 220 7112; Eduardo Rossi, Dipartimento di Scienze Economiche ed Aziendali, University of Pavia, 27100 Pavia, Italy. E-mail: eduardo.rossi@unipv.it

Does macroeconomics help in predicting stock markets volatility comovements? A nonlinear approach*

Andrea Bucci *Giulio Palomba* *Eduardo Rossi*

1 Introduction

Understanding how financial volatility and co-volatilities react to changes in macroeconomic and financial conditions is critical for investors and financial institutions. The study of volatility and co-volatility dynamics contributes to the knowledge of the links through which the macroeconomic and financial shocks propagate across markets and asset classes. It is worth noticing that the existing literature on volatility determinants focuses almost exclusively on univariate volatility and linear models. In a first attempt, Schwert (1989) analyses the relation between volatility and the level of economic activity, reporting little effects of macroeconomic variables on volatility dynamics. More recently, Mele (2007) suggests that determinants of the time-varying risk premium are viable candidates for volatility forecasting. Paye (2012) further shows that volatility is persistent and countercyclical, while the predictive ability of the exogenous variables seems to be poor in his findings. Christiansen, Schmeling, and Schrimpf (2012) extends this work, including a larger set of predictors. Their linear framework underlines the utility of financial variables as volatility determinants, while macroeconomic variables seem not able to produce superior out of sample forecasts. In this paper, we contribute to this literature by extending the analysing to the relationship between volatilities and co-volatilities and macroeconomic and financial variables in a multivariate framework.

Our analysis investigates the potential of macroeconomic variables in predict-

* *We wish to thank all the participants of the Eighth Italian Congress of Econometrics and Empirical Economics (ICEEE) held in Lecce from January 24 to 26, 2019.*

ing realized volatilities and covariances of major stock market indexes. We propose a new nonlinear specification of volatilities and co-volatilities that links macroeconomic and financial factors to the multivariate volatility process. In this way, we are able to investigate the spillover effects of macroeconomic and financial shocks among markets. This is in line with the growing interest in the nature of volatility spillovers between markets. Diebold and Yilmaz (2009), for example, provide evidence of volatility spillovers during financial crisis, while Beirne, Caporale, Schulze-Ghattas, and Spagnolo (2009) show that major markets volatility affects conditional variances in many emerging markets and that this behaviour is time-varying. Xiong and Han (2015), on the contrary, analyse the volatility spillover effect between stock market and foreign exchange market, proving that the nature of the spillover effect is bi-directional and asymmetric, since a change in foreign market has a larger effect on stock market respect to the inverse relation.

Relying on the results of the literature on ex-post volatility measurement based on high-frequency prices (see Ait-Sahalia and Jacod, 2014), we construct monthly realized measures using low-frequency daily data, which are immune from microstructure noise. However, low frequency prices (e.g. monthly frequency) are characterized by discretization error which is inevitably larger than the one generated with high frequency sampling. Recently, Bollerslev, Patton, and Quaedvlieg (2017) have proposed dynamic attenuation models to limit the impact of heteroskedastic measurement errors on the parameter estimation. They introduce dynamic specifications for the high-frequency realized covariances, in which the autoregressive parameters of the models depend linearly on the measurement errors of covariance matrix estimates. Since we model the Cholesky factors of the monthly realized covariance matrices in a nonlinear set up, this strategy cannot be implemented.

A few nonlinear models have been proposed in the realized volatility literature. Martens, De Pooter, and Van Dijk (2004) firstly introduce a long memory model with asymmetries and structural breaks for realized volatility, while McAleer and Medeiros (2008) extend this approach, testing for the presence of structural breaks and defining nonlinearity tests. These authors rely on a smooth transition specification and consider the possibility to include external exogenous variables into the model structure. Recently, Bucci (2019a,b) studies whether using a non-linear tool, like neural networks, combined with the use of exogenous variables may help to predict realized volatility in the univariate and multivariate framework.

Building on the existing literature, we propose a new multivariate specification based on the Vector Smooth Transition Autoregressive model with lagged exogenous variables both in linear and nonlinear components (hereafter VLSTAR). This class of models explicitly assumes that the regime switch is determined by an observable switch or transition variable. The use of this kind of models is coherent with the empirical finding that volatility is higher in presence of unexpected news and may lead to a better understanding of the relationships between a set of exogenous variables and the volatility. The forerunner model of this literature, introduced by Quandt (1958), proposed a coefficient changing model related to the values of an observable stochastic variable. Later, several extensions, like the smooth transition model introduced and developed by Chan and Howell (1986), the logistic STAR (LSTAR) by May (1976) and Tong (1990) and the exponential STAR (ESTAR) by Haggan and Ozaki (1981), have been considered. In this paper, we are interested in the multivariate extension of the STAR model and, in particular, in a Vector Logistic Smooth Transition Autoregressive model, firstly appeared in Anderson and Vahid (1998).

We consider a quite large set of financial and economic predictors: monthly industrial production indexes, monthly inflation rates, unemployment rates, oil price, dividend yield and earning price ratio of S&P 500 index, Fama and French factors, and the Economic Policy Uncertainty (EPU) index of Baker, Bloom, and Davis (2016). This latter is a weighted average of three measures of economic policy uncertainty. The first, with the greatest weight, is the frequency of major news discussing economic policy-related uncertainty. Baker, Bloom, and Davis (2016) point out that there exists significant relationship between EPU index and real macroeconomic variables. Pástor and Veronesi (2012) finds that introducing new policies with an uncertain impact increases the volatility of the stochastic discount factor. The increase in the volatility of the discount factor leads to increases in risk premia which in turn result in high volatility in stock market.

Notably, there are many empirical papers that investigate the effects of EPU on stock market return or volatility (see, e.g., Amengual and Xiu, 2018; Antonakakis, Chatziantoniou, and Filis, 2013, Ajmi, Aye, Balcilar, Montasser, and Gupta, 2015, Brogaard and Detzel, 2015, Sum and Fanta, 2012, Johnson and Lee, 2014, Kang and Ratti, 2013, Lam, Zhang, and Zhang, 2019). In this study, we contribute to the literature by examining the role of the EPU as predictor of stock market covolatilities, jointly with other variables, both in a linear and a non linear modelling framework. To the best of our knowledge, this issue has not been addressed in the existing studies..

To assess the validity of our approach, we implement an out of sample forecasting analysis, using data from 1990 to 2018. In the analysis, we consider direct methods, such as Root Mean Square Error (RMSE), Diebold and Mariano, (DM,1995) and Giacomini and White, (GW, 2006) tests and non-direct methods, such as portfolio optimization. Our findings show that the set of macroeconomic and financial variables improve the predictive accuracy of the low frequency realized covariance. As in Paye (2012) and Christiansen, Schmeling, and Schrimpf (2012), purely macroeconomic variables hardly show up as important predictors of financial volatility. The results are robust to a number of direct and non-direct methods for evaluating out of sample forecasts.

The paper is organized as follows. In section 2, we introduce the multivariate volatility model, including the structural breaks and linearity tests performed. Section 3 presents the data set, the model and the estimation results, while Section 4 discusses the evaluation of the out of sample forecasts. Section 5 concludes.

2 The Cholesky-VLSTAR

In this section we introduce the model used to forecast monthly variances and covariances of n risky stocks. As it is well known, the estimate of the covariance matrix must be at least positive definite. Generally, unconstrained methods, such as Cholesky decomposition or matrix logarithmic transformation, are preferred to guarantee parsimony, especially in high-dimensional problems. For example, Halbleib-Chiriac and Voev (2011) implement a VARFIMA model on the elements of the Cholesky decomposition of the covariance matrix, while Bauer and Vorkink (2011) rely on the matrix exponential transformation of the covariance matrix to ensure a positive semi-definite covariance matrix. In this paper, the forecast matrix is guaranteed positive definite through a Cholesky decomposition. Let $r_\tau = \begin{bmatrix} r_{1,\tau} & r_{2,\tau} & \dots & r_{n,\tau} \end{bmatrix}'$ be the n-dimensional column vector of returns between day τ and day $\tau - 1$ calculated as $r_\tau = 100\left(\ln P_\tau - \ln P_{\tau-1}\right)$, where P_τ is the vector of stock prices in the τ-th trading day. Hence, the monthly return is

$$r_t = \begin{bmatrix} r_{1,t} & r_{2,t} & \dots & r_{n,t} \end{bmatrix}' = \sum_{\tau=1}^{N_t} r_\tau,$$

where N_t is the number of trading days in the t-th month. The object of interest is the $n \times n$ conditional covariance matrix of the monthly returns, $Var(r_t \mid I_{t-1}) = \Sigma_t$, where I_{t-1} is the information available at time $t-1$. We estimate the conditional covariance matrix through the realized covariance measure, computed as

$$\mathrm{RC}_t = \sum_{\tau=1}^{N_t} r_\tau r'_\tau \tag{1}$$

for $t = 1, 2, \dots, T$. Recently, Barndorff-Nielsen and Shephard (2002) and Andersen, Bollerslev, Diebold, and Labys (2003) show that RC_t converges in probability to Σ_t, *i.e.* the quadratic variation of the price process, under very general assumptions.

Once obtained the realized covariance matrix, we compute the Cholesky decomposition as

$$\mathrm{RC}_t = Y_t Y'_t, \tag{2}$$

where Y_t is a full rank $n \times n$ lower triangular matrix, and the vectorization of the Cholesky factors is the column vector $y_t = vech(Y_t)$, $\tilde{n} = n(n+1)/2$ elements. Clearly, the use of the Cholesky decomposition allows us to have a positive definite matrix Σ_t without setting any parameter restriction on y_t (see Halbleib-Chiriac and Voev, 2011; Becker, Clements, and O'Neill, 2010).

In order to model the $\tilde{n}$ elements of y_t, we specify the relationship between future market realized variances and covariances and economic predictors as a multivariate smooth transition model, which is an extension of the smooth transition regression model introduced by Bacon and Watts (1971) (see also Anderson and Vahid, 1998). The general model is

$$\begin{aligned} y_t &= \mu_0 + \sum_{j=1}^{p} \Phi_{0,j}\, y_{t-j} + A_0 x_t + G_t(s_t;\gamma,c) \left[\mu_1 + \sum_{j=1}^{p} \Phi_{1,j}\, y_{t-j} + A_1 x_t\right] + \varepsilon_t \\ &= \mu_0 + G_t(s_t;\gamma,c)\mu_1 + \sum_{j=1}^{p} \left[\Phi_{0,j} + G_t(s_t;\gamma,c)\Phi_{1,j}\right] y_{t-j} + \left[A_0 + G_t(s_t;\gamma,c)A_1\right] x_t + \varepsilon_t \end{aligned} \tag{3}$$

where μ_0 and μ_1 are the $\tilde{n} \times 1$ vectors of intercepts, $\Phi_{0,j}$ and $\Phi_{1,j}$ are square $\tilde{n} \times \tilde{n}$ matrices of parameters for lags $j = 1, 2, \ldots, p$, A_0 and A_1 are $\tilde{n} \times k$ matrices of parameters, x_t is the $k \times 1$ vector of exogenous variables and ε_t is the innovation. Finally, $G_t(s_t; \gamma, c)$ is a $\tilde{n} \times \tilde{n}$ diagonal matrix of transition function at time t, such that

$$G_t(s_t; \gamma, c) = \text{diag}\left\{G_{1,t}(s_{1,t}; \gamma_1, c_1), G_{2,t}(s_{2,t}; \gamma_2, c_2), \ldots, G_{\tilde{n},t}(s_{\tilde{n},t}; \gamma_{\tilde{n}}, c_{\tilde{n}})\right\}. \qquad (4)$$

Every scalar transition function $G_{i,t}(s_{i,t}, \gamma_i, c_i)$, with $i = 1, 2, \ldots, \tilde{n}$, is a continuous function of the transition variables $s_{i,t}$ with scale parameter γ_i and threshold c_i.

The model in equation (3) can be written as

$$y_t = \left[\mu_0 + G_t\mu_1 \quad \Phi_{0,1} + G_t\Phi_{1,1} \quad \Phi_{0,2} + G_t\Phi_{1,2} \quad \ldots \quad \Phi_{0,p} + G_t\Phi_{1,p} \quad A_0 + G_tA_1\right] z_t + \varepsilon_t,$$

where $G_t \equiv G_t(s_t; \gamma, c)$ and $z_t = [1 \quad y'_{t-1} \quad y'_{t-2} \quad \ldots \quad y'_{t-p} \quad x'_t]'$ is a $(1 + p\tilde{n} + k) \times 1$ vector containing the constant, the exogenous explanatory variables and all the lags of y_t. Given that G_t is nonsingular, the model becomes

$$y_t = G_tB'z_t + \varepsilon_t, \qquad (5)$$

where $B' = \left[G_t^{-1}\mu_0 + \mu_1 \quad G_t^{-1}\Phi_{0,1} + \Phi_{1,1} \quad G_t^{-1}\Phi_{0,2} + \Phi_{1,2} \quad \ldots \quad G_t^{-1}\Phi_{0,p} + \Phi_{1,p} \quad G_t^{-1}A_0 + A_1\right]$ is $\tilde{n} \times (1 + p\tilde{n} + k)$.

The model can be extended to include $m - 1$ regime changes, in such case (3) becomes

$$\begin{aligned} y_t \quad = \quad & \mu_0 + \sum_{j=1}^{p} \Phi_{0,j}\, y_{t-j} + A_0x_t + G_t(s_t; \gamma, c)\left[\mu_1 + \sum_{j=1}^{p} \Phi_{1,j}\, y_{t-j} + A_1x_t\right] + \ldots \\ & + G_t^{m-1}(s_t; \gamma, c)\left[\mu_m + \sum_{j=1}^{p} \Phi_{m,j}\, y_{t-j} + A_mx_t\right] + \varepsilon_t. \qquad (6) \end{aligned}$$

Let $B_i = [\mu_i \quad \Phi_{i,1} \quad \ldots \quad \Phi_{i,p} \quad A_i]$ be a $(\tilde{n} \times (1 + k + \tilde{n}))$ matrix with $i = 1, \ldots, m$, the linear equation (5) is accordingly modified as follows:

$$y_t = \left\{ \sum_{r=1}^{m} G_t^{r-1} B_r' \right\} z_t + \varepsilon_t = \begin{bmatrix} I_{\tilde{n}} & G_t^1 & \dots & G_t^{m-1} \end{bmatrix} \begin{bmatrix} B_1 \\ B_2 \\ \vdots \\ B_m \end{bmatrix} z_t + \varepsilon_t = \tilde{G}_t \tilde{B}' \, z_t + \varepsilon_t, \quad (7)$$

where $\tilde{G}_t$ has dimension $\tilde{n} \times m\tilde{n}$, $\tilde{B}'$ is a $m\tilde{n} \times (1 + k + p\tilde{n})$ matrix and $G_t^0 = I_{\tilde{n}}$ (identity matrix) indicates that there is no transition before the first break.

Finally, we specify each diagonal element $G_{i,t}^r$ as a logistic cumulative density functions, i.e.

$$G_{i,t}^r(s_{i,t}^r; \gamma_i^r, c_i^r) = \left[1 + \exp\left\{ -\gamma_i^r (s_{i,t}^r - c_i^r) \right\} \right]^{-1}, \tag{8}$$

for $i = 1, 2, \dots, \tilde{n}$ and $r = 0, 1, \dots, m-1$, so that (7) is a Vector Logistic Smooth Transition AutoRegressive (VLSTAR) model.

2.1 The nonlinear model specification

The nonlinear model specification can be thought of as made up of three steps, such as: the choice of the relevant exogenous explanatory variables (x_t), the determination of the number of regimes (m), the selection of the transition variable(s) (s_t).

Firstly, the set of determinants of the dependent variables shall be defined. Possible candidates can be derived from the literature on the macroeconomic and financial determinants of volatility. In a seminal paper, Schwert (1989) analysed the volatility dynamics related to the business cycle. His study involved the use of several macroeconomic variables as determinants. Despite the lack of significance, Schwert (1989) found counter-cyclical movements of volatility compared to the level of the economy. These findings stimulated an interest in analysing the effects of economic and financial activities on volatility. Mele (2007, 2008), among others, suggested the use of stock returns predictors as volatility determinants in a study on risk premium. More recently, Christiansen, Schmeling, and Schrimpf (2012), Paye (2012), Conrad and Loch (2014) and Bucci (2019a,b) analysed the role of macroeconomic and financial variables in predicting the realized volatility.

Once a set of exogenous explicative variables is chosen, the next step is to test the presence of structural breaks in the time series of the Cholesky factors, in order to determine the number of regimes in the model. For this purpose, we employ the Bai and Perron (1998, 2003) procedure. Let $\sup F_t(l)$ be the test statistic for the null hypothesis of no structural breaks versus an alternative hypothesis containing an arbitrary number of breaks, with a maximum of M. We employ an equally weighted version of such test defined as

$$\mathrm{UD}_{max} = \max_{1 \leq m \leq M} \quad F_t(\hat{\lambda}_1, \ldots, \hat{\lambda}_m),$$

and a not-equally weighted version

$$\mathrm{WD}_{max} = \max_{1 \leq m \leq M} \quad w_m F_t(\hat{\lambda}_1, \ldots, \hat{\lambda}_m),$$

where $\hat{\lambda}_r = T_r/T$, T_r is the sample size in regime r, with $r = 1, \ldots, m$ and w_m is a weight which depends on m. For both tests, the null hypothesis is the absence of structural breaks against the alternative hypothesis of an unknown number of breaks. Once the presence of breaks has been detected via the UD_{max} and WD_{max} tests, the optimal number of breaks is determined using the $F_t(l+1 \mid l)$ test in which the null hypothesis is l breaks in the time series against the alternative of $l+1$ breaks (see Bai and Perron, 2003, for details).

The choice of the transition variables s_t may be based on economic theory or may be data driven. Whether the economic theory does not allow us to define a unique transition variable, a common choice is to test the linearity of the model for each candidate transition variable and to select the one that exhibits the lowest p-value. In the multivariate framework, linearity can be tested jointly assuming a common transition variable; alternatively, a different threshold variable can be chosen for each equation, see also Camacho (2004), Luukkonen, Saikkonen, and Teräsvirta (1988) and Teräsvirta and Yang (2014).

Testing linearity is essential before fitting a nonlinear model to time series data. Specifically, it must be checked if a linear model is nested in the nonlinear framework or may be an adequate representation of the data-generating process, thus simplifying the estimation process. Moreover, nonlinear models, like the smooth transition regression and the switching regression model, are not identified if the nested linear model is the data-generating process.

As discussed in Luukkonen, Saikkonen, and Teräsvirta (1988), Teräsvirta (1994) and Teräsvirta, Tjøstheim, and Granger (2010), linearity in smooth transition models may be tested through a Lagrange Multiplier (LM) test, assuming that the nonlinear model of the alternative hypothesis is a 2-state dynamic smooth transition regression. The transition function $G^r(s_t^r;\gamma^r, c_{ij}^r)$ is approximated by a third-order Taylor expansion around the null hypothesis of linearity $H_0\colon \gamma = 0$. After merging and reparameterizing, this yields the linear regression model

$$y_t = X_t\beta_0 + X_t s_t\beta_1 + X_t s_t^2\beta_2 + X_t s_t^3\beta_3 + \varepsilon_t, \tag{9}$$

where s_t is the transition variable and β_i, for $i = 0, 1, 2, 3$, are coefficients vectors functions of the original parameters. Testing for linearity is equivalent to testing the null hypothesis $H_{0:} : \beta_i = 0$ for each $i > 1$ in the previous regression. In a multivariate framework, if more than one equation share the same transition variable, it is possible to apply a joint linearity test.

2.2 Estimation and forecasting

Assuming $\varepsilon_t \sim i.i.d.N(0, \Omega)$, the model (7) can be represented by the following multivariate conditional density function

$$f(y_t|I_t;\theta) = (2\pi)^{-\frac{\tilde{n}}{2}}|\Omega|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(y_t - \tilde{G}_t B'\, z_t)'\Omega^{-1}(y_t - \tilde{G}_t B'\, z_t)\right\}, \tag{10}$$

where I_t is the information set at time t which contains all the exogenous variables x_t and all the lags of y_t. As a consequence, the overall conditional loglikelihood function is

$$\ell(y_t|I_t;\theta) = -\frac{T\tilde{n}}{2}\ln(2\pi) - \frac{T}{2}\ln|\Omega| - \frac{1}{2}\sum_{t=1}^{T}(y_t - \tilde{G}_t B\, z_t)'\Omega^{-1}(y_t - \tilde{G}_t B\, z_t), \tag{11}$$

where T is the sample size and the vector θ contains all the unknown parameters in matrices $\tilde{B}$, Ω and $\tilde{G}_t$ which, in general, depends on scale parameters and thresholds

$$\Gamma = \begin{bmatrix} \gamma_1^1 & \gamma_1^2 & \dots & \gamma_1^m \\ \gamma_2^1 & \gamma_2^2 & \dots & \gamma_2^m \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_{\tilde{n}}^1 & \gamma_{\tilde{n}}^2 & \dots & \gamma_{\tilde{n}}^m \end{bmatrix} \quad \text{and} \quad C = \begin{bmatrix} c_1^1 & c_1^2 & \dots & c_1^m \\ c_2^1 & c_2^2 & \dots & c_2^m \\ \vdots & \vdots & \ddots & \vdots \\ c_{\tilde{n}}^1 & c_{\tilde{n}}^2 & \dots & c_{\tilde{n}}^m \end{bmatrix}.$$

The total amount of unknown parameters in equation (11) is $\tilde{n}\,[1 + 0.5(\tilde{n}+1) + p\tilde{n} + k + 2m]$. Model (7) can be estimated through nonlinear least squares (NLS) or maximum likelihood (ML). In this paper, we opted for the NLS.

The NLS estimator is defined as the solution of the following optimisation problem

$$\hat{\theta}_{NLS} = \arg\min_{\theta} \sum_{t=1}^{T} (y_t - \tilde{G}_t \tilde{B}' z_t)'(y_t - \tilde{G}_t \tilde{B}' z_t). \tag{12}$$

The algorithm of optimization may converge to a local minimum instead of the global, therefore the choice of the starting values for θ is crucial. For this reason, (see Teräsvirta and Yang, 2014, for details), it is necessary to implement the following algorithm for obtaining the starting values for Γ and C:

1. construct a discrete grid of Γ and C values;
2. estimate B_{NLS} conditionally to the values of the grid, calculating the corresponding residuals sum of squares, Q_T;
3. find the smallest Q_T, and choose the related pair of Γ and C as starting values, Γ_0 and C_0.

The values of the grid for Γ ranged from 0 to 100, while the values of C ranged from minimum to maximum of each dependent variable. For example, using a sequence of 50 values of Γ and C, we had 2500 couples of values for each dependent variable.

The NLS estimates of vec(B) for step 3, given the values of Γ and C, are equal to:

$$\text{vec}(\hat{B})_{NLS} = \left[T^{-1} \sum_{t=1}^{T} (\tilde{G}_t \tilde{G}_t') \otimes (z_t z_t')\right]^{-1} \left[T^{-1} \sum_{t=1}^{T} \text{vec}(z_t y_t' \tilde{G}_t')\right]. \tag{13}$$

The estimated errors covariance matrix is given by

$$\hat{\Omega}_{NLS} = T^{-1}\hat{E}'\hat{E}, \tag{14}$$

where $\hat{E} = (\hat{\varepsilon}_1, \ldots, \varepsilon_t)'$ is a $T \times \tilde{n}$ matrix, and $\varepsilon_t = y_t - \tilde{G}_t\hat{B}'_{NLS}z_t$ is a column vector of residuals.

Once obtained the initial values of Γ and C from the previous algorithm, a new algorithm is implemented to obtain an estimate of θ_{NLS} and Ω_{NLS}, without increasing the computational complexity of the estimation process:

1. estimate B through Equation (13), relying on Γ_0 and C_0;
2. use $\hat{B}$, estimated in step 1, to obtain an estimate of Γ and C by Equation (12);
3. estimate the new B through Equation (13).
4. repeat step 2 and 3 until convergence.

The forecasts of the nonlinear model, for more than one step ahead, can be generalised via numerical techniques. Given a nonlinear model

$$y_t = g(z_t, \theta) + \varepsilon_t, \tag{15}$$

where θ is a vector of parameters to be estimated, z_t is a combination of lagged values of y_t and of exogenous variables x_t and ε_t is a white noise with zero mean and constant variance σ^2. Let I_T be the information set at time T and ε_t be independent of I_{T-1}, the forecast of y_{T+h} made at time T is equal to the conditional mean

$$\hat{y}_{T+h} = E\left\{y_{T+h}|I_t\right\} = E\left\{g(z_{T+h-1})|I_t\right\}. \tag{16}$$

When $h = 1$ the forecast $\hat{y}_{T+1} = g(z_t)$ is obtained from equation (16); otherwise, if $h \geq 2$, the prediction can only be calculated recursively using numerical techniques. See Hubrich and Teräsvirta (2013), Kock and Teräsvirta (2011) and Teräsvirta, Tjøstheim, and Granger (2010) for a detailed analysis of forecasting methods.

In this paper, we forecast the lower triangular, $\hat{Y}_{t+h}$, by means of the Cholesky-VLSTAR model presented above, then we get the forecast covariance matrix as $\widehat{\mathrm{RC}}_{t+h} = \hat{Y}_{t+h}\hat{Y}'_{t+h}$ which is by construction a symmetric and positive definite matrix.

3 Data

Our analysis relies on a sample of $n = 4$ time series of stock market index returns, namely S&P 500 for USA, Nikkei for Japan, FTSE 100 for UK and DAX for Germany (this is the ordering of series in r_t). The sample period ranges from Wednesday 1st August 1990 to friday 29th June 2018 and includes 7283 daily observations for each series. Thomson DataStream is the data provider. Monthly realized volatility matrices are defined as in equation (1), for a total amount of $T = 335$ monthly observations. In this context, following the procedure described in section 2, the monthly realized covariance matrices are transformed using Cholesky decomposition according equation (2), thus obtaining $\tilde{n} = 10$ Cholesky factors y_t. The covariance stationarity of the series included in y_t is confirmed by Augmented Dickey-Fuller (ADF), ADF-GLS, Phillips-Perron and Kwiatkowsky- Phillips-Schmidt-Shin tests for unit roots.[1]

The set of macroeconomic and financial variables used as predetermined variables in the estimation process of the VLSTAR model are selected according to the results of the literature on risk premia predictability, e.g. Mele (2007). Specifically, they are all sampled at monthly frequency. The level of macroeconomic activity is measured through the same variables used by Schwert (1989), such as the inflation rate (π_{t-1}), the industrial production growth rate (g_{t-1}) of the United States, Japan and EU, the unemployment growth rate (u_{t-1}). The set contains two variables such as the dividend price ratio (DP_{t-1}) and the earning price ratio (EP_{t-1}) of the S&P 500 used to predict the excess returns, see also Welch and Goyal (2008). We also include the Fama and French's factors (see Fama and French, 1993) of each country as a measure of market risk (MKT, SMB, HML) and to capture the leverage effect. We also consider the Economic Policy Uncertainty Index (EPU_{t-1}) for the U.S., Japan and EU as computed by Baker, Bloom, and Davis (2015). In fact, it has been showed (see Liu and Zang, 2015) that stock market volatility tends to increase in presence of higher levels of economic policy uncertainty. Finally, we include the oil price growth rate even though we do not disentangle between oil price shocks originated from the demand and supply side, as in Bastianin and Manera (2018). Table 1 provides a description of the $k = 26$ available exogenous variables (x_t).

[1] The results are available upon request from the authors.

Table 1: Exogenous explanatory variables

Symbol	Variable description	Source
$\Delta\pi_{t-1}^{US}$	First difference of US monthly inflation rate	Datastream
$\Delta\pi_{t-1}^{UK}$	First difference of UK monthly inflation rate	Datastream
$\Delta\pi_{t-1}^{GE}$	First difference of German monthly inflation rate	Datastream
$\Delta\pi_{t-1}^{JP}$	First difference of Japanese monthly inflation rate	Datastream
g_{t-1}^{US}	US monthly Industrial Production growth	OECD Database
g_{t-1}^{EU}	EU monthly Industrial Production growth	OECD Database
g_{t-1}^{JP}	Japan monthly Industrial Production growth	OECD Database
DP_{t-1}	Dividend Yield Ratio S&P 500 growth rate over the past year relative to current market prices; S&P500 index	Robert Shiller's website
EP_{t-1}	Earning Price Ratio S&P 500 growth rate over the past year relative to current market prices; S&P500 index	Robert Shiller's website
MKT_{t-1}^{US}	Fama-French's market factor for U.S.	Kenneth French's website
SMB_{t-1}^{US}	Fama-French's SMB for U.S.	Kenneth French's website
HML_{t-1}^{US}	Fama-French's HML for U.S.	Kenneth French's website
MKT_{t-1}^{JP}	Fama-French's market factor for Japan	Kenneth French's website
SMB_{t-1}^{JP}	Fama-French's SMB for Japan	Kenneth French's website
HML_{t-1}^{JP}	Fama-French's HML for Japan	Kenneth French's website
MKT_{t-1}^{EU}	Fama-French's market factor for the European Union	Kenneth French's website
SMB_{t-1}^{EU}	Fama-French's SMB for the European Union	Kenneth French's website
HML_{t-1}^{EU}	Fama-French's HML for the European Union	Kenneth French's website
u_{t-1}^{US}	Unemployment rate growth in U.S.	Datastream
u_{t-1}^{UK}	Unemployment rate growth in the United Kingdom	Datastream
u_{t-1}^{JP}	Unemployment rate growth in Japan	Datastream
u_{t-1}^{GE}	Unemployment rate growth in Germany	Datastream
EPU_{t-1}^{US}	Economic Policy Uncertainty in U.S.	Economic Policy Uncertainty's website
EPU_{t-1}^{JP}	Economic Policy Uncertainty in Japan	Economic Policy Uncertainty's website
EPU_{t-1}^{EU}	Economic Policy Uncertainty in European Union	Economic Policy Uncertainty's website
Oil	Growth rate of the oil price	Datastream

4 Estimation Results

We first estimate the "Cholesky-VAR" model with exogenous explanatory variables (`CholVARX`). This can be considered a baseline model since it is a linear stationary VAR(1) where the number of lags has been set using the Akaike and the Bayesian information criteria. The equation is

$$\Phi(L)y_t = \mu_0 + Ax_{t-1} + \delta WTC_t + \varepsilon_t, \tag{17}$$

where y_t is the $\tilde{n} \times 1$ vector of Cholesky factors, $\Phi(L) = I_{\tilde{n}} - \Phi_1 L$, x_{t-1} is the set of exogenous predictors at time $t-1$ and WTC_t is a dummy variable for the World Trade Center attack of September 11, 2001.

Table 2: `CholVARX` model estimates

	$y_{1,t}$	$y_{2,t}$	$y_{3,t}$	$y_{4,t}$	$y_{5,t}$	$y_{6,t}$	$y_{7,t}$	$y_{8,t}$	$y_{9,t}$	$y_{10,t}$
$\Delta\pi^{US}_{t-1}$	-0.1414 (0.2474)	0.0509 (0.259)	-0.1863 (0.2124)	0.0507 (0.3068)	-0.3073 (0.3695)	-0.0613 (0.1734)	0.2687 (0.2378)	-0.0088 (0.1841)	0.0768 (0.2417)	0.2704 (0.196)
$\Delta\pi^{UK}_{t-1}$	-0.1159 (0.1728)	-0.0595 (0.1809)	-0.0765 (0.1484)	-0.0969 (0.2143)	-0.5304^{**} (0.2581)	0.0978 (0.1211)	0.1976 (0.1662)	-0.2363^{*} (0.1286)	-0.2642 (0.1689)	-0.3119^{**} (0.1369)
$\Delta\pi^{GE}_{t-1}$	-0.1583 (0.1365)	-0.1124 (0.1429)	0.0021 (0.1172)	0.0401 (0.1693)	-0.0468 (0.2039)	0.0247 (0.0957)	-0.1005 (0.1313)	-0.1426 (0.1016)	-0.1535 (0.1334)	0.0109 (0.1082)
$\Delta\pi^{JP}_{t-1}$	0.1207 (0.2439)	-0.1072 (0.2554)	0.1498 (0.2095)	-0.2709 (0.3025)	0.3735 (0.3643)	-0.1628 (0.171)	-0.0798 (0.2345)	0.0597 (0.1815)	0.1948 (0.2384)	-0.0087 (0.1933)
g^{US}_{t-1}	-0.2148 (0.1408)	0.3768^{**} (0.1474)	-0.2593^{**} (0.1209)	-0.0828 (0.1746)	-0.1209 (0.2103)	-0.094 (0.0987)	-0.0344 (0.1354)	-0.0171 (0.1048)	0.1896 (0.1376)	0.0463 (0.1116)
g^{EU}_{t-1}	-0.1749 (0.1085)	-0.0878 (0.1136)	-0.0372 (0.0932)	-0.1038 (0.1345)	-0.047 (0.162)	0.0326 (0.076)	-0.009 (0.1043)	-0.1435^{*} (0.0807)	-0.1497 (0.106)	-0.0079 (0.086)
g^{JP}_{t-1}	0.0755 (0.0462)	0.0418 (0.0484)	0.0082 (0.0397)	0.0388 (0.0573)	-0.1242^{*} (0.069)	0.0473 (0.0324)	-0.0667 (0.0444)	0.0867^{**} (0.0344)	$9e-04$ (0.0452)	0.0568 (0.0366)
DP_{t-1}	32.6939^{***} (2.837)	0.1244 (2.9703)	14.3548^{***} (2.4361)	21.0419^{***} (3.5182)	26.8544^{***} (4.2374)	17.7602^{***} (1.9883)	12.4174^{***} (2.7278)	21.0629^{***} (2.1113)	18.0052^{***} (2.7723)	12.5173^{***} (2.2482)
MKT^{US}_{t-1}	0.0755^{***} (0.0284)	-0.0458 (0.0297)	0.0877^{***} (0.0244)	0.118^{***} (0.0352)	0.1208^{***} (0.0424)	0.0512^{**} (0.0199)	-0.001 (0.0273)	0.0416^{**} (0.0211)	0.0553^{**} (0.0277)	0.067^{***} (0.0225)
SMB^{US}_{t-1}	0.0082 (0.0278)	-0.0207 (0.0291)	-0.0035 (0.0239)	-0.0167 (0.0345)	-0.1124^{***} (0.0415)	0.0061 (0.0195)	0.0363 (0.0267)	-0.0389^{*} (0.0207)	-0.0906^{***} (0.0272)	-0.0043 (0.022)
HML^{US}_{t-1}	0.014 (0.0363)	-0.0405 (0.038)	0.0347 (0.0312)	0.1127^{**} (0.045)	0.0786 (0.0542)	0.0028 (0.0254)	$3e-04$ (0.0349)	-0.0395 (0.027)	-0.0398 (0.0355)	0.0084 (0.0288)
MKT^{JP}_{t-1}	0.0052 (0.0153)	-0.0145 (0.0161)	-0.0031 (0.0132)	-0.0088 (0.019)	-0.0484^{**} (0.0229)	-0.007 (0.0107)	-0.0138 (0.0147)	0.0064 (0.0114)	0.0094 (0.015)	-0.0257^{**} (0.0122)
SMB^{JP}_{t-1}	-0.048^{*} (0.0266)	0.0282 (0.0279)	0.0066 (0.0228)	0.0138 (0.033)	-0.0227 (0.0397)	-0.0099 (0.0186)	-0.0355 (0.0256)	-0.0301 (0.0198)	-0.046^{*} (0.026)	-0.0597^{***} (0.0211)
HML^{JP}_{t-1}	-0.0112 (0.0312)	0.0048 (0.0327)	-0.0192 (0.0268)	-0.0281 (0.0387)	-0.0848^{*} (0.0466)	0.0124 (0.0219)	-0.0445 (0.03)	0.0097 (0.0232)	-0.0242 (0.0305)	-0.0445^{*} (0.0247)
MKT^{EU}_{t-1}	0.0054 (0.045)	-0.0252 (0.0471)	0.0175 (0.0387)	0.0546 (0.0558)	0.0576 (0.0672)	-0.0268 (0.0315)	-0.0378 (0.0433)	0.0092 (0.0335)	0.0518 (0.044)	0.0718^{**} (0.0357)
SMB^{EU}_{t-1}	0.0353 (0.0441)	-0.0075 (0.0462)	-0.0307 (0.0379)	-0.0287 (0.0547)	0.0998 (0.0658)	-0.0111 (0.0309)	-0.0103 (0.0424)	$8e-04$ (0.0328)	-0.0059 (0.0431)	0.0097 (0.0349)
HML^{EU}_{t-1}	1.5254 (2.2721)	-4.2057^{*} (2.3789)	0.1403 (1.9511)	1.7325 (2.8177)	0.1828 (3.3937)	-0.215 (1.5924)	0.1334 (2.1847)	-2.0034 (1.6909)	-1.4034 (2.2204)	0.8765 (1.8006)
u^{US}_{t-1}	10.7092^{***} (3.2195)	-1.2646 (3.3708)	2.6318 (2.7646)	3.5421 (3.9925)	9.1597^{*} (4.8087)	0.6815 (2.2563)	0.4975 (3.0956)	5.436^{**} (2.396)	6.9104^{**} (3.1461)	2.1659 (2.5513)
u^{UK}_{t-1}	-0.5775 (5.0407)	5.3293 (5.2775)	7.5388^{*} (4.3284)	10.213 (6.2509)	1.2401 (7.5289)	0.4927 (3.5327)	0.958 (4.8467)	0.9731 (3.7513)	-0.6446 (4.9258)	-4.9604 (3.9945)
u^{JP}_{t-1}	6.3511^{**} (2.6276)	1.4371 (2.7511)	1.4132 (2.2563)	2.6218 (3.2585)	5.747 (3.9247)	1.6466 (1.8415)	1.0262 (2.5265)	3.0986 (1.9555)	-0.535 (2.5677)	1.4807 (2.0823)
u^{GE}_{t-1}	-0.8931 (7.0486)	3.2662 (7.3798)	-7.1361 (6.0527)	8.5025 (8.741)	11.6306 (10.528)	-1.4417 (4.9399)	-1.2796 (6.7774)	2.2589 (5.2457)	-6.3236 (6.888)	17.6095^{***} (5.5858)
EPU^{US}_{t-1}	$5e-04$ (0.0027)	0.0018 (0.0029)	0.0026 (0.0023)	0.0092^{***} (0.0034)	0.0085^{**} (0.0041)	0.0032^{*} (0.0019)	0.0025 (0.0026)	0.0016 (0.002)	-0.0024 (0.0027)	0.0051^{**} (0.0022)
EPU^{JP}_{t-1}	$-3e-04$ (0.0037)	0.0079^{**} (0.0039)	0.0089^{***} (0.0032)	0.0101^{**} (0.0046)	0.0055 (0.0055)	-0.0033 (0.0026)	$-7e-04$ (0.0035)	-0.0049^{*} (0.0027)	-0.0053 (0.0036)	$-2e-04$ (0.0029)
EPU^{EU}_{t-1}	$-2e-04$ (0.0025)	-0.0042 (0.0026)	-0.0028 (0.0021)	-0.0069^{**} (0.0031)	-0.0061^{*} (0.0037)	-0.0016 (0.0017)	$-6e-04$ (0.0024)	0.0025 (0.0018)	0.0062^{**} (0.0024)	$1e-04$ (0.002)
Oil	-0.0143 (0.0186)	0.0052 (0.0195)	0.0129 (0.016)	0.0172 (0.0231)	0.0072 (0.0278)	-0.0239^{*} (0.013)	-0.0114 (0.0179)	0.0123 (0.0138)	0.0039 (0.0182)	$1e-04$ (0.0147)
WTC_t	-0.5663 (1.4641)	2.8774^{*} (1.5328)	2.1108^{*} (1.2572)	1.5451 (1.8156)	1.3807 (2.1867)	-6.3925^{***} (1.0261)	-7.5734^{***} (1.4077)	2.8404^{***} (1.0896)	7.8248^{***} (1.4307)	1.1807 (1.1602)
SSR	544.209	596.551	401.280	836.900	1214.080	267.295	503.132	301.408	519.682	341.758
S.E.	1.356	1.420	1.164	1.681	2.025	0.950	1.304	1.009	1.325	1.075
R^2	0.762	0.160	0.549	0.626	0.536	0.487	0.305	0.684	0.603	0.527
LL	-555.454	-570.789	-504.573	-627.326	-689.456	-436.720	-542.347	-456.779	-547.752	-477.760
AIC	1188.907	1219.579	1087.146	1332.651	1456.912	951.4395	1162.694	991.5574	1173.504	1033.521
BIC	1337.542	1368.213	1235.781	1481.286	1605.546	1100.074	1311.329	1140.192	1322.139	1182.155
DW	2.222	1.942	2.218	2.272	2.113	1.896	1.946	2.137	2.141	2.347
LB(5)	17.378^{**}	2.891	11.877^{**}	13.957^{**}	4.359	5.756	7.302	5.054	6.799	33.453^{***}
ARCH(5)	16.118^{***}	2.471	7.510	8.464	4.157	39.438^{***}	2.041	4.534	22.059^{***}	14.210^{**}

*** indicates $p-$ value $< 1\%$, ** indicates $p-$ value $< 5\%$ and * indicates $p-$ value $< 10\%$; SSR is the sum of squared residuals, SE is the standard error of the regresssion, R^2 is the R-squared index, LL is the loglikelihood, AIC and BIC are the Akaike and Bayesian information criteria; DW is the Durbin-Watson statistic, LB(5) is the Ljung-Box test ststistic with 5 lags and ARCH(5) is the test for conditional heteroskedasticity with 5 lags.

Table 2 reports the results of the Cholesky-VAR estimation on the full sample (only the estimated coefficients of the exogenous variables are reported to save space). All the eigenvalues of the companion form matrix are smaller than one in modulus (the maximum eigenvalue is $\lambda_{\max} \approx 0.76$), hence the estimated VAR(1) is likely to be stationary. Moreover, the diagnostic tests highlight that the model residuals are not autocorrelated, but they show some conditional heteroskedasticity. As a consequence, despite the `CholVARX` could be misspecified in variance, surely it represents a good specification for the conditional mean of y_t and may be used as an initial reference model for the specification of an alternative nonlinear one.

The Cholesky decomposition guarantees that the conditional covariance matrix is at least semidefinite positive for each t. Since the aim of this paper is to provide a suitable statistical model for the conditional covariance matrix RC_t, through the modelling of the dynamic of the conditional mean of the Cholesky factors, the estimated coefficients of equation (17) do not have a straightforward economic interpretation. Yet, some exogenous variables in x_{t-1} seem to be good predictors for the Cholesky factors. The variables affecting the dependent variables in the linear framework are used as potential predictors in the VLSTAR model, in order to keep the model as parsimonious as possible. Plainly, the choice of the set of determinants could be different in the nonlinear framework, nevertheless no clear methods are available for such tasks.

In order to estimate the VLSTAR model for the Cholesky factors, the latter should be tested for the presence of structural breaks and nonlinearity. As we mentioned in section 2.1, we employ the UD_{max}, the WD_{max} and the $\sup F_t(l+1 \mid l)$ tests to determine the number of regimes m. As shown in table 3, the most part of the tests are significant, for $l = 1, 2, \ldots, 8$, for each time series, confirming that at least one break is present in the factor series. Moreover, in most of the cases the sequential $\sup F_t(l + 1 \mid l)$ test highlights that the number of regimes m is smaller than 4. Further, since for the most of the Cholesky factors the sequential $\sup F_T(l + 1 \mid l)$ test rejects the null hypothesis when $l = 0$, we conclude that a single break is an acceptable approximation, hence $m = 2$ regimes. We further have performed the test of common breaks among dependent variables provided by Bai, Lumsdaine, and Stock (1998) and Aue, Hörmann, Horváth, and Reimherr (2009) which exhibits the presence of at least one common break.

Table 3: Multiple structural changes tests

	$y_{1,t}$	$y_{2,t}$	$y_{3,t}$	$y_{4,t}$	$y_{5,t}$	$y_{6,t}$	$y_{7,t}$	$y_{8,t}$	$y_{9,t}$	$y_{10,t}$
$\sup F_T(1)$	**19.55**	1.70	**17.29**	**50.05**	5.85	**5.96**	5.92	**11.01**	**38.86**	**44.93**
$\sup F_T(2)$	**10.00**	3.59	**23.87**	**32.93**	2.95	**3.45**	4.26	**14.61**	**21.77**	**35.75**
$\sup F_T(3)$	**15.51**	4.10	**42.71**	**33.51**	5.76	**4.18**	4.67	**18.22**	**10.89**	**43.99**
$\sup F_T(4)$	**36.59**	**6.88**	**42.92**	**28.98**	**6.34**	**4.37**	4.01	**19.62**	**9.09**	**32.93**
$\sup F_T(5)$	**34.16**	**7.59**	**32.39**	**28.13**	**4.52**	**3.45**	3.27	**18.58**	**7.09**	**22.67**
$\sup F_T(2 \mid 1)$	**19.93**	4.37	**26.45**	**11.33**	1.93	3.95	3.70	**13.73**	**10.19**	**14.42**
$\sup F_T(3 \mid 2)$	**19.77**	2.58	**23.20**	6.23	**12.05**	3.24	5.69	9.23	1.68	**14.42**
$\sup F_T(4 \mid 3)$	**55.28**	6.37	**16.02**	**15.04**	3.05	0.19	3.07	**12.53**	6.89	0.36
$\sup F_T(5 \mid 4)$	0.00	0.60	0.00	0.00	0.00	0.00	0.92	0.00	0.00	0.00

Trimming set to 0.15 and $M = 5$, the rejections at 5% significance level are in bold

The specification of the nonlinear model in (7), based on Akaike and Bayesian Information criteria, has one lag ($p = 1$). Therefore, in line with Camacho (2004), the 2-state `CholVLSTARX` model is

$$y_t = \Big(B_1 + G_t^1 B_2\Big) z_t + \varepsilon_t, \tag{18}$$

where $z_t = [1 \quad y'_{t-1} \quad \eta'_{t-1} \quad d'_t]'$ has dimension $3 + \tilde{n} + b$, where b is the number of exogenous variables affecting the dependent variable in the VAR(1) model and $\eta_{t-1} =[\ \mathrm{DP}_{t-1},\ \Delta\pi^{UK}_{t-1},\ g^{US}_{t-1},\ g^{JP}_{t-1},\ \mathrm{MKT}^{US}_{t-1},\ \mathrm{SMB}^{US}_{t-1},\ \mathrm{SMB}^{JP}_{t-1},\ \mathrm{HML}^{JP}_{t-1},\ \mathrm{u}^{US}_{t-1},\ \mathrm{u}^{GE}_{t-1},\ \mathrm{EPU}^{US}_{t-1},\ \mathrm{EPU}^{JP}_{t-1},\ \mathrm{EPU}^{EU}_{t-1}\ ]'$. Once the number of regimes is determined, the nonlinear specification procedure foresees the selection of the transition variable and testing for linearity of the model. Since the economic theory does not provide any specific insight for choosing the transition variable, we repeat the test for each predicting variable, equation by equation. We choose the transition variable according to the minimum p-value associated with the univariate linearity test. When two or more equations share the same transition variable as a candidate, we perform a joint linearity test, described in Appendix 6, assuming that nonlinear dynamics are driven by one single transition variable. Given a 2-regimes VLSTAR model, as in (18), the null of nonlinearity equals to $H_0\colon \gamma_j = 0,\ j = 1, 2, \ldots, \tilde{n}$, while the alternative is that at least one shape parameter, γ_j, is greater than 0. The test, as the univariate version in Luukkonen, Saikkonen, and Teräsvirta (1988), is based on third-order Taylor expansion of the transition variable and it is further analysed in appendix 6. The test statistics of the linearity tests for each equation and for each candidate transition variable are reported in Table 4. In some cases, the highest test statistics is associated with the first lag of g^{US}. For other equations, the linearity test is mainly significant for the first lag of $y_{1,t}$.

Table 4: Linearity Test

s_t	$y_{1,t}$	$y_{2,t}$	$y_{3,t}$	$y_{4,t}$	$y_{5,t}$	$y_{6,t}$	$y_{7,t}$	$y_{8,t}$	$y_{9,t}$	$y_{10,t}$
$y_{1,t-1}$	0.081	0.542	3.892	3.445	1.49	0.832	7.234	3.493	15.297	0.024
$y_{2,t-1}$	0.985	0.11	2.545	0.349	0.546	0.074	0.041	0.011	0.055	7.15
$y_{3,t-1}$	0.097	1.14	6.232	2.694	0.339	0.482	5.383	0.719	9.741	0.369
$y_{4,t-1}$	0.054	1.295	4.002	16.041	1.411	0.504	2.053	0.961	7.002	4.609
$y_{5,t-1}$	0.174	0.436	0.081	0.345	1.294	0.284	1.307	0.003	3.428	1.058
$y_{6,t-1}$	0.081	0.282	0.682	0.361	0.342	3.267	6.191	1.368	13.486	1.058
$y_{7,t-1}$	0.408	1.36	1.269	1.858	0.061	4.399	4.214	0.654	6.014	7.598
$y_{8,t-1}$	0.176	0.956	4.072	7.121	1.116	1.252	1.512	0.686	10.612	0.19
$y_{9,t-1}$	0.011	3.972	3.464	5.624	0.482	0.636	0.96	0.194	9.259	3.051
$y_{10,t-1}$	0.177	0.1	0.461	0.957	1.715	0.162	0.005	0.094	1.954	24.003
DP_{t-1}	0.078	3.918	3.224	0.877	1.073	0.14	4.241	2.411	15.739	4.827
$\Delta\pi^{UK}_{t-1}$	5.231	5.354	4.391	4.397	1.353	2.608	0.005	1.708	4.921	4.757
g^{US}_{t-1}	2.763	2.767	11.03	11.005	24.141	2.137	0.293	11.113	1.972	2.893
g^{JP}_{t-1}	0.755	1.579	0.066	0.08	1.671	0.77	0.463	2.072	1.537	0.294
MKT^{US}_{t-1}	1.861	3.193	1.389	0.339	1.664	0.002	1.772	0.002	11.095	0.351
SMB^{US}_{t-1}	3.276	1.473	0.947	0.332	0.611	5.123	0.018	1.772	0.641	0.404
SMB^{JP}_{t-1}	0.073	1.276	0.281	0.937	0.313	4.076	0.773	0.005	3.698	3.157
HML^{JP}_{t-1}	0.344	0.48	2.005	0.906	0.423	0.793	1.014	0.833	0.825	0.900
u^{US}_{t-1}	11.609	0.597	5.478	2.487	16.543	27.556	3.79	5.456	2.436	0.51
u^{GER}_{t-1}	6.794	0.019	4.933	0.089	0.721	25.083	5.407	0.949	0.146	3.624
EPU^{US}_{t-1}	0.107	1.664	0.899	0.432	0.444	0.177	1.029	0.601	8.996	2.241
EPU^{JP}_{t-1}	0.982	0.41	4.414	3.662	0.036	4.812	2.941	0.024	6.899	2.857
EPU^{EU}_{t-1}	0.066	6.444	0.456	0.053	0.289	0.052	1.511	2.242	0.776	0.316

* and ** indicate highest test statistics

Since it turns out that several equations share more than one candidate transition variables, there is no clear cut conclusion that can be drawn upon the linearity tests. To circumvent this issue we run a joint linearity test on the whole model, assuming a unique transition variable for all the equations. Table 5 shows the results of the LM test, introduced by Yang and Teräsvirta (2013). Even from the results of the joint linearity test do not emerge a preferred transition variable. Now, since we put the monthly return on S&P 500 as first series in the vector r_t, it follows that in the Cholesky decomposition the first factor corresponds to the realized volatility (square root of the realized variance) of the US stock market. This choice is motivated by the the relevance of the US stock market in the global scenario. Thus selecting as transition variable $y_{1,t-1}$ for all the equations corresponds to assuming that the switch between the two regimes is governed by the evolution of the volatility on the US market.

Table 5: LM Test

Transition variable	LM	p-value	Transition variable	LM	p-value
$y_{1,t-1}$	813.850	0.008	$\Delta\pi^{UK}_{t-1}$	794.543	0.028
$y_{2,t-1}$	739.462	0.299	g^{US}_{t-1}	784.120	0.049
$y_{3,t-1}$	755.646	0.173	g^{JP}_{t-1}	783.133	0.051
$y_{4,t-1}$	819.626	0.006	DP_{t-1}	1021.201	<0.001
$y_{5,t-1}$	746.603	0.239	MKT^{US}_{t-1}	899.868	<0.001
$y_{6,t-1}$	833.207	0.002	SMB^{US}_{t-1}	744.741	0.254
$y_{7,t-1}$	928.817	<0.001	SMB^{JPN}_{t-1}	744.498	0.256
$y_{8,t-1}$	742.627	0.272	HML^{JPN}_{t-1}	740.156	0.293
$y_{9,t-1}$	785.386	0.045	u^{US}_{t-1}	735.820	0.333
$y_{10,t-1}$	782.313	0.053	u^{GER}_{t-1}	790.199	0.035
EPU^{US}_{t-1}	772.733	0.085	EPU^{JPN}_{t-1}	762.252	0.134
EPU^{EU}_{t-1}	682.255	0.840			

Table 6 reports the estimated parameters of the nonlinear `CholVLSTARX` model, where the transition variable, $y_{1,t-1}$, is common to all the equations. The information criteria and the value of the log-likelihood computed for the nonlinear model are similar to those found in the case of the linear specification. Based on Akaike and Bayesian information criteria, the model includes a single lag of the dependent variable y_t and the set of statistically significant leading variables in the linear `CholVARX` model. The estimated effects significantly vary among the equations and regressors. Table 7 reports the 5% significant-level leading variables in each equation in both regimes. In general, we can say that the leading variables are more significant in the second than in the first regime, like the macroeconomic variables, namely $\Delta\pi^{UK}_{t-1}, g^{US}_{t-1}, g^{JP}_{t-1}$. Further, across the equations the variables which are significant in the first regime tend to be less significant in the second regime and vice versa, like for instance MKT. In first regime, the EPU's variables are significant in the first 5 equations whereas in the second regime they are mainly significant in the last five equations. A notable exception is DP that is significant in both regimes. The evidence for the Fama and French factors is rather mixed.

Table 6: `CholVLSTARX` model estimates

$r = 1$										
μ_0, A_0, Φ_0	$y_{1,t}$	$y_{2,t}$	$y_{3,t}$	$y_{4,t}$	$y_{5,t}$	$y_{6,t}$	$y_{7,t}$	$y_{8,t}$	$y_{9,t}$	$y_{10,t}$
const	1.8658^{***} (0.2029)	0.8422^{***} (0.2315)	0.8099^{***} (0.1878)	1.1014^{***} (0.2643)	2.6177^{***} (0.3203)	-0.0374 (0.1496)	0.3247 (0.1986)	1.3884^{***} (0.1572)	0.511^{**} (0.2054)	1.1206^{***} (0.1657)
$y_{1,t-1}$	0.5685^{***} (0.0544)	-0.101 (0.0621)	0.0597 (0.0504)	0.1948^{***} (0.0709)	0.0613 (0.0859)	0.2282^{***} (0.0401)	0.1302^{**} (0.0533)	0.0858^{**} (0.0422)	0.0826 (0.0551)	0.0914^{**} (0.0444)
$y_{2,t-1}$	-0.1037^{**} (0.0446)	0.0582 (0.0509)	-0.0628 (0.0413)	-0.1324^{**} (0.0581)	0.1772^{**} (0.0705)	-0.0645^{*} (0.0329)	-0.049 (0.0437)	-0.125^{***} (0.0346)	-0.1443^{***} (0.0452)	-0.0761^{**} (0.0364)
$y_{3,t-1}$	-0.1702^{**} (0.0816)	0.0314 (0.0931)	0.2429^{***} (0.0755)	-0.0172 (0.1063)	-0.1181 (0.1288)	0.1571^{***} (0.0602)	-0.129 (0.0799)	0.0464 (0.0632)	-0.0218 (0.0826)	-0.1158^{*} (0.0666)
$y_{4,t-1}$	0.0328 (0.0483)	0.0816 (0.0551)	0.0368 (0.0447)	0.3513^{***} (0.0629)	-0.1187 (0.0762)	-0.1043^{***} (0.0356)	0.0972^{**} (0.0473)	0.0262 (0.0374)	0.183^{***} (0.0489)	0.0321 (0.0394)
$y_{5,t-1}$	-0.0269 (0.0323)	$9e-04$ (0.0369)	-0.0276 (0.0299)	-0.0504 (0.0421)	0.4708^{***} (0.051)	-0.0209 (0.0238)	-0.0629^{**} (0.0316)	$-3e-04$ (0.025)	-0.0203 (0.0327)	-0.017 (0.0264)

Continued on next page

Table 6 — continued from previous page

$y_{6,t-1}$	-0.1638* (0.0945)	0.2684 ** (0.1079)	-0.0685 (0.0875)	0.0571 (0.1231)	-0.473 *** (0.1492)	0.0297 (0.0697)	-0.2801*** (0.0925)	0.0339 (0.0732)	0.1584 * (0.0957)	-0.2042*** (0.0772)
$y_{7,t-1}$	-0.03 (0.0682)	-0.1581** (0.0778)	0.0395 (0.0631)	-0.0948 (0.0889)	0.1021 (0.1077)	0.0436 (0.0503)	0.1857 *** (0.0668)	-0.0806 (0.0528)	-0.089 (0.0691)	0.0185 (0.0557)
$y_{8,t-1}$	0.2195 *** (0.0831)	0.1468 (0.0948)	0.2269 *** (0.0769)	0.2456 ** (0.1083)	0.3684 *** (0.1312)	0.1557 ** (0.0613)	0.2709 *** (0.0814)	0.3703 *** (0.0644)	0.1995 ** (0.0842)	0.1077 (0.0679)
$y_{9,t-1}$	-0.0247 (0.0519)	-0.1329** (0.0592)	0.0557 (0.048)	0.0046 (0.0676)	0.0896 (0.0819)	-0.0773** (0.0382)	-0.1373*** (0.0508)	-0.0109 (0.0402)	0.0949 * (0.0525)	-0.048 (0.0423)
$y_{10,t-1}$	-0.1677*** (0.0569)	-0.0791 (0.065)	-0.1594*** (0.0527)	-0.283 *** (0.0742)	-0.1933** (0.0899)	-0.0282 (0.042)	0.0952 * (0.0557)	0.0514 (0.0441)	0.2015 *** (0.0576)	0.5766 *** (0.0465)
$\Delta\pi^{UK}_{t-1}$	0.1805 (0.1123)	-0.1271 (0.1281)	-0.0202 (0.1039)	0.0189 (0.1462)	0.0732 (0.1772)	0.0823 (0.0828)	0.1798 (0.1099)	0.0159 (0.0869)	-0.041 (0.1137)	-0.1348 (0.0917)
g^{US}_{t-1}	-0.0439 (0.1085)	0.2612 ** (0.1239)	-0.0348 (0.1005)	-0.2372* (0.1414)	-0.0213 (0.1714)	-0.2347*** (0.08)	0.0658 (0.1063)	0.0315 (0.0841)	0.1212 (0.1099)	0.0134 (0.0886)
g^{JP}_{t-1}	-0.013 (0.0354)	-0.0307 (0.0403)	-0.0468 (0.0327)	-0.054 (0.0461)	-0.3564*** (0.0558)	0.0453 * (0.0261)	-0.0725** (0.0346)	0.0572 ** (0.0274)	0.0197 (0.0358)	0.0037 (0.0289)
DP_{t-1}	29.7139*** (2.2178)	-8.4746*** (2.5308)	14.7237*** (2.0525)	15.5528*** (2.8891)	16.5135*** (3.501)	18.8107*** (1.6354)	8.3942*** (2.171)	17.713 *** (1.7178)	21.42 *** (2.2455)	5.5552 *** (1.8109)
MKT^{US}_{t-1}	0.1385 *** (0.0206)	-0.0955*** (0.0235)	0.09 *** (0.0191)	0.1353 *** (0.0269)	-0.0019 (0.0326)	0.0634 *** (0.0152)	-0.0161 (0.0202)	0.021 (0.016)	0.0915 *** (0.0209)	0.0241 (0.0169)
SMB^{US}_{t-1}	-0.0807*** (0.0205)	-0.023 (0.0234)	-0.0106 (0.019)	-0.0205 (0.0267)	-0.098 *** (0.0324)	0.0064 (0.0151)	$1e-04$ (0.0201)	-0.025 (0.0159)	-0.0697*** (0.0208)	0.004 (0.0168)
SMB^{JP}_{t-1}	-0.0421** (0.0203)	0.0498 ** (0.0231)	0.0078 (0.0188)	0.0127 (0.0264)	-0.0919*** (0.032)	-0.0238 (0.0149)	-0.0414** (0.0198)	-0.0449*** (0.0157)	-0.0261 (0.0205)	-0.0593*** (0.0165)
HML^{JP}_{t-1}	-0.0384* (0.0226)	0.0264 (0.0258)	-0.0032 (0.0209)	0.0145 (0.0294)	-0.1361*** (0.0356)	0.0048 (0.0167)	-0.0597*** (0.0221)	-0.0609*** (0.0175)	-0.0445* (0.0229)	-0.0423** (0.0184)
u^{US}_{t-1}	2.8199 (2.597)	-5.9024** (2.9635)	-0.406 (2.4034)	-4.2547 (3.383)	-1.0448 (4.0996)	2.0902 (1.915)	0.8249 (2.5422)	4.0813 ** (2.0115)	4.514 * (2.6294)	-2.7284 (2.1205)
u^{GE}_{t-1}	4.1211 (5.5943)	6.7323 (6.3837)	-7.4938 (5.1773)	5.45 (7.2875)	12.4442 (8.831)	2.6719 (4.1252)	0.3309 (5.4762)	5.5726 (4.333)	-6.6639 (5.6641)	15.2056*** (4.5679)
EPU^{US}_{t-1}	0.0065 *** (0.0022)	-0.0032 (0.0025)	0.003 (0.0021)	0.0061 ** (0.0029)	$-5e-04$ (0.0035)	0.005 *** (0.0016)	0.004 * (0.0022)	0.0018 (0.0017)	-0.0027 (0.0023)	0.0013 (0.0018)
EPU^{JP}_{t-1}	-0.0065** (0.003)	0.017 *** (0.0034)	0.008 *** (0.0027)	0.0118 *** (0.0039)	0.0125 *** (0.0047)	-0.0012 (0.0022)	0.0032 (0.0029)	-0.0038* (0.0023)	-0.0037 (0.003)	0.0062 ** (0.0024)
EPU^{EU}_{t-1}	-0.0029 (0.002)	-0.0046** (0.0023)	-0.0038** (0.0018)	-0.0082*** (0.0026)	-0.0074** (0.0031)	-0.0024* (0.0015)	-0.0037* (0.0019)	$3e-04$ (0.0015)	0.0048** (0.002)	-0.0026 (0.0016)
WTC_t	-0.7149 (1.1936)	0.3900 (1.3621)	1.0112 (1.1047)	-0.2301 (1.5549)	0.034 (1.8842)	-6.648 *** (0.8802)	-6.7121*** (1.1684)	4.1781 *** (0.9245)	8.2415 *** (1.2085)	2.5601 *** (0.9746)

$r = 2$

μ_1, A_1, Φ_1	$y_{1,t}$	$y_{2,t}$	$y_{3,t}$	$y_{4,t}$	$y_{5,t}$	$y_{6,t}$	$y_{7,t}$	$y_{8,t}$	$y_{9,t}$	$y_{10,t}$
$y_{1,t-1}$	0.4557 *** (0.0544)	-0.0837 (0.0621)	0.0619 (0.0504)	-0.0755 (0.0709)	0.1979 ** (0.0859)	-0.0819** (0.0401)	-0.0492 (0.0533)	0.0107 (0.0422)	-0.0582 (0.0551)	-0.1029** (0.0444)
$y_{2,t-1}$	-0.1398*** (0.0446)	-0.1248** (0.0509)	-0.0545 (0.0413)	-0.2548*** (0.0581)	-0.023 (0.0705)	0.1629 *** (0.0329)	-0.0961** (0.0437)	-0.1152*** (0.0346)	-0.138 *** (0.0452)	-0.0746** (0.0364)
$y_{3,t-1}$	0.7333 *** (0.0816)	0.2694 *** (0.0931)	0.1857 ** (0.0755)	0.1522 (0.1063)	0.6179 *** (0.1288)	0.3298 *** (0.0602)	0.6434 *** (0.0799)	0.2733 *** (0.0632)	1.1329 *** (0.0826)	0.1067 (0.0666)
$y_{4,t-1}$	-0.2949*** (0.0483)	-0.0893 (0.0551)	0.1934 *** (0.0447)	0.5681 *** (0.0629)	-0.499 *** (0.0762)	-0.2965*** (0.0356)	-0.5503*** (0.0473)	-0.1019*** (0.0374)	-0.3842*** (0.0489)	-0.1123*** (0.0394)
$y_{5,t-1}$	-0.1834*** (0.0323)	0.129 *** (0.0369)	-0.1005*** (0.0299)	-0.1374*** (0.0421)	0.1014** (0.051)	0.0871 *** (0.0238)	0.1298 *** (0.0316)	-0.0586** (0.025)	-0.6034*** (0.0327)	-0.1382*** (0.0264)
$y_{6,t-1}$	0.0961 (0.0945)	0.3176 *** (0.1079)	0.1334 (0.0875)	-0.0131 (0.1231)	0.4857 *** (0.1492)	-0.122 * (0.0697)	0.28 *** (0.0925)	0.0835 (0.0732)	0.1845 * (0.0957)	0.0557 (0.0772)
$y_{7,t-1}$	0.1415 ** (0.0682)	-0.0985 (0.0778)	0.0726 (0.0631)	0.1839 ** (0.0889)	0.0165 (0.1077)	0.0617 (0.0503)	-0.1416** (0.0668)	0.123 ** (0.0528)	0.028 (0.0691)	0.1774 *** (0.0557)
$y_{8,t-1}$	0.0515 (0.0831)	-0.4265*** (0.0948)	0.2107 *** (0.0769)	-0.0826 (0.1083)	-0.1665 (0.1312)	0.1603 *** (0.0613)	0.0327 (0.0814)	0.3224 *** (0.0644)	0.971 *** (0.0842)	0.0415 (0.0679)
$y_{9,t-1}$	0.056 (0.0519)	0.2413 *** (0.0592)	0.0817* (0.048)	0.4562 *** (0.0676)	0.1312 (0.0819)	0.1876 *** (0.0382)	0.1981 *** (0.0508)	0.0411 (0.0402)	-0.2693*** (0.0525)	0.2093 *** (0.0423)
$y_{10,t-1}$	0.3199 *** (0.0569)	0.1849*** (0.065)	-0.228 *** (0.0527)	0.1053 (0.0742)	0.3251 *** (0.0899)	0.0715* (0.042)	0.0532 (0.0557)	0.3385 *** (0.0441)	0.8003 *** (0.0576)	0.513 *** (0.0465)
$\Delta\pi^{UK}_{t-1}$	-0.4899*** (0.1123)	0.0088 (0.1281)	0.3338 *** (0.1039)	-0.1628 (0.1462)	-0.7289*** (0.1772)	-0.2307*** (0.0828)	-0.4519*** (0.1099)	-1.3017*** (0.0869)	-3.6859*** (0.1137)	-0.8985*** (0.0917)
g^{US}_{t-1}	-0.2765** (0.1085)	0.5392 *** (0.1239)	-0.5646*** (0.1005)	-0.0322 (0.1414)	-0.1518 (0.1714)	0.3817*** (0.08)	0.4488 *** (0.1063)	-0.0511 (0.0841)	0.7865 *** (0.1099)	0.1804 ** (0.0886)
g^{JP}_{t-1}	0.1694 *** (0.0354)	0.0987 ** (0.0403)	0.2818 *** (0.0327)	0.255 *** (0.0461)	0.1762 *** (0.0558)	-0.0217 (0.0261)	-0.0728** (0.0346)	0.102 *** (0.0274)	-0.6207*** (0.0358)	0.0918 *** (0.0289)
DP_{t-1}	37.5419*** (2.2178)	5.929 ** (2.5308)	18.2143*** (2.0525)	31.9434*** (2.8891)	35.0621*** (3.501)	14.6097*** (1.6354)	15.1424*** (2.171)	26.5778*** (1.7178)	-16.7307*** (2.2455)	18.9172*** (1.8109)
MKT^{US}_{t-1}	-0.0131 (0.0206)	-0.0346 (0.0235)	0.0453 ** (0.0191)	0.0416 (0.0269)	0.0993 *** (0.0326)	0.0091 (0.0152)	-0.0127 (0.0202)	0.0392** (0.016)	-0.2333*** (0.0209)	0.0397 ** (0.0169)
SMB^{US}_{t-1}	0.0728 *** (0.0205)	0.004 (0.0234)	-0.0093 (0.019)	-0.0257 (0.0267)	-0.1438*** (0.0324)	0.0099 (0.0151)	0.1267 *** (0.0201)	0.0016 (0.0159)	0.1052 *** (0.0208)	-0.019 (0.0168)
SMB^{JP}_{t-1}	-0.0556*** (0.0203)	-0.0201 (0.0231)	0.0304 (0.0188)	0.0325 (0.0264)	0.0765** (0.032)	0.0113 (0.0149)	-0.163 *** (0.0198)	-0.0138 (0.0157)	-0.4187*** (0.0205)	-0.0052 (0.0165)
HML^{JP}_{t-1}	-0.0034 (0.0226)	-0.019 (0.0258)	-0.0692*** (0.0209)	-0.0437 (0.0294)	0.0129 (0.0356)	$-9e-04$ (0.0167)	0.0494 ** (0.0221)	0.0233 (0.0175)	-0.2414*** (0.0229)	-0.0466** (0.0184)
u^{US}_{t-1}	12.4653*** (2.597)	13.462 *** (2.9635)	4.5257 * (2.4034)	12.4692*** (3.383)	19.5058*** (4.0996)	3.7171* (1.915)	10.8071*** (2.5422)	5.7265 *** (2.0115)	14.2514*** (2.6294)	12.2716*** (2.1205)
u^{GE}_{t-1}	7.0139 (5.5943)	10.0588 (6.3837)	-8.4888 (5.1773)	8.6778 (7.2875)	18.9269** (8.831)	2.9962 (4.1252)	1.9312 (5.4762)	6.9806 (4.333)	-5.7167 (5.6641)	20.1928*** (4.5679)

Continued on next page

Table 6 — continued from previous page

EPU^{US}_{t-1}	-0.0013 (0.0022)	0.0078^{***} (0.0025)	0.0028 (0.0021)	0.0138^{***} (0.0029)	0.0152^{***} (0.0035)	-0.0087^{***} (0.0016)	-0.0079^{***} (0.0022)	0.0051^{***} (0.0017)	0.0197^{***} (0.0023)	0.0097^{***} (0.0018)
EPU^{JP}_{t-1}	0.0043 (0.003)	-0.0017 (0.0034)	0.0131^{***} (0.0027)	0.0087^{**} (0.0039)	0.0056 (0.0047)	-0.0051^{**} (0.0022)	-0.008^{***} (0.0029)	-0.0069^{***} (0.0023)	-0.0068^{**} (0.003)	-0.0036 (0.0024)
EPU^{EU}_{t-1}	-0.0029 (0.002)	-0.0114^{***} (0.0023)	$9e-04$ (0.0018)	-0.0065^{**} (0.0026)	-0.0059^{*} (0.0031)	0.0058^{***} (0.0015)	0.0052^{***} (0.0019)	0.0021 (0.0015)	0.0214^{***} (0.002)	0.0017 (0.0016)
γ	1.9398	15.7955	100.0477	4.1681	2.4756	100.0005	42.1346	61.4113	22.0723	17.9613
c	4.4546	4.7538	5.2712	4.6020	4.4083	6.8594	5.8317	5.4265	7.3171	5.6202
σ	1.1237	1.2823	1.0400	1.4638	1.7739	0.8286	1.1000	0.8704	1.1377	0.9175
SSR	421.752	549.184	361.226	715.703	1050.981	229.334	404.138	253.014	432.350	281.189
LL	-513.918	-558.141	-487.970	-602.500	-666.856	-411.871	-506.772	-428.330	-518.075	-446.015
AIC	1127.836	1216.282	1075.940	1305.001	1433.712	923.742	1113.545	956.660	1136.151	992.030
BIC	1318.543	1406.988	1266.647	1495.707	1624.419	1114.448	1304.251	1147.367	1326.857	1182.737

*** indicates $p-$ value $< 1\%$, ** indicates $p-$ value $< 5\%$ and * indicates $p-$ value $< 10\%$; SSR is the sum of squared residuals, SE is the standard error of the regresssion, R^2 is the R-squared index, LL is the loglikelihood, AIC and BIC are the Akaike and Bayesian information criteria; DW is the Durbin-Watson statistic, LB(5) is the Ljung-Box test ststistic with 5 lags and ARCH(5) is the test for conditional heteroskedasticity with 5 lags.

Table 7: Significance of leading variables

	$y_{1,t}$	$y_{2,t}$	$y_{3,t}$	$y_{4,t}$	$y_{5,t}$	$y_{6,t}$	$y_{7,t}$	$y_{8,t}$	$y_{9,t}$	$y_{10,t}$
$\Delta\pi^{UK}_{t-1}$										
g^{US}_{t-1}		*				*				
g^{JP}_{t-1}						*		*	*	
DP_{t-1}	*	*	*	*	*	*	*	*	*	*
MKT^{US}_{t-1}	*	*	*	*		*			*	
SMB^{US}_{t-1}	*				*				*	
SMB^{JP}_{t-1}	*	*			*		*	*		*
HML^{JP}_{t-1}					*		*	*		*
u^{US}_{t-1}		*						*		
u^{GE}_{t-1}										*
EPU^{US}_{t-1}	*			*		*				
EPU^{JP}_{t-1}	*	*	*	*	*					*
EPU^{EU}_{t-1}		*	*	*	*				*	
	$y_{1,t}$	$y_{2,t}$	$y_{3,t}$	$y_{4,t}$	$y_{5,t}$	$y_{6,t}$	$y_{7,t}$	$y_{8,t}$	$y_{9,t}$	$y_{10,t}$
$\Delta\pi^{UK}_{t-1}$	*		*		*	*	*	*	*	*
g^{US}_{t-1}	*	*	*			*	*		*	*
g^{JP}_{t-1}	*	*	*	*	*		*	*	*	*
DP_{t-1}	*	*	*	*	*	*	*	*	*	*
MKT^{US}_{t-1}			*		*			*	*	*
SMB^{US}_{t-1}	*				*		*		*	
SMB^{JP}_{t-1}	*				*		*		*	
HML^{JP}_{t-1}			*				*		*	*
u^{US}_{t-1}	*	*		*	*		*	*	*	*
u^{GE}_{t-1}					*					*
EPU^{US}_{t-1}		*			*	*	*	*	*	*
EPU^{JP}_{t-1}			*	*		*	*	*	*	
EPU^{EU}_{t-1}		*		*		*	*		*	

Columns refer to equations in the `CholVLSTARX` model. The asterisk indicates that the variable is significant at least at 5% level.

5 Forecasting

In this section we present the results on the forecast accuracy of the proposed Cholesky-VLSTAR model for the realized covariance matrix. We consider one-step ahead forecasts from a rolling window estimation, with a rolling window of 223 observations. The whole sample is split into an in-sample subset from August 1990 to February 2009 (223 monthly observations) and an out-of-sample subset from March 2009 to June 2018 (112 monthly observations). At each step, the parameters of the model are re-estimated.

In order to evaluate the forecast accuracy of our nonlinear model with exogenous variables (`CholVLSTARX`), we compare the out-of-sample forecasts with those obtained from the following competing models:

- `CholVLSTAR`: the nonlinear model without macroeconomic and financial exogenous regressors,

- `CholVARX` and `CholVAR`: linear models on the Cholesky factors with and without macroeconomic and financial exogenous regressors,

- `VARX` and `VAR`: VAR(1) model on the realized covariances with and without macroeconomic and financial exogenous regressors,

- `logVARX` and `logVAR`: based on the log-volatilities introduced by Bauer and Vorkink (2011), with and without macroeconomic and financial exogenous regressors,

- `BEKK` and `DCC`: the most commonly used multivariate GARCH (see Engle and Kroner, 1995; Engle, 2002).

5.1 Forecasting evaluation: statistical measures

Direct and indirect methods are implemented to evaluate the forecast accuracy of the `CholVLSTARX` model. Statistical evaluation of the forecasts, such as the test for equal predictive accuracy, can be considered direct methods, while the indirect evaluation of the predictions is based on economic of financial loss functions, like the optimal portfolio allocation. In this section, we focus on direct methods.

Due to the latent nature of the object of interest (*i.e.* co-volatility is unobservable) direct methods shall rely on a proxy. We assume that the realized covariance is an unbiased co-volatility proxy, this allows us to apply direct methods, such as Diebold and Mariano (1995) test, Giacomini and White (2006) test and Mincer and Zarnowitz (1969) regression.

Moreover, direct evaluation of the forecasts involves the use of a loss function. As pointed out in Patton and Sheppard (2009), the use of "non-robust" loss functions may lead to a misleading ranking of forecasts (see Patton, 2011). Patton (2011) verifies that many commonly-used loss functions lead to severe biases when used with a noisy proxy. Thus, the statistical evaluation of the forecasts presented in this paper is based on "robust" functions, such as the univariate Mean Square Error (hereafter MSE) loss function, $L(\hat{\sigma}_t, h_t) = (\hat{\sigma}_t - h_t)^2$, where $\hat{\sigma}_t$ is the realized out-of-sample volatitlity and h_t is the forecast volatility. Similarly, the Frobenius norm between forecast covariance matrices, is defined as

$$L_F(\hat{\Sigma}_t, H_t) = \sum_{i=1}^{n}\sum_{j=1}^{n} L(\hat{\sigma}_{ij,t}, h_{ij,t}),$$

where i and j indicate the rows and the columns, therefore the above MSE loss function is calculated as all the element of the realized volatility matrix $\hat{\Sigma}_t$ and the forecast conditional volatility matrix H_t. Finally, the Euclidean distance between vectors $\tilde{n} = n(n+1)/2$ elements of the covariance matrices is defined as

$$L_E(\hat{\sigma}_t, h_t) = \sum_{i=1}^{\tilde{n}} (\hat{\sigma}_{i,t} - h_{i,t})^2.$$

Volatility forecasts are pairwise compared via Diebold-Mariano (DM) test and Giacomini-White (GW) test. The DM test is based on the difference, d_t, between the loss functions of two models f_{1t} and f_{2t} based on the forecast errors. However, since the estimated models are in large part nested, the standard Diebold and Mariano (1995) inference on the equal predictive accuracy is not valid, as shown by Clark and McCracken (2001) and Clark and West (2007), this because the statistics based on average comparisons of prediction errors have a degenerate limiting variance under the null hypothesis and they are not asymptotically normally distributed. To allow for a unified treatment of nested and non-nested models, Giacomini and

White (2006) (GW) suggest to approach the problem of the forecast evaluation as a problem of inference about conditional (rather than unconditional) expectations of forecast errors. The GW is a test of finite-sample predictive ability. The GW approach holds with a rolling window scheme and, in general, with a wider class of models respect to the DM test, including nonlinear, semi-parametric, parametric, nested and non-nested models.

Table 8: Out of sample forecast accuracy tests

	Diebold and Mariano (1995)		Giacomini and White (2006)	
Model (f_{2t})	Frobenius	Eucidean	Frobenius	Eucidean
`CholVLSTAR`	-1.318 (0.190)	-2.342^{**} (0.021)	-2.366^{**} (0.018)	-3.631^{***} (<0.001)
`CholVAR`	-1.526 (0.130)	-2.815^{***} (0.006)	-2.777^{***} (0.005)	-4.552^{***} (<0.001)
`CholVARX`	-2.103^{**} (0.038)	-2.631^{***} (0.009)	-2.912^{***} (0.004)	-3.967^{***} (<0.001)
`VAR`	-2.218^{**} (0.029)	-3.789^{***} (<0.001)	-3.849^{***} (<0.001)	-6.143^{***} (<0.001)
`VARX`	-1.974^{*} (0.051)	-3.321^{***} (0.001)	-3.461^{***} (<0.001)	-6.604^{***} (<0.001)
log`VAR`	-1.587 (0.115)	-2.227^{**} (0.028)	-2.195^{**} (0.031)	-2.995^{***} (0.006)
log`VARX`	-1.781^{*} (0.078)	-2.570^{**} (0.011)	-2.721^{***} (0.007)	-2.995^{***} (0.003)
`BEKK`	-2.184^{**} (0.031)	-4.369^{***} (<0.001)	-4.189^{***} (<0.001)	-8.735^{***} (<0.001)
`DCC`	-2.181^{**} (0.032)	-4.367^{***} (<0.001)	-4.187^{***} (<0.001)	-8.728^{***} (<0.001)

p-values in parentheses, * statistically significant at 10%, ** statistically significant at 5%, *** statistically significant at 1%

Table 8 provides the results of the tests for predicting accuracy with multivariate loss functions. The reference model is the Cholesky-VLSTAR with exogenous explanatory variables (`CholVLSTARX`) estimated in section 4. A negative and statistically significant test statistics means that the nonlinear model with exogenous variables performs better than the competing models. The results of the DM test with Euclidean loss function say that the `CholVLSTARX` model outperforms all models compared. Quite the same results are obtained with the Frobenius norm.

The GW test rejects the null of equal predictive accuracy in several cases. As for the DM test, the sign of the test statistics indicates that the `CholVLSTARX` model performs better than the competing models. When the loss function is the Euclidean distance, our proposed model strongly overperforms the competing models. Interestingly, the highest test statistics are shown in the comparison with the BEKK and DCC models.

Alternatively, forecasts may be compared for more than two models. For this purpose, Hansen, Lunde, and Nason (2011) introduced the model confidence set (MCS) to compare all forecasts against each other. For a given confidence level, the MCS defines the set of models containing the best out of sample forecasts. The MCS approach consists in a sequential test that allows us to test the equal predictive ability of the compared models, to discard any inferior model and to define the set of superior models (SSM). Given a set of H forecasts, the MCS procedure test whether all models in H have equal forecasting ability. The performance is measured pairwise via the loss functions difference, $d_{j_1,j_2,t} = L(\hat{y}_{j_1,t};\sigma_t) - L(\hat{y}_{j_2,t};\sigma_t)$, for all $j_1, j_2 = 1, 2, \ldots, H$ and $j_1 \neq j_2$. Assuming that $d_{j_1,j_2,t}$ is stationary, the null hypothesis takes the following form:

$$H_0 : E(d_{j_1,j_2,t}) = 0 \tag{19}$$

for each $j_1 \neq j_2$. A model is discarded if the null is rejected at a given confidence level α. The test is sequentially repeated until H_0 is not rejected. The remaining models define the set of statistically equivalent models with respect to a given loss function. Since MCS procedure strongly relies on the ordering imposed by the loss function, we implement the Frobenius and the Euclidean distances in a multivariate framework.

The ranking of the models through MCS are reported in Table 9 for the multivariate Frobenius and Euclidean loss functions. The Cholesky-VLSTAR model, with or without exogenous variables, stands in the top ranked models amongst the equal predictive models because they exhibit the highest probability of being included in the SSM. In practice, their inclusion occurs 10000 times on 10000 attempts. On the other hand, the linear models (`VAR` and `VARX`) and the multivariate GARCH models (`BEKK` and `DCC`) show the lowest probabilities. This leads us to conclude that combining the Cholesky decomposition with nonlinearity and proper exogenous variables improves the predictive accuracy. Thus, the nonlinear model introduced in this work (`CholVLSTARX`) appears to perform better than any other competing models, contrary to what the literature on nonlinear models emphasizes. Moreover, the use of a transformation of the realized covariance to ensure semi-positiveness seems to improve the predictive ability of the models, while the use of exogenous variables seems to slightly improve the forecast accuracy.

Table 9: MCS with $\alpha = 0.1$ and 10 000 bootstraps

Rank	Frobenius Model	Loss	P_{MCS}	Euclidean Model	Loss	P_{MCS}
1	CholVLSTARX	2678	1.000**	CholVLSTARX	32.61	1.000**
2	CholVLSTAR	4492	0.964**	logVARX	39.44	0.015
3	CholVARX	4458	0.899**	logVAR	40.13	0.009
4	logVARX	4657	0.196*	CholVLSTAR	40.36	0.001
5	logVAR	5173	0.164*	CholVARX	41.21	<0.001
6	CholVAR	5338	0.006	CholVAR	43.79	<0.001
7	VAR	7442	0.004	VAR	54.09	<0.001
8	VARX	6829	0.003	VARX	55.14	<0.001
9	BEKK	12823	0.001	BEKK	73.35	<0.001
10	DCC	12782	<0.001	DCC	73.27	<0.001

Note: The table displays the average loss over the testing sample as well as the Model Confidence Set p-values and the ranking in the SSM. * and ** indicate a probability of being in the SSM higher than 10% and 30% respectively.

5.2 Forecasting evaluation: financial measures

While direct methods are useful to rank volatility forecasts, indirect evaluation allows to measure the performance of the models in economic applications. In this section we focus on portfolio optimization in absolute risk-total return space (see Markowitz, 1952) that relies on the first two conditional moments of asset returns. In this context, using an accurate measure of forecast volatility is crucial for portfolio management or risk hedging. Hence, we use a Global Minimum Variance (GMV) approach. According to Merton (1980) GMV does not require modelling the conditional expected returns and several papers show that GMV portfolio performs better than mean-variance portfolio, see also Chan, Karceski, and Lakonishok (1999), Jagannathan and Ma (2003) and Kyj, Ostdiek, and Ensor (2009), when the covariance matrix is predicted. Assuming a risk adverse investor, the GMV portfolio weights at time t are the solution of the optimization problem

$$\begin{cases} \min\limits_{w} & w_t' \hat{\Sigma}_{t+1} w_t \\ \text{sub} & w_t' \iota = 1 \end{cases} \tag{20}$$

where ι is a vector of ones, $\hat{\Sigma}_{t+1}$ is the forecast covariance matrix and w_t is the $n \times 1$ vector of portfolio weights in the t-th month.

In order to assess the economic value of the forecasts produced by the model, we create a portfolio for each model considered in the direct methods. Table 10 reports the GMV portfolio performance. In addition to the models considered in the statistical evaluation, we consider also a naïve portfolio with $1/n$ weights, in line with the contribution of DeMiguel, Garlappi, and Uppal (2009). The portfolio weights are monthly rebalanced using the forecasts from the rolling window estimation.

Table 10: GMV Portfolio Performance

Portfolio	μ_p(%)	σ_p(%)	SR	Treynor	alpha	VaR(95%)	CVaR(95%)	Ledoit-Wolf	
								t-stat	p-value
`CholVLSTARX`	1.032	3.505	0.303	0.153	2.373	6.265*	7.496**	-	-
`CholVLSTAR`	0.974	3.362**	0.300	0.149	2.015	6.401	7.617	0.300	0.764
`CholVARX`	1.053	3.533	0.308**	0.154	2.511	6.423	7.644	-0.214	0.831
`CholVAR`	0.963	3.400	0.294	0.145	1.715	6.368	7.574	0.625	0.532
log`VARX`	1.011	3.495	0.298	0.149	2.025	6.593	7.756	0.327	0.743
log`VAR`	0.994	3.400	0.302	0.150	2.150	6.421	7.651	0.131	0.896
`VARX`	1.099*	3.771	0.302	0.157*	2.828*	7.254	8.066	0.141	0.888
`VAR`	0.982	3.419	0.298	0.149	1.991	6.361	7.637	0.399	0.690
`BEKK`	1.020	3.336*	0.315*	0.156**	2.578**	6.314**	7.535	-0.381	0.703
`DCC`	0.884	3.364	0.275	0.137	1.027	6.361	7.272*	0.843	0.399
`Naïve`	1.065**	3.693	0.300	0.152	2.474	6.377	8.216	0.199	0.842

Note: * and ** indicates the best performing portfolios for each measure

We use the Sharpe Ratio (SR), the Treynor Ratio and the Jensen's alpha, in order to evaluate the portfolio performance. We include also pure risk measures, such as the portfolio standard deviation (σ_P), the Value at Risk (VaR) and the Conditional Value at Risk (CVaR). The `CholVLSTARX` model exhibits the lowest Value at Risk (95%). Similarly, the Conditional Value at Risk (95%) is lower for `DCC` and `CholVLSTARX`. Only the portfolio obtained from a BEKK model seems to outperform the competing ones in terms of SR, Treynor and Jensen's alpha. However, the Ledoit and Wolf (2008) test for the equality of the Sharpe ratios of two investment strategies does not highlight any relevant difference between our proposed model and all the other competing approaches. From the results in Table 10, it may be concluded that a matrix parametrization may help in reducing the overall risk of a portfolio, nevertheless the risk-return performance of the compared portfolios remains quite similar.

6 Concluding remarks

In this paper, we introduce the Vector Logistic Smooth Transition (VLSTAR), a new statistical approach for the specification of multivariate conditional covariance matrices. Our proposed methodology benefits of the Cholesky decomposition in order to obtain positive definite estimated covariance matrices, and the possibility to add macroeconomic and financial variables as exogenous explanatory variables. Moreover, we apply a nonlinear version with changes in regime in order to account for asymmetric dynamics of volatility.

We provide evidence that our proposed model is able to significantly improve the out-of-sample volatility forecasts, compared to the standard techniques as, for example, the multivariate GARCH models, the linear `VAR` models or the `logVAR` models. A portfolio optimization exercise is also carried out in order to assess the accuracy of the forecasts in economic applications. The results confirm that, in terms of risk-return performance, the portfolio obtained by using the Cholesky-VLSTAR approach is comparable and sometimes seems to perform better than those obtained via the competing models. Variables capturing time-varying risk show up as robust predictors for Cholesky factors. Conversely, macroeconomic variables are less informative about future volatility. Economic Policy Uncertainty, instead, plays a crucial role in the definition of the movements of volatility.

Our work leaves space for future work and further studies. A possible extension of this work should check its feasibility in a higher dimensional framework. Future research in this field could be conducted on the use of machine learning algorithms to identify nonlinear dynamics in a more parsimonious way than smooth transition model.

References

Ait-Sahalia, Y., and J. Jacod (2014): *High-frequency financial econometrics.* Princeton University Press.

Ajmi, A. N., G. C. Aye, M. Balcilar, G. E. Montasser, and R. Gupta (2015): "Causality Between US Economic Policy and Equity Market Uncertainties: Evidence from Linear and Nonlinear Tests," *Journal of Applied Economics*, 18(2), 225 – 246.

Amengual, D., and D. Xiu (2018): "Resolution of policy uncertainty and sudden declines in volatility," *Journal of Econometrics*, 203(2), 297 – 315.

Andersen, T. G., T. Bollerslev, F. X. Diebold, and P. Labys (2003): "Modeling and Forecasting Realized Volatility," *Econometrica*, 71, 579–625.

Anderson, H. M., and F. Vahid (1998): "Testing multiple equation systems for common nonlinear components," *Journal of Econometrics*, 84, 1–36.

Antonakakis, N., I. Chatziantoniou, and G. Filis (2013): "Dynamic comovements of stock market returns, implied volatility and policy uncertainty," *Economics Letters*, 120, 87–92.

Aue, A., S. Hörmann, L. Horváth, and M. Reimherr (2009): "Break detection in the covariance structure of multivariate time series models," *The Annals of Statistics*, 37, 4046–4087.

Bacon, D. W., and D. G. Watts (1971): "Estimating the transition between two intersecting straight lines," *Biometrika*, 58, 525–534.

Bai, J., R. L. Lumsdaine, and J. H. Stock (1998): "Testing For and Dating Common Breaks in Multivariate Time Series," *Review of Economic Studies*, 65, 395–432.

Bai, J., and P. Perron (1998): "Estimating and Testing Linear Models With Multiple Structural Changes," *Econometrica*, 66, 47–78.

——— (2003): "Computation and Analysis of Multiple Structural Change Models," *Journal of Applied Econometrics*, 18, 1–22.

Baker, S., N. Bloom, and S. Davis (2015): "Measuring Economic Policy Uncertainty," Centre for Economic Performance, Discussion Paper No. 1379.

Baker, S. R., N. Bloom, and S. J. Davis (2016): "Measuring Economic Policy Uncertainty," *The Quarterly Journal of Economics*, 131(4), 1593–1636.

Barndorff-Nielsen, O. E., and N. Shephard (2002): "Econometric analysis of realised volatility and its use in estimating stochastic volatility models," *Journal of the Royal Statistical Society*, 64(2), 253–280.

Bastianin, A., and M. Manera (2018): "How does stock market volatility react to oil price shocks?," *Macroeconomic Dynamics*, 22(3), 666–682.

Bauer, G. H., and K. Vorkink (2011): "Forecasting multivariate realized stock market volatility," *Journal of Econometrics*, 160(1), 93–101.

Becker, R., A. Clements, and R. O'Neill (2010): "A Cholesky-MIDAS model for predicting stock portfolio volatility," Discussion paper, Working paper, Centre for Growth and Business Cycle Research Discussion Paper Series.

Beirne, J., G. M. Caporale, M. Schulze-Ghattas, and N. Spagnolo (2009): "Volatility spillovers and contagion from mature to emerging stock markets," Working Paper Series 1113, European Central Bank.

Bollerslev, T., A. Patton, and R. Quaedvlieg (2017): "Realized SemiCovariances: Looking for Signs of Direction Inside the Covariance Matrix," Economic Research Initiatives at Duke, Working Papers Series.

Brogaard, J., and A. Detzel (2015): "The Asset-Pricing Implications of Government Economic Policy Uncertainty," *Management Science*, 61(1), 3–18.

Bucci, A. (2019a): "Cholesky-ANN models for predicting multivariate realized volatility," MPRA Working paper No. 95137.

——— (2019b): "Realized Volatility Forecasting with Neural Networks," MPRA Working paper No. 95443.

Camacho, M. (2004): "Vector smooth transition regression models for US GDP and the composite index of leading indicators," *Journal of Forecasting*, 23, 173–196.

Chan, K., and T. Howell (1986): "On Estimating Thresholds in Autoregressive Models," *Journal of Time Series Analysis*, 7(3), 179–190.

Chan, L. K., J. Karceski, and J. Lakonishok (1999): "On Portfolio Optimization: Forecasting Covariances and Choosing the Risk Model," *The Review of Financial Studies*, 12, 937–974.

Christiansen, C., M. Schmeling, and A. Schrimpf (2012): "A comprehensive look at financial volatility prediction by economic variables," *Journal of Applied Econometrics*, 27(6), 956–977.

Clark, T. E., and M. W. McCracken (2001): "Tests of equal forecast accuracy and encompassing for nested models," *Journal of Econometrics*, 105(1), 85 – 110, Forecasting and empirical methods in finance and macroeconomics.

Clark, T. E., and K. D. West (2007): "Approximately normal tests for equal predictive accuracy in nested models," *Journal of Econometrics*, 138(1), 291–311.

Conrad, C., and K. Loch (2014): "Anticipating Long-Term Stock Market Volatility," *Journal of Applied Econometrics*, 30(7), 1090–1114.

DeMiguel, V., L. Garlappi, and R. Uppal (2009): "Optimal Versus Naive Diversification: How Inefficient is the 1/N Portfolio Strategy?," *Review of Financial Studies*, 22(5), 1915–1953.

Diebold, F. X., and R. S. Mariano (1995): "Comparing predictive accuracy," *Journal of Business and Economic Statistics*, 13, 253–263.

Diebold, F. X., and K. Yilmaz (2009): "Measuring Financial Asset Return and Volatility Spillovers, with Application to Global Equity Markets," *Economic Journal*, 119, 158–171.

Engle, R. F. (2002): "Dynamic conditional correlation: a simple class of multivariate GARCH models," *Journal of Business and Economic Statistics*, 20, 339–350.

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

FAMA, E., AND K. FRENCH (1993): "Common Risk Factors in the Returns on Stocks and Bonds," *Journal of Financial Economics*, 33, 3–56.

GIACOMINI, R., AND H. WHITE (2006): "Tests of conditional predictive ability," *Econometrica*, 74(6), 1545–1578.

HAGGAN, V., AND T. OZAKI (1981): "Modelling Nonlinear Random Vibrations Using an Amplitude-Dependent Autoregressive Time Series Model," *Biometrika*, 68(1), 189–196.

HALBLEIB-CHIRIAC, R., AND V. VOEV (2011): "Modelling and Forecasting Multivariate Realized Volatility," *Journal of Applied Econometrics*, 26, 922–947.

HANSEN, P. R., A. LUNDE, AND J. M. NASON (2011): "The Model Confidence Set," *Econometrica*, 79(2), 435–497.

HUBRICH, K., AND T. TERÄSVIRTA (2013): "Thresholds and Smooth Transitions in Vector Autoregressive Models," CREATES Research Paper, Aarhus University.

JAGANNATHAN, R., AND T. MA (2003): "Risk Reduction in Large Portfolios: Why Imposing the Wrong Constraints Helps," *Journal of Finance*, 58(4), 1651–1683.

KANG, W., AND R. A. RATTI (2013): "Oil shocks, policy uncertainty and stock market return," *Journal of International Financial Markets, Institutions and Money*, 26, 305 – 318.

KOCK, A. B., AND T. TERÄSVIRTA (2011): "Forecasting with nonlinear time series models," CREATES Research Paper, Aarhus University.

KYJ, L., B. OSTDIEK, AND K. ENSOR (2009): "Realized Covariance Estimation in Dynamic Portfolio Optimization," Working Paper.

LAM, S.-S., H. ZHANG, AND W. ZHANG (2019): "Does Policy Instability Matter for International Equity Markets?," *International Review of Finance*, 0(0), 1–10.

LEDOIT, O., AND M. WOLF (2008): "Robust performance hypothesis testing with the Sharpe ratio," *Journal of Empirical Finance*, 15, 850–859.

LIU, L., AND T. ZANG (2015): "Economic policy uncertainty and stock market volatility," *Finance Research Letters*, 15, 99–105.

Luukkonen, R., P. Saikkonen, and T. Teräsvirta (1988): "Testing Linearity Against Smooth Transition Autoregressive Models," *Biometrika*, 75, 491–499.

Markowitz, H. (1952): "Portfolio selection," *Journal of Finance*, 7(1), 77–91.

Martens, M., M. De Pooter, and D. J. Van Dijk (2004): "Modeling and Forecasting S&P 500 Volatility: Long Memory, Structural Breaks and Nonlinearity," Tinbergen Institute Discussion Paper No. 04-067/4.

May, R. M. (1976): "Simple mathematical models with very complicated dynamics," *Nature*, 261, 459–467.

McAleer, M., and M. Medeiros (2008): "A multiple regime smooth transition Heterogeneous Autoregressive model for long memory and asymmetries," *Journal of Econometrics*, 147(1), 104–119.

Mele, A. (2007): "Asymmetric stock market volatility and the cyclical behavior of expected returns," *Journal of Financial Economics*, 86, 446–478.

——— (2008): "Understanding Stock Market Volatility - A Business Cycle Perspective," Working Paper.

Merton, R. C. (1980): "On estimating the expected return on the market: An explanatory investigation," *Journal of Financial Economics*, 8, 323–361.

Mincer, J. A., and V. Zarnowitz (1969): "The Evaluation of Economic Forecasts," in *Economic Forecasts and Expectations: Analysis of Forecasting Behavior and Performance*, NBER Chapters, pp. 3–46. National Bureau of Economic Research, Inc.

Pástor, L., and P. Veronesi (2012): "Uncertainty about Government Policy and Stock Prices," *The Journal of Finance*, 67(4), 1219–1264.

Patton, A. J. (2011): "Volatility forecast comparison using imperfect volatility proxies," *Journal of Econometrics*, 160(1), 246 – 256.

Patton, A. J., and K. Sheppard (2009): *Evaluating Volatility and Correlation Forecasts*in T. Mikosch, J.P. Kreiß, A. Richard and G.T. Andersen (*eds.*), Handbook of Financial Time Series, Springer Berlin Heidelberg.

PAYE, B. S. (2012): "Dèjà vol: Predictive regressions for aggregate stock market volatility using macroeconomic variables," *Journal of Financial Economics*, 106(3), 527–546.

QUANDT, R. E. (1958): "The Estimation of the Parameters of a Linear Regression System Obeying Two Separate Regimes," *Journal of the American Statistical Association*, 53(284), 873–880.

SCHWERT, W. G. (1989): "Why does stock market volatility change over time?," *Journal of Finance*, 44(5), 1115–1153.

SUM, V., AND F. FANTA (2012): "Long-run relation and speed of adjustment of economic policy uncertainty and excess return volatility," *Int.Res.J.Finance Econ.*, 102, 6–12.

TERÄSVIRTA, T. (1994): "Specification, estimation and evaluation of vector smooth transition autoregressive models," *Journal of the American Statistical Association*, 89, 208–218.

TERÄSVIRTA, T., D. TJØSTHEIM, AND C. W. GRANGER (2010): *Modelling Nonlinear Economic Time Series*. Oxford University Press.

TERÄSVIRTA, T., AND Y. YANG (2014): "Specification, estimation and evaluation of vector smooth transition autoregressive models with applications," Discussion paper, Department of Economics and Business Economics, Aarhus University, CREATES Research Paper, Aarhus University.

TONG, H. (1990): *Non-linear time series: a dynamical system approach*. Oxford University Press.

WELCH, I., AND A. GOYAL (2008): "A comprehensive look at the empirical performance of equity premium prediction," *Review of Financial Studies*, 21(4), 1455–1508.

XIONG, Z., AND L. HAN (2015): "Volatility spillover effect between financial markets: evidence since the reform of the RMB exchange rate mechanism," *Financial Innovation*, 1(1), 1–9.

YANG, Y., AND T. TERÄSVIRTA (2013): "Linearity and misspecification tests for vector smooth transition regression models," CREATES Research Paper, Aarhus University.

Appendix: Joint linearity test

Given a VLSTAR model with a unique transition variable, $s_{1t} = s_{2t} = \cdots = s_{\widetilde{n}t} = s_t$, a generalization of the linearity test presented in Luukkonen, Saikkonen, and Teräsvirta (1988) may be implemented.

Assuming a 2-state VLSTAR model, such that

$$y_t = B_1 z_t + G_t B_2 z_t + \varepsilon_t. \tag{21}$$

Where the null $H_0 : \gamma_j = 0$, $j = 1, \ldots, \widetilde{n}$, is such that $G_t \equiv (1/2)/\widetilde{n}$ and the (21) are linear. When the null cannot be rejected, an identification problem of the parameter c_j in the transition function emerges, that can be solved through a first-order Taylor expansion around $\gamma_j = 0$.

The approximation of the logistic function with a first-order Taylor expansion is given by

$$\begin{aligned} G(s_t; \gamma_j, c_j) &= (1/2) + (1/4)\gamma_j(s_t - c_j) + r_{jt} \\ &= a_j s_t + b_j + r_{jt} \end{aligned} \tag{22}$$

where $a_j = \gamma_j/4$, $b_j = 1/2 - a_j c_j$ and r_j is the error of the approximation. If G_t is specified as follows

$$\begin{aligned} G_t &= \operatorname{diag}\{a_1 s_t + b_1 + r_{1t}, \ldots, a_{\widetilde{n}} s_t + b_{\widetilde{n}} + r_{\widetilde{n}t}\} \\ &= A s_t + B + R_t \end{aligned} \tag{23}$$

where $A = \operatorname{diag}\{a_1, \ldots, a_{\widetilde{n}}\}$, $B = \operatorname{diag}\{b_1, \ldots, b_{\widetilde{n}}\}$ and $R_t = \operatorname{diag}\{r_{1t}, \ldots, r_{\widetilde{n}t}\}$, y_t

can be written as

$$
\begin{aligned}
y_t &= B_1 z_t + (A s_t + B + R_t) B_2 z_t + \varepsilon_t \\
&= (B_1 + B B_2) z_t + A B_2 z_t s_t + R_t B_2 z_t + \varepsilon_t \\
&= \Theta_0 z_t + \Theta_1 z_t s_t + \varepsilon_t^*
\end{aligned} \tag{24}
$$

where $\Theta_0 = B_1 + B_2' B$, $\Theta_1 = B_2' A$ and $\varepsilon_t^* = R_t B_2 + \varepsilon_t$. Under the null, $\Theta_0 = B_1$ and $\Theta_1 = 0$, while the (24) model is linear, with $\varepsilon_t^* = \varepsilon_t$. It follows that the Lagrange multiplier test, under the null, is derived from the score

$$
\frac{\partial \log L(\widetilde{\theta})}{\partial \Theta_1} = \sum_{t=1}^{T} z_t s_t (y_t - \widetilde{B}_1 z_t)' \widetilde{\Omega}^{-1} = S(Y - Z\widetilde{B}_1)\widetilde{\Omega}^{-1}, \tag{25}
$$

where

$$
Y = \begin{bmatrix} y_1' \\ \vdots \\ y_t' \end{bmatrix}, \quad Z = \begin{bmatrix} z_1' \\ \vdots \\ z_t' \end{bmatrix}, \quad S = \begin{bmatrix} z_1' s_1 \\ \vdots \\ z_t' s_t \end{bmatrix}
$$

and where $\widetilde{B}_1$ and $\widetilde{\Omega}$ are estimated from the model in H_0. If $P_Z = Z(Z'Z)^{-1}Z'$ is the projection matrix of Z, the LM test is specified as follows

$$
LM = tr\big\{\widetilde{\Omega}^{-1}(Y - Z\widetilde{B}_1)' S \big[S'(I_t - P_Z)S\big]^{-1} S'(Y - Z\widetilde{B}_1)\big\}. \tag{26}
$$

Under the null, the test statistics is distributed as a χ^2 with $\widetilde{n}(p \cdot \widetilde{n} + k)$ degrees of freedom.