Volatility forecasting using global stochastic financial trends extracted from non-synchronous data

Lyudmila Grigoryeva
Juan-Pablo Ortega
Anatoly Peresetsky

August 2016
Graduate School of Decision Sciences

All processes within our society are based on decisions – whether they are individual or collective decisions. Understanding how these decisions are made will provide the tools with which we can address the root causes of social science issues.

The GSDS offers an open and communicative academic environment for doctoral researchers who deal with issues of decision making and their application to important social science problems. It combines the perspectives of the various social science disciplines for a comprehensive understanding of human decision behavior and its economic and political consequences.

The GSDS primarily focuses on economics, political science and psychology, but also encompasses the complementary disciplines computer science, sociology and statistics. The GSDS is structured around four interdisciplinary research areas: (A) Behavioural Decision Making, (B) Intertemporal Choice and Markets, (C) Political Decisions and Institutions and (D) Information Processing and Statistical Analysis.
Volatility forecasting using global stochastic financial trends extracted from non-synchronous data

Lyudmila Grigoryeva†, Juan-Pablo Ortega²,³, and Anatoly Peresetsky⁴

Abstract

This paper introduces a method based on various linear and nonlinear state space models that are used to extract global stochastic financial trends (GST) out of non-synchronous financial data. More specifically, these models are constructed to take advantage of the intraday arrival of closing information coming from different international markets to improve volatility description and forecasting. A set of three major asynchronous international stock market indices is used in order to empirically show that this forecasting scheme is capable of significant performance gains when compared to standard models like the dynamic conditional correlation (DCC) family.

Keywords: multivariate volatility modeling and forecasting, global stochastic trend, extended Kalman filter, dynamic conditional correlations (DCC), non-synchronous data.

Acknowledgments: LG acknowledges partial financial support from the Faculty for the Future Program of the Schlumberger Foundation. LG and JPO acknowledge partial financial support of the Région de Franche-Comté (Convention 2013C-5493) as well as the hospitality of the International Laboratory of Quantitative Finance of the Higher School of Economics, Moscow, that made possible the collaborative work that lead to the results presented in this paper. The research of AP was funded by the Russian Federation government grant 14.A12.31.0007.
# Contents

1 Introduction .................................................................................. 3

2 State-space model for the global stochastic trend ................... 5
   2.1 The linear and nonlinear state-space models ....................... 6
   2.2 The linear and extended Kalman filters for state and parameter estimation .......... 7

3 Volatility forecasting using the global stochastic trend .......... 10
   3.1 GST-based volatility forecasting using the nonlinear state-space model ......... 11
   3.2 GST-based volatility forecasting using the linear state-space model ............. 12

4 Empirical performance of the GST-based volatility forecasting schemes .......... 16
   4.1 Dataset and competing models ........................................... 16
   4.2 Volatility forecasting study .............................................. 18
   4.3 Model confidence sets based on covariance and KLIC loss functions .......... 19

5 Conclusions ............................................................................... 25

Bibliography .................................................................................. 26

A Appendices and supplementary material .................................. 28
   A.1 Proof of Proposition 2.1 ..................................................... 28
   A.2 Proof of Proposition 3.1 ..................................................... 31
   A.3 Empirical performance of the GST-based volatility forecasting schemes using a smaller estimation sample .................................................. 32
1 Introduction

Many frameworks for the description of financial returns have as their first building block a factor model of the form

$$r_t = \alpha + \beta y_t + u_t \quad \text{with} \quad \{u_t\} \sim \text{WN}(0, \sigma^2),$$

where the instantaneous returns $r_t$ at time $t$ of individual assets are presented as a mean-zero stochastic stationary additive perturbation of an affine function of a common factor $y_t$. This factor usually accounts for a common market feature to which all the assets under study are exposed. Consequently, this functional dependence allows to determine, for a given asset return $r_t$, how much of it has to do with the market situation (through the coefficient $\beta$, which is a function of the correlation between $r_t$ and $y_t$) and how much comes from an idiosyncratic perturbation $u_t$ specifically related to the individual asset. In particular applications of this model, the factor values $y_t$ are sometimes computed by using an index constructed out of a set of assets that represent the class to which $r_t$ is naturally associated.

An alternative to this approach consists of treating the common factor values as a non-observable variable and of extracting them using observed individual returns via a Kalman-type state-space model. This direction has already been profusely explored in the literature. In the early available works (see for instance [Jeon 91, Kasa 92, Chun 94, Sikl 01, Rang 01, Rang 02, Phen 04]) the authors consider low sampling frequencies in order to be able to neglect asynchronicity issues. In several more recent references [Dung 01, Chou 07, Chan 09, Luce 11, Cart 11, Bae 12, Feli 12, Bent 15] daily quoted data are used but always using a synchronized approach.

In this work we use a different point of view introduced in [Korh 13] and later on extended in [Durd 14, Pere 15], in which the market returns are thought of as a non-observable global stochastic trend (GST) whose value is ruled by the arrival of information coming from different local markets. In this framework, the returns of the GST are estimated several times per day at time points that are synchronized with the closing times of the markets that are assumed to drive it. This approach is implemented by setting a state-space model in which the observation equation writes the different observable individual market returns that we are interested in as a stochastic perturbation of an affine function of the estimated GST return accumulated during the 24 hours that precede this quote. It is assumed that the observed returns are those that drive the GST and hence its returns are estimated as many times per day as different closing times are included in the list of markets considered.

This point of view has been studied in [Korh 13] using three different setups, namely: three world indices (NIKKEI, MICEX, S&P500) with three different closing times, five world indices (NIKKEI, MICEX, DAX, PX, S&P500) with four different closing times, and ten world indices (NIKKEI, HSI, SENSEX, MICEX, DAX, PX, FTSE, IBOV, DJI, S&P500) with seven different closing times. The estimates of the GST obtained in these different situations are remarkably similar. The robustness that these results indicate allowed the authors to identify, for each market, the relative importance of local with respect to global news in stock prices formation.

The main goal of our work is modifying this approach in order to make it amenable to volatility forecasting and to prove the pertinence of the resulting method when compared to more standard families of models designed to specifically carry out this task. The rationale behind this attempt is that the error inherent to the filtering and forecasting of an unobserved variable like the GST is compensated by the more frequent information updates that the use of asynchronicity information carries in its wake. Since the models introduced in [Korh 13, Pere 15] are intrinsically homoscedastic, they are not appropriate to handle financial volatility modeling and forecasting. The heteroscedastic generalization needed for this purpose can be naturally implemented by using two different approaches. The simplest one consists of using the linear state-space approach in [Korh 13] in a first step to estimate the GST and to subsequently model the volatility and conditional correlation of the resulting global trend and idiosyncratic term using an adapted multivariate correlation model; for this purpose, we consider in this work adapted scalar and non-scalar versions of the dynamic conditional correlation (DCC) model.
introduced in [Engl 02, Tse 02]. The non-scalar models are estimated using the techniques introduced in [Chre 14, Bauw 15]. The adjustments of these standard models for the handling of the GST are implemented at the level of the so called “deGARCHing” or “first estimation step” in which a model for the conditional variances of the assets of interest is chosen; in our context, we put to work in this step two different GARCH-type models that take into account in their specification the chronology with which the different intraday trend returns are quoted.

A more sophisticated approach that we also study is the inclusion of the heteroscedasticity assumption on the GST returns directly in the formulation of the state-space model by using a GARCH-type and GST-adapted prescription of the type that we just described. The main complication that arises in this setup is the nonlinearity of the resulting modeling scheme that we handle using the extended Kalman filter (EKF) (see [Durb 12] and references therein).

The paper is organized as follows. In Section 2.1 we explain in detail the linear and nonlinear state-space models that we propose. We give details on how they handle the asynchronous character of the observable data and prove rigorous sufficient conditions that ensure their proper identification. Section 2.2 contains details on the Kalman filter based model estimation techniques that we use in the paper, as well as on the model specifications for the conditional variances incorporated in the nonlinear state-space model, together with the positivity and stationarity constraints that need to be imposed at the time of estimation. The GST-based volatility forecasting scheme is described in Section 3. Section 4 contains an empirical study using the adjusted closing values of three major indices (NIKKEI, FTSE, and S&P500) that are quoted at different times due to the time zones in which they are geographically based. In this experiment, we use the model confidence set (MCS) approach of [Hans 03, Hans 11] and we implement it with loss functions constructed with the conditional covariance matrices implied by the different models under consideration. The results show that the proposed forecasting scheme exhibits excellent and statistically significant performance improvements when compared to the use of standard multivariate parametric correlation models. Section 5 concludes the paper. The proofs of various technical results in the paper are contained in the appendices in Section A. Additionally, the last appendix in this section contains a replicate of the empirical study in Section 4 in which the estimation period does not include the Fall 2008 volatility events and the out-of-sample forecasting period comprises the Great Recession; this experiments aims at illustrating the robustness of our empirical results with respect to the estimation period used and the pertinence in some situations of the nonlinear state-space model.

Notation and conventions: column vectors are denoted by a bold lower case symbol like \( v \) and \( v^\top \) indicates its transpose. Given a vector \( v \in \mathbb{R}^n \), we denote its entries by \( v_i \), with \( i \in \{1, \ldots, n\} \); we also write \( v = (v_i)_{i \in \{1, \ldots, n\}} \). The symbols \( 1_n, 0_n \in \mathbb{R}^n \) stand for the vectors of length \( n \) consisting of ones and zeros, respectively. We denote by \( \mathbb{M}_{n,m} \) the space of real \( n \times m \) matrices with \( m, n \in \mathbb{N} \). When \( n = m \), we use the symbols \( \mathbb{M}_n \) and \( \mathbb{D}_n \) to refer to the space of square and diagonal matrices of order \( n \), respectively. Given a matrix \( A \in \mathbb{M}_{n,m} \), we denote its components by \( A_{ij} \) and we write \( A = (A_{ij}) \), with \( i \in \{1, \ldots, n\}, j \in \{1, \ldots, m\} \). We use \( \mathbb{S}_n \) to denote the subspace \( \mathbb{S}_n \subset \mathbb{M}_n \) of symmetric matrices:

\[
\mathbb{S}_n = \{ A \in \mathbb{M}_n \mid A^\top = A \},
\]

and we use \( \mathbb{S}_n^+ \) (respectively \( \mathbb{S}_n^- \)) to refer to the cone \( \mathbb{S}_n^+ \subset \mathbb{S}_n \) (respectively \( \mathbb{S}_n^- \subset \mathbb{S}_n \)) of positive (respectively negative) semidefinite matrices. When \( A \in \mathbb{S}_n^+ \) (respectively, \( A \in \mathbb{S}_n^- \)) we write \( A \succeq 0 \) (respectively, \( A \preceq 0 \)). The symbol \( I_n \in \mathbb{D}_n \) denotes the identity matrix. Given two matrices \( A, B \in \mathbb{M}_{n,m} \), we denote by \( A \odot B \in \mathbb{M}_{n,m} \) their elementwise multiplication matrix or Hadamard product, that is:

\[
(A \odot B)_{ij} := A_{ij} B_{ij} \text{ for all } i \in \{1, \ldots, n\}, \ j \in \{1, \ldots, m\}.
\]  

We denote as \( \text{Diag} \) the operator \( \text{Diag} : \mathbb{M}_n \rightarrow \mathbb{D}_n \) that sets equal to zero all the components of a square matrix except for those that are on the main diagonal. The operator \( \text{diag} : \mathbb{R}^n \rightarrow \mathbb{D}_n \) takes a given vector and constructs a diagonal matrix with its entries in the main diagonal.
2 State-space model for the global stochastic trend

We start by recalling the description of the global stochastic trend as the state variable in a state-space model, as it was introduced in [Korh 13]. In order to keep the presentation simple, we will carry this out for only three different non-synchronous assets, which is the framework in which our subsequent empirical analysis takes place. The generalization to more assets and quoting times is straightforward.

Let \( r_t \in \mathbb{R}^3 \) be a vector containing three non-synchronous stock market returns (typically based on adjusted closing prices) quoted at different times of the same calendar date \( t \in \mathbb{N} \). The different intraday quoting times have typically to do with lags in the closing times of the different markets. The intraday moments of time \( t_i, i \in \{1, 2, 3\} \) of the given day \( t \) at which the components of \((r_{1,t}, r_{2,t}, r_{3,t})^\top\) of \( r_t \) become available are labeled as \( t_i := 3(t-1) + i, t \in \mathbb{N} \).

We now assume the existence of an underlying and non-observable global stochastic trend and we denote by \( s_{t_i}, i \in \{1, 2, 3\} \), its intraday log-values for the given calendar date \( t \in \mathbb{N} \). We now define \( \varepsilon_t \in \mathbb{R}^3 \) as the vector that contains the intra-day stochastic trend log-return components of a given calendar day \( t \), that is,

\[
\varepsilon_t = \begin{pmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \\ \varepsilon_{3,t} \end{pmatrix} := \begin{pmatrix} s_{t_1} - s_{(t-1)_3} \\ s_{t_2} - s_{t_1} \\ s_{t_3} - s_{t_2} \end{pmatrix}.
\tag{2.1}
\]

Following the factor model scheme discussed in the introduction, for every \( t \in \mathbb{Z} \) we write the log-returns of the components of \( r_t \) as excess returns with respect to an affine function of the entries of the vector \( y_t \in \mathbb{R}^3 \), which is constructed with the daily global stochastic trend returns computed at the moments in time in which the components of \( r_t \) are quoted. More specifically \( y_t := (s_{t_1} - s_{(t-1)_1}, s_{t_2} - s_{(t-1)_2}, s_{t_3} - s_{(t-1)_3})^\top \) and

\[
r_{i,t} = \alpha_i + \beta_i y_{i,t} + u_{i,t}, \quad i = \{1, 2, 3\}, \ t \in \mathbb{Z},
\tag{2.2}
\]

with the regression intercepts \( \alpha_i \in \mathbb{R}, i = \{1, 2, 3\} \), and the parameters \( \beta := (\beta_1, \beta_2, \beta_3)^\top \in \mathbb{R}^3 \). For the time being, in this relation we only assume that the residuals \( u_i \) are serially uncorrelated (they are a white noise) with mean zero and unconditional diagonal covariance matrix \( \Sigma_u \in \mathbb{S}_3^+ \), that is \( \{u_t\} \sim \text{WN}(0, \Sigma_u) \), \( \Sigma_u := \text{diag}(\sigma_{u,1}^2, \sigma_{u,2}^2, \sigma_{u,3}^2) \) with entries \( \sigma_{u,1}, \sigma_{u,2}, \sigma_{u,3} \in \mathbb{R}^+ \).

Using the definition (2.1), we write the returns \( r_t \) in (2.2) in terms of the global stochastic trend returns in the preceding twenty-four hours. It is easy to verify that using (2.1) in the regression expression (2.2) yields the following identity

\[
r_t = \alpha + B \varepsilon_t + u_t,
\tag{2.3}
\]

with \( \varepsilon_t := (\varepsilon_{1,t}, \varepsilon_{2,t}, \varepsilon_{3,t}, \varepsilon_{2,t-1}, \varepsilon_{3,t-1})^\top \), where \( \alpha \in \mathbb{R}^3 \), \( \{u_t\} \sim \text{WN}(0, \Sigma_u) \) as in (2.2), and the matrix \( B \in \mathbb{M}_{5,3} \) is of the form

\[
B := \begin{pmatrix} \beta_1 & 0 & 0 & \beta_1 \\ \beta_2 & \beta_2 & 0 & \beta_2 \\ \beta_3 & \beta_3 & \beta_3 & 0 \end{pmatrix}.
\tag{2.4}
\]

This equation describes the returns dynamics in terms of the non-observable global stochastic trend returns in the preceding twenty-four hours. In order to estimate this model and to filter out of it the values of the GST, we will proceed by considering (2.3) as the observation equation of several linear and nonlinear state-space models that we design in order make possible their subsequent use for volatility forecasting.
2.1 The linear and nonlinear state-space models

The linear state-space model. The first model that we present in this subsection is identical to the one originally considered in [Korh 13]:

\[
\begin{align*}
\mathbf{r}_t &= \alpha + B\mathbf{e}_t + \mathbf{u}_t, & \mathbf{u}_t &\sim \text{WN} (0, \Sigma_u), \text{ with } \Sigma_u := \text{diag} (\sigma^2_{u,1}, \sigma^2_{u,2}, \sigma^2_{u,3}), \quad (2.5a) \\
\mathbf{e}_t &= T\mathbf{e}_{t-1} + R\mathbf{v}_{t-1}, & \mathbf{v}_t &\sim \text{WN}(0, I_3), \quad (2.5b)
\end{align*}
\]

where \( \mathbf{e}_t \in \mathbb{R}^5, \alpha \in \mathbb{R}^3 \), the matrix \( B \in \mathbb{M}_{5,3} \) is provided in (2.4), the matrices \( R \in \mathbb{M}_{3,5} \) and \( T \in \mathbb{M}_5 \) are given by

\[
R := \begin{pmatrix}
\sigma_{u,1} & 0 & 0 \\
0 & \sigma_{v,2} & 0 \\
0 & 0 & \sigma_{v,3}
\end{pmatrix}, \quad T := \begin{pmatrix}
0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0
\end{pmatrix}, \quad (2.6)
\]

and \( \sigma_{u,1}, \sigma_{u,2}, \sigma_{u,3}, \sigma_{v,1}, \sigma_{v,2}, \sigma_{v,3} \in \mathbb{R}^+. \) We emphasize that (2.5a)-(2.5b) constitutes a linear state-space model in which the dynamical behavior of the GST is prescribed by the corresponding state equation and where the observation equation establishes a relation between the time evolution of the GST and the observed returns.

In the following proposition, whose proof is provided in an appendix, we study the identification of this model and state sufficient conditions that ensure it.

**Proposition 2.1** The linear state-space model (2.5a)-(2.5b) is well identified with respect to the natural invariance properties spelled out in the Appendix A.1 if one of the elements of the vector \( \beta = (\beta_1, \beta_2, \beta_3)^\top \) that define the matrix \( B \) in (2.4) is set equal to a constant or, alternatively, when one of the unconditional variances \( \sigma^2_{v,1}, \sigma^2_{v,2}, \sigma^2_{v,3} \) that define \( R \) in (2.5b) is set equal to a positive constant.
The nonlinear state-space model. The dynamic specification (2.5b) does not introduce any dependence between the components \((\varepsilon_{1,t}, \varepsilon_{2,t}, \varepsilon_{3,t})\) of the GST \(\varepsilon_t\), even though this feature is empirically observed. This has motivated the introduction in [Durd 14, Pere 15] of a VAR-type prescription for the dynamics of \(\varepsilon_t\) that allows for correlation between its components while preserving the linearity of the corresponding state-space model. Since in this work we focus on volatility forecasting, we will handle separately in a second step.

The nonlinear state-space model incorporates the modeling of the conditional variances and then the conditional covariances are made advisable the adoption of an intermediate two-steps solution in which the nonlinear state-space dynamics of both the GST and the residuals. In unreported numerical experiments, we observed that this approach obliges us to formulate a nonlinear state-space model which makes us use an extended Kalman filter (EKF) instead of a standard one.

There are many different approaches that can be taken in order to implement this strategy. The most comprehensive one consists of building into the state-space model the entire conditional covariance dynamics of both the GST and the residuals. In unreported numerical experiments, we observed that the complexity of the resulting specification makes its estimation difficult to implement. This limitation makes advisable the adoption of an intermediate two-steps solution in which the nonlinear state-space model incorporates the modeling of the conditional variances and then the conditional covariances are handled separately in a second step.

Consider the following modified nonlinear state-space model:

\[
\begin{align*}
\{ \mathbf{r}_t \} &= \alpha + B\mathbf{e}_t + \mathbf{u}_t, \\
\{ \mathbf{e}_t \} &= \mathbf{T}\mathbf{e}_{t-1} + R_{t-1}(\mathbf{e}_{t-1})\mathbf{v}_{t-1},
\end{align*}
\]

where \(\mathbf{e}_t \in \mathbb{R}^5\), \(\alpha \in \mathbb{R}^3\), the matrix \(B \in \mathbb{M}_{5,3}\) is provided in (2.4), \(\sigma_{u,1}, \sigma_{u,2}, \sigma_{u,3} \in \mathbb{R}^+\), and the matrices \(R \in \mathbb{M}_{3,5}\) and \(T \in \mathbb{M}_{5}\) are defined as

\[
R_{t-1}(\mathbf{e}_{t-1}) := \begin{pmatrix}
\sigma_{1,t}(\mathbf{e}_{t-1}) & 0 & 0 \\
0 & \sigma_{2,t}(\mathbf{e}_{t-1}) & 0 \\
0 & 0 & \sigma_{3,t}(\mathbf{e}_{t-1}) \\
0 & 0 & 0 \\
0 & 0 & 0
\end{pmatrix}, \quad T := \begin{pmatrix}
0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0
\end{pmatrix}.
\]

Additionally, we set

\[
Q_t := R_{t-1}(\mathbf{e}_{t-1})R_{t-1}(\mathbf{e}_{t-1})^\top.
\]

The choice of matrix forms of the type (2.8)-(2.9) has two important consequences: first, the resulting model falls in the framework of the EKF, which can be hence put to work to estimate the GST. Second, as it can be seen using the EKF iterations that we spell out later on in the text, the model provides a dynamic description of the conditional variances that we are interested in forecasting, but it is static as far as the covariances is concerned. These will be modeled in a second step.

Later on in the paper we describe two specific functional dependences for the specifications \(\sigma_{i,t}(\mathbf{e}_{t-1})\), \(i \in \{1, 2, 3\}\) in (2.8) that we work with and that are consistent with the chronology with which the components \(\varepsilon_{i,t}\) of the GST \(\varepsilon_t\) are quoted.

2.2 The linear and extended Kalman filters for state and parameter estimation

We now recall the linear (LKF) and extended (EKF) Kalman filters corresponding to the models (2.5a)-(2.5b) and (2.7a)-(2.7b), respectively. An in-depth treatment of this topic can be found in [Durb 12].

Let \(\mathbf{r} := \{\mathbf{r}_1, \ldots, \mathbf{r}_T\}\) be a sample containing \(T\) three-dimensional observed log-returns and for any \(t \leq T\) denote by \(\mathcal{F}_t\) the information set generated by the observed returns up to time \(t\), that is, \(\mathcal{F}_t = \sigma(\mathbf{r}_1, \ldots, \mathbf{r}_t)\). The Kalman recursions yield minimum variance linear unbiased estimates of the forecasted and updated (or filtered) state vectors and of their covariance matrices. We denote by
\( \epsilon_{t|t} := E[\epsilon_t | F_t] \) (respectively, \( \epsilon_{t+1} := E[\epsilon_{t+1} | F_t] \)) the **updated or filtered** (respectively, **forecasted**) state vector and by \( P_{t|t} := E[e_t e_t^\top | F_t] \) (respectively, \( P_{t+1} := E[\epsilon_{t+1} \epsilon_{t+1}^\top | F_t] \)) the corresponding covariance matrices. Additionally, let \( H_t := E[(r_t - \alpha)(r_t - \alpha)^\top | F_{t-1}] \) be the forecasted conditional covariance matrices of the returns. The elements that we just introduced can be recursively obtained out of the Kalman recursions (see [Durb 12]) once \( \epsilon_1 \) and \( P_1 \) have been provided. More specifically:

\[
\begin{align*}
    u_t &= r_t - (\alpha + B \epsilon_t), \\
    H_t &= BP_t B^\top + \Sigma_u, \\
    \epsilon_{t|t} &= \epsilon_t + K_t u_t, \quad \text{with} \quad K_t := P_t B^\top H_t^{-1}, \\
    P_{t|t} &= P_t - K_t B P_t, \\
    \epsilon_{t+1} &= T \epsilon_{t|t}, \\
    P_{t+1} &= T P_{t|t} T^\top + Q_t, \quad \text{with} \quad Q_t := R_t (\epsilon_{t|t}) R_t (\epsilon_{t|t})^\top.
\end{align*}
\]

The matrix \( K_t \in \mathbb{M}_{5,3} \) is referred to as the **Kalman gain**, the relations (2.12)-(2.13) are called the **updating step**, and (2.14)-(2.15) are the **prediction step** of the Kalman filter, respectively.

If the model parameters are known, the Kalman recursions make possible the filtering of the state vectors for a given observed sample. Otherwise, the model parameters can be estimated via quasi-Newton updating step.

Proposition 2.2, one has to make sure that the volatility specifications \( \sigma_{i,t} \) yield positive values and that the resulting process has stationary solutions. This obviously depends on the specific parametric dependence chosen to define the functions \( \sigma_{i,t}(\epsilon_{t-1}) \). We will consider two different models in the empirical work that we spell out in detail in the following paragraphs.

**Model 1 for the conditional variances in the nonlinear state-space model.** In this first model we define recursively the values \( \sigma_{i,t}(\epsilon_{t-1}), \ i \in \{1, 2, 3\} \), using a GARCH-type functional dependence adapted to the chronology of the GST components. We set:

\[
\begin{align*}
    \sigma_{1,t}^2 &= a_1 + \delta_1 \sigma_{3,t-1}^2 + \gamma_1 \epsilon_{3,t-1}^2, \\
    \sigma_{2,t}^2 &= a_2 + \delta_2 \sigma_{1,t}^2, \\
    \sigma_{3,t}^2 &= a_3 + \delta_3 \sigma_{2,t}^2.
\end{align*}
\]
the stationarity of the process, we rewrite (2.17)-(2.19) as
\[
\begin{align*}
\sigma_{1,t}^2 &= a_1 + \delta_1 \sigma_{1,t-1}^2 + \gamma_1 \varepsilon_{1,t-1}^2, \\
\sigma_{2,t}^2 &= (a_2 + \alpha_1 \delta_2) + \delta_1 \sigma_{2,t-1}^2 + \gamma_1 \varepsilon_{2,t-1}^2, \\
\sigma_{3,t}^2 &= (a_3 + \alpha_2 \delta_3 + \alpha_3 \delta_4 \delta_5) + \delta_1 \sigma_{3,t-1}^2 + \gamma_1 \varepsilon_{3,t-1}^2.
\end{align*}
\]  
(2.20)  
(2.21)  
(2.22)
If we think of these relations as those defining a VEC model (see [Boll 88]), stationarity can be ensured by imposing that the spectral radius of the matrix \( A \) given by
\[
A := \begin{pmatrix}
0 & 0 & \delta_1 + \gamma_1 \\
0 & \delta_2(\delta_1 + \gamma_1) & 0 \\
0 & 0 & \delta_2 \delta_3(\delta_1 + \gamma_1)
\end{pmatrix}
\]  
(2.23)
is smaller than one [Gour 97]. It is easy to verify that this results in the inequality
\[
\delta_2 \delta_3(\delta_1 + \gamma_1) < 1,
\]  
(2.24)
which we treat later on as a nonlinear parameter constraint imposed at the time of the model estimation.

**Model 2 for the conditional variances in the nonlinear state-space model.** Based on the same arguments that we used for Model 1, we consider another GARCH-type variant for the functions \( \sigma_{i,t}(e_{t-1}), \ i \in \{1, 2, 3\} \), that determine the nonlinear state-space model (2.7a)-(2.7b) by allowing this time the possibility of autoregressive behavior in the volatilities and in the components of the GST. We set:
\[
\begin{align*}
\sigma_{1,t}^2 &= a_1 + \delta_1 \sigma_{3,t-1}^2 + \gamma_1 \varepsilon_{3,t-1}^2 + \rho_1 \sigma_{1,t-1}^2 + \tau_1 \varepsilon_{1,t-1}^2, \\
\sigma_{2,t}^2 &= (a_2 + \alpha_1 \delta_2) + \delta_1 \sigma_{2,t-1}^2 + \gamma_1 \varepsilon_{2,t-1}^2 + \rho_2 \sigma_{2,t-1}^2 + \tau_2 \varepsilon_{2,t-1}^2, \\
\sigma_{3,t}^2 &= (a_3 + \alpha_2 \delta_3 + \alpha_3 \delta_4 \delta_5) + \delta_1 \sigma_{3,t-1}^2 + \gamma_1 \varepsilon_{3,t-1}^2 + \rho_3 \sigma_{3,t-1}^2 + \tau_3 \varepsilon_{3,t-1}^2.
\end{align*}
\]  
(2.25)  
(2.26)  
(2.27)
where, again, we ensure positivity by requiring that \( \gamma_1 \geq 0 \) and \( a_i > 0, \delta_i, \rho_i, \tau_i \geq 0, \) for all \( i \in \{1, 2, 3\} \).

In order to find sufficient stationarity conditions we proceed by rewriting Model 2 as
\[
\begin{align*}
\sigma_{1,t}^2 &= a_1 + \delta_1 \sigma_{3,t-1}^2 + \gamma_1 \varepsilon_{3,t-1}^2 + \rho_1 \sigma_{1,t-1}^2 + \tau_1 \varepsilon_{1,t-1}^2, \\
\sigma_{2,t}^2 &= (a_2 + \alpha_1 \delta_2) + \delta_1 \sigma_{2,t-1}^2 + \gamma_1 \varepsilon_{2,t-1}^2 + \rho_2 \sigma_{2,t-1}^2 + \tau_2 \varepsilon_{2,t-1}^2, \\
\sigma_{3,t}^2 &= (a_3 + \alpha_2 \delta_3 + \alpha_3 \delta_4 \delta_5) + \delta_1 \sigma_{3,t-1}^2 + \gamma_1 \varepsilon_{3,t-1}^2 + \rho_3 \sigma_{3,t-1}^2 + \tau_3 \varepsilon_{3,t-1}^2 + \tau_1 \delta_2 \delta_3 \varepsilon_{1,t-1}^2 + \rho_2 \delta_3 \sigma_{2,t-1}^2 + \tau_2 \delta_3 \varepsilon_{2,t-1}^2.
\end{align*}
\]  
(2.28)  
(2.29)  
(2.30)
As for Model 1 we ensure stationarity by requiring that the spectral radius \( \rho(A) \) of the matrix \( A \) defined by:
\[
A := \begin{pmatrix}
\rho_1 + \tau_1 & \delta_1 + \gamma_1 & 0 \\
0 & \rho_2 + \tau_2 & \delta_2(\delta_1 + \gamma_1) \\
\delta_2(\rho_1 + \tau_1) & \rho_2 + \tau_2 & \delta_2(\delta_1 + \gamma_1) + \tau_3 + \rho_3
\end{pmatrix}
\]  
(2.31)
is smaller than one. Since in this case the general expression of the eigenvalues of \( A \) is very convoluted, we take advantage of the fact that for any matrix norm \( \| \cdot \| \) the inequality \( \rho(A) \leq \| A \| \) is satisfied and hence it suffices to require that \( \| A \| < 1 \) to ensure that \( \rho(A) < 1 \). We implement this condition by using the so called maximum column and row sum norms (see [Horn 13]). In the case of the maximum column sum norm, the inequality \( \| A \| < 1 \) amounts to the following three conditions
\[
\begin{align*}
(\rho_1 + \tau_1)(1 + \delta_2(1 + \delta_3)) &< 1, \\
(\rho_2 + \tau_2)(1 + \delta_2(1 + \delta_3)) &< 1, \\
(\delta_1 + \gamma_1)(1 + \delta_2(1 + \delta_3)) + \tau_3 + \rho_3 &< 1.
\end{align*}
\]  
(2.32a)  
(2.32b)  
(2.32c)
Volatility forecasting using global stochastic financial trends extracted from non-synchronous data

while the use of the maximum row sum norm results in three other different conditions, namely,

\[
\begin{align*}
\delta_1 + \gamma_1 + \rho_1 + \tau_1 &< 1, \\
\delta_2 (\delta_1 + \gamma_1 + \rho_1 + \tau_1) + \rho_2 + \tau_2 &< 1, \\
\delta_2 \delta_3 (\delta_1 + \gamma_1 + \rho_1 + \tau_1) + \delta_3 (\rho_2 + \tau_2) + \tau_3 + \rho_3 &< 1.
\end{align*}
\] (2.33a) (2.33b) (2.33c)

Any of these two sets of inequalities may be used as parameter constraints at the time of model estimation in order to ensure stationarity. Nevertheless, imposing different parameter constraints may obviously produce different estimation results.

We conclude by stating a result, whose proof can be obtained by mimicking that of Proposition 2.1, that provides conditions that ensure the proper identification of the nonlinear state-space model.

**Proposition 2.2** The nonlinear state-space model (2.7a)-(2.7b) is well identified with respect to the natural invariance properties spelled out in the Appendix A.1 if one of the elements of the vector \( \beta = (\beta_1, \beta_2, \beta_3) \) that define the matrix \( B \) in (2.4) is set equal to a constant or, alternatively, when one of the components of the vector \( a := (a_1, a_2, a_3) \) that determine the models 1 or 2 is set equal to a positive constant.

3 Volatility forecasting using the global stochastic trend

The objective of this section is producing one-step ahead forecasts for the covariance matrices of the observed asset returns. The state-space models that we introduced in the previous section have such a forecast naturally associated. Indeed, by expression (2.11), the conditional covariances \( H_t := E[(r_t - \alpha) (r_t - \alpha)^\top \mid F_{t-1}] \) are given by \( H_t = B P_t B^\top + \Sigma_u \) and, moreover, expression (2.15) provides the conditional covariance among the different components of the GST.

However, the values of the conditional covariance provided by these formulas are not compatible with the empirically observed properties of estimates of the global trend \( \{e_t\} \) or of the innovations \( \{u_t\} \) obtained with both the linear (2.5a)-(2.5b) and the nonlinear (2.7a)-(2.7b) state-space models. These processes are systematically heteroscedastic and exhibit time-varying correlation among its different components. However, regarding the innovations, both models capture by construction only the unconditional covariance \( \Sigma_u \). As to the time-varying correlation in the GST, we analyze separately the linear and the nonlinear state-space models.

First, the linear model specification (2.5a)-(2.5b) is intrinsically homoscedastic, that is, the conditional covariances \( H_t := E[(r_t - \alpha) (r_t - \alpha)^\top \mid F_{t-1}] \) and \( P_t := E [e_t e_t^\top \mid F_{t-1}] \) of both the returns \( \{r_t\} \) and the GST associated to it are asymptotically constant in time. Indeed, the time independence of the matrices \( R \) and \( \Sigma_u \) implies that after a certain number of iterations, the model reaches a steady state solution in which \( P_t \) converges to the constant matrix \( \overline{P} \) determined by the matrix equation (see [Durb 12, page 86]):

\[
\overline{P} = T \overline{P} T^\top - T \overline{P} B^\top \overline{P}^{-1} B \overline{P} T + R R^\top, \quad \text{with} \quad \overline{P} = B \overline{P} B^\top + \Sigma_u.
\]

Second, in the nonlinear case (2.7a)-(2.7b), the matrix \( Q_t \) has a nontrivial time evolution and hence so does \( P_t \). Nevertheless, a straightforward computation shows that the dynamics prescribed by either Model 1 (2.17)-(2.19) or Model 2 (3.11)-(3.13) induces a non-trivial dynamical behavior of the conditional variances of the components of the GST but makes zero the correlation between them.

These observations entail that the description of the GST associated to the models (2.5a)-(2.5b) or (2.7a)-(2.7b) is not complete enough to be used for volatility forecasting. We solve this limitation by using appropriate multivariate heteroscedastic models on both the filtered estimates \( \{e_{i,t}\} \) of the GST \( \{e_t\} \) and on the associated residuals \( \{u_t\} \). This strategy comes down to using the state-space models and the Kalman filters that go with them as a first step that provides an estimation of the GST and the
residuals out of the observed returns; this allows us to construct a larger filtration whose elements \( F^*_t \) are the pseudo-information sets generated by both the observed returns and the filtered values \( \hat{e}_t := \epsilon_{t|t} \) of the GST up to time \( t \), that is,

\[
F^*_t := \sigma (\{ r_1, \ldots, r_t \} \cup \{ \hat{e}_1, \ldots, \hat{e}_t \}).
\]

We refer to \( \{ F^*_t \} \) as the **extended information set**. We subsequently estimate multivariate volatility models that are, first, designed to take into account the asynchronicity between the components of the GST and, second, are predictable with respect to \( \{ F^*_t \} \). More specifically, we will use adapted models that will produce predictable matrix processes \( \{ P^*_t \} \) and \( \{ \Sigma^*_u,t \} \) with respect to \( \{ F^*_t \} \) such that:

\[
\hat{e}_t|_{F^*_{t-1}} \sim \text{WN}(0, P^*_t) \quad \text{and} \quad u_t|_{F^*_{t-1}} \sim \text{WN}(0, \Sigma^*_u,t).
\]

Combining these ingredients together with the expression (2.3), we can easily produce a forecast for the covariance \( H^*_t \) of the returns based on the information set \( F^*_{t-1} \):

\[
H^*_t := E \left[ (r_t - \alpha) (r_t - \alpha)^\top | F^*_{t-1} \right] = BP^*_t B^\top + \Sigma^*_u,t.
\] (3.1)

In the following paragraphs we provide the implementation details of this forecasting scheme for the linear and the nonlinear state-space models. Its empirical performance is evaluated later on in Section 4.

### 3.1 GST-based volatility forecasting using the nonlinear state-space model

The nonlinear state-space model (2.7a)-(2.7b) prescribes a conditionally heteroscedastic behavior on the components of the GST with respect to the extended information set \( \{ F^*_t \} \). Indeed, the relation (2.7b) implies that

\[
E \left[ \hat{e}_{i,t}^2 | F^*_{t-1} \right] = \sigma_{i,t}^2, \quad E \left[ \hat{e}_{i,t} \hat{e}_{j,t} | F^*_{t-1} \right] = 0, \quad \text{for any } i, j \in \{ 1, 2, 3 \}. \tag{3.2}
\]

where the functional prescription \( \sigma_{i,t}(\hat{e}_{i-1}) \) is given by one of the models (2.17)-(2.19) or (3.11)-(3.13) under consideration. The second identity in (3.2) shows that this model neglects the conditional correlation between the components of the GST that is nevertheless empirically observed [Durd 14, Pere 15]. This leads us to refine the description by introducing a dynamic conditional correlation (DCC) model for the filtered GST values \( \{ \hat{e}_t \} \) constructed out of GST returns that have been standardized using the conditional covariances \( \{ \sigma_{i,t}(\hat{e}_{i-1}) \} \), \( i \in \{ 1, 2, 3 \} \). This strategy introduces time-varying correlation between the components of the GST while preserving the conditional variance (3.2) captured by the non-linear state-space model.

More specifically, construct the standardized return vector \( \zeta_t \in \mathbb{R}^n \) via the component-wise assignment \( \zeta_{i,t} := \hat{e}_{i,t}/\sigma_{i,t}(\hat{e}_{i-1}) \) and let \( D_t := \text{diag}(\sigma_{1,t}(\hat{e}_{1-1}), \ldots, \sigma_{n,t}(\hat{e}_{n-1})) \) denote the corresponding diagonal matrix of conditional standard deviations. We now specify the dynamics of the conditional correlation matrix \( R_t \) of the standardized returns \( \zeta_t \) as:

\[
R_t = Q_t^{-1/2} Q_t^{*-1/2}, \quad Q_t := I_3 \odot Q_t, \tag{3.3}
\]
\[
Q_t = (I_3 I_3^\top - A - B) \odot Q + A \odot (\zeta_{t-1} \zeta_{t-1}^\top) + B \odot Q_{t-1}, \tag{3.4}
\]

where \( \odot \) denotes the Hadamard (or component-wise) matrix product, the parameter matrices \( A \) and \( B \) are symmetric of order 3, \( Q \) is a positive semidefinite parameter matrix of order three, and \( I_3 \) is the column vector of three elements all equal to one. Equation (3.4) is the most general DCC prescription proposed by [Engl 02]; we call it the **Hadamard DCC model**. A simplified and much
more parsimonious version of this model is the scalar subfamily in which all the elements of $A$ are considered identical and likewise those of $B$; in that case the expression (3.4) is replaced by

$$Q_t = (1 - a - b)Q + a(\zeta_{t-1}\zeta_{t-1}^T) + bQ_{t-1}, \quad (3.5)$$

with $a, b \in \mathbb{R}^+$ such that $a + b < 1$. The matrix $Q$ is obtained following an approximate targeting procedure that consists in assuming that $Q = \mathbb{E}\left[\zeta_t\zeta_t^T\right]$ and can thus be estimated by $\hat{Q} := \sum_{i=1}^T \zeta_i\zeta_i^T / T$ prior to estimating the model parameters. Despite the fact that $Q$ is not equal to the second moment matrix of $\{\zeta_t\}$ and, as a consequence, $\hat{Q}$ is not a consistent estimator of $Q$ (see [Aiel 13]), this targeting procedure is used in almost all applications of the DCC model which, according to simulation results in [Aiel 13], does not lead to strong biases in practice. The use of this model yields a conditional covariance matrix

$$\Sigma_{\hat{e},t} := \mathbb{E}\left[\hat{e}_t\hat{e}_t^T | \mathcal{F}_{t-1}^*\right] = D_tR_tD_t^T, \quad (3.6)$$

and hence

$$P_{t}^* := \mathbb{E}\left[\hat{e}_t\hat{e}_t^T | \mathcal{F}_{t-1}^*\right] = \left(\begin{array}{cc} \Sigma_{\hat{e},t} & 0 \\ 0 & \hat{e}_{2,t-1}\hat{e}_{3,t-1} / \hat{e}_{3,t-1}^T \end{array}\right) \quad (3.7)$$

Finally, we proceed analogously by modeling the conditional covariance $\{\Sigma_{u,t}^*\}$ of the residuals $\{u_t\}$ by using another DCC model also based on standardized returns constructed using the conditional variances associated to GARCH models of the type 1 or 2. The use of the resulting conditional covariance matrices $\Sigma_{u,t}^*$ and $P_{t}^*$ in (3.7) in the expression (3.1) leads to the required forecast for the covariance of the returns

$$H_{t|t-1}^* := \mathbb{E}\left[(r_t - \alpha)(r_t - \alpha)^T | \mathcal{F}_{t-1}^*\right] = BP_{t}^*B^T + \Sigma_{u,t}^*,$$

based on the information set $\{\mathcal{F}_{t}^*\}$.

### 3.2 GST-based volatility forecasting using the linear state-space model

As we already pointed out, the linear state-space model is intrinsically homoscedastic. We will hence proceed again by estimating appropriate DCC models on the filtered estimates $\{\hat{e}_t\}$ of the GST and on the associated residuals $\{u_t\}$ but, this time, the state space model does not produce any conditional variance on the components of the GST that needs to be preserved in the correlation modeling stage.

**Models for the conditional variances of the GST and the residuals.** We will carry out this construction for the estimates $\{\hat{e}_t\}$ of the GST but keeping in mind that the same approach is applicable to the residuals. We proceed by using a DCC model constructed using a strategy that is reminiscent of the one used for the nonlinear state-space model case, that is, we will use standardized returns in the construction of the dynamical correlation model obtained out of conditional variances whose parametric prescription respects the chronology with which the different components $\{\varepsilon_{1,t}, \varepsilon_{2,t}, \varepsilon_{3,t}\}$ of the GST $\varepsilon_t$ are disclosed. An important difference with respect to the approach taken in the nonlinear state-space model is that, this time, we are not obliged to respect functional prescriptions for the conditional variances of the form $\sigma_{i,t}(\hat{e}_{i,t-1})$ that were imposed by the use of the extended Kalman filter. In particular, we can use intraday dependences that will allow us to update more frequently the information set and hence to improve the forecasts. More specifically, we use two parameter families of conditional variance dynamics that generalize the models 1 and 2 that we used in the context of the nonlinear state-space model, namely:
Model 1 for the conditional variances.

\[
\sigma_{1,t}^2 = a_1 + \delta_1 \sigma_{3,t-1}^2 + \gamma_1 \hat{\varepsilon}_{3,t-1}^2,
\]

(3.8)

\[
\sigma_{2,t}^2 = a_2 + \delta_2 \sigma_{1,t}^2 + \gamma_2 \hat{\varepsilon}_{1,t}^2,
\]

(3.9)

\[
\sigma_{3,t}^2 = a_3 + \delta_3 \sigma_{2,t}^2 + \gamma_3 \hat{\varepsilon}_{2,t}^2.
\]

(3.10)

Model 2 for the conditional variances.

\[
\sigma_{1,t}^2 = a_1 + \delta_1 \sigma_{3,t-1}^2 + \gamma_1 \hat{\varepsilon}_{3,t-1}^2 + \rho_1 \sigma_{1,t-1}^2 + \tau_1 \hat{\varepsilon}_{1,t-1}^2,
\]

(3.11)

\[
\sigma_{2,t}^2 = a_2 + \delta_2 \sigma_{1,t}^2 + \gamma_2 \hat{\varepsilon}_{1,t}^2 + \rho_2 \sigma_{2,t-1}^2 + \tau_2 \hat{\varepsilon}_{2,t-1}^2,
\]

(3.12)

\[
\sigma_{3,t}^2 = a_3 + \delta_3 \sigma_{2,t}^2 + \gamma_3 \hat{\varepsilon}_{2,t}^2 + \rho_3 \sigma_{3,t-1}^2 + \tau_3 \hat{\varepsilon}_{3,t-1}^2.
\]

(3.13)

Now, for each trading date \( t \) we single out three different intraday extended information sets \( F_{t_1}^*, F_{t_2}^*, \) and \( F_{t_3}^* \) defined by

\[
F_{t_1}^* := \sigma (\{ r_1, \ldots, r_t \} \cup \{ \hat{\varepsilon}_{1,t}, \ldots, \hat{\varepsilon}_{t-1}, \hat{\varepsilon}_{t,t} \}), \quad F_{t_2}^* := F_{t_1}^* \cup \sigma (\hat{\varepsilon}_{t_2}), \quad F_{t_3}^* := F_{t_1}^*.
\]

(3.14)

The importance of the filtrations determined by these information sets in the context of the models that just introduced lays in the fact that the conditional variances that they determine are such that \( \sigma_{1,t}^2 = \sigma_{1,t}^2 \) is \( F_{t_1}^* \)-predictable, \( \sigma_{2,t}^2 = \sigma_{2,t}^2 \) is \( F_{t_2}^* \)-predictable, and \( \sigma_{3,t}^2 = \sigma_{3,t}^2 \) is \( F_{t_3}^* \)-predictable. This observation will prompt us, at the time of carrying out forecasting using these intraday information sets, to use subfamilies of the models 1 and 2 for which the entire volatility vectors \( \sigma_{1,t}, \sigma_{2,t}, \sigma_{3,t} \) are \( F_{t_3-1}^*, F_{t_1}^*, \) or \( F_{t_2}^* \)-predictable.

Positivity and stationarity of the conditional variance models. Before we proceed with the implementation of a forecasting scheme using these models, we study the conditions that need to be imposed in their parameters in order to ensure that the conditional variances that they produce are positive and that they exhibit second order stationary solutions. A way to approach this question consists of thinking of the three dimensional time series \( \{ \varepsilon_t \} \) as the one-dimensional process \( \{ \varepsilon_t \} \) obtained by ordering the components of each element \( \varepsilon_t \) according to the intraday time at which they have been disclosed. Using this point of view, the models 1 and 2 become one-dimensional GARCH models with time varying (periodic in this case) coefficients that are usually designated with the acronym tvGARCH (see [Dahl 06, Cize 09, Roha 13], and references therein). More specifically, they can be considered as tvGARCH(1,1) and tvGARCH(3,3) models, respectively, if we rewrite them as:

\[
\sigma_{1,t}^2 = a_t + \delta_t \sigma_{1,t-1}^2 + \gamma_t \hat{\varepsilon}_{1,t-1}^2, \quad \text{and} \quad \sigma_{2,t}^2 = a_t + \delta_t \sigma_{2,t-1}^2 + \gamma_t \hat{\varepsilon}_{2,t-1}^2 + \rho_t \sigma_{2,t-3}^2 + \tau_t \hat{\varepsilon}_{3,t-3},
\]

(3.15)

with \( i \in \{ 1, 2, 3 \} \), \( a_t := a_i, \quad \delta_t := \delta_i, \quad \gamma_t := \gamma_i, \quad \rho_t := \rho_i, \quad \tau_t := \tau_i \), and where \( t_i - 1 \) and \( t_i - 3 \) are defined by using recursively the convention

\[
t_i - 1 := \begin{cases} (t_i - 1)_3 & \text{when } i = 1, \\ t_{i-1} & \text{when } i \in \{ 2, 3 \}.
\end{cases}
\]

The positivity of the conditional variances implied by these models can be obtained by using only positive coefficients in the expressions that define them. Regarding stationarity, a sufficient condition of widespread use in the tvGARCH context (see for example [Roha 13]) is that \( \delta_t + \gamma_t < 1 \) for the model 1, and that \( \delta_t + \gamma_t + \rho_t + \tau_t < 1 \) for the model 2, with \( i \in \{ 1, 2, 3 \} \). Numerical experiments show that, in our particular situation, these conditions lack sharpness and produce mediocre estimation results.

In the following proposition, whose proof is provided in Appendix A.2, we establish less restrictive stationarity solutions that take advantage of the periodicity of the GARCH coefficients. In order to
formulate them, we need to introduce the matrices $A_i$, associated to the Markov representations of the recursions in (3.15) corresponding to the second model, as well as their expectations $A_i := E[A_i]$ (see Section 2.2.2 in [Fran 10] for the details). Let $\{v_t\} \sim WN(0, I_3)$ be the innovations introduced in the definition of the state-space model (2.5b). Then:

$$A_i := \begin{pmatrix} \gamma_i v_i^2 & 0 & \tau_i v_i^2 & \delta_i v_i^2 & 0 & \rho_i v_i^2 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ \gamma_i & \tau_i & \delta_i & 0 & \rho_i \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix} \quad \text{and} \quad A_i := E[A_i] = \begin{pmatrix} \gamma_i & \tau_i & \delta_i & 0 & \rho_i \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ \gamma_i & \tau_i & \delta_i & 0 & \rho_i \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

**Proposition 3.1** Consider the GARCH models with time-varying coefficients defined by the recursions in the expression (3.15). If the innovations $\{v_t\}$ that drive them are independent, then the following conditions imply the existence of a unique periodic (with period equal to three) stationary solution:

(i) For Model 1: $(\delta_1 + \gamma_1)(\delta_2 + \gamma_2)(\delta_3 + \gamma_3) < 1.$

(ii) For Model 2: $\rho(A_3 A_2 A_1) < 1$, where $\rho(\cdot)$ denotes the spectral radius.

The stationarity condition $\rho(A_3 A_2 A_1) < 1$ for Model 2 cannot be implemented as such at the time of estimation due to the convoluted analytic expression of the spectrum of $A_3 A_2 A_1$. Indeed, a straightforward computation shows that:

$$A_3 A_2 A_1 = \begin{pmatrix} \gamma_1 \xi_3 + \tau_3 & \tau_2 \xi_3 + \gamma_2 \xi_2 \xi_3 & \delta_1 \xi_3 + \rho_3 \xi_3 & \rho_2 \xi_3 & \rho_1 \xi_3 \\ \gamma_1 \xi_2 \xi_3 + \tau_2 \xi_3 & \tau_1 \xi_2 \xi_3 & \delta_1 \xi_2 \xi_3 + \rho_3 \xi_2 \xi_3 & \rho_2 \xi_2 \xi_3 & \rho_1 \xi_2 \xi_3 \\ \gamma_1 \xi_2 \xi_3 & \tau_2 \xi_3 & \tau_1 \xi_2 \xi_3 & \delta_1 \xi_2 \xi_3 + \rho_3 \xi_2 \xi_3 & \rho_2 \xi_2 \xi_3 & \rho_1 \xi_2 \xi_3 \\ \gamma_1 \xi_2 \xi_3 + \tau_3 & \tau_2 \xi_3 & \tau_1 \xi_2 \xi_3 & \delta_1 \xi_2 \xi_3 + \rho_3 \xi_2 \xi_3 & \rho_2 \xi_2 \xi_3 & \rho_1 \xi_2 \xi_3 \\ \gamma_1 \xi_2 \xi_3 & \tau_2 \xi_3 & \tau_1 \xi_2 \xi_3 & \tau_3 \xi_3 & \xi_1 \xi_3 & \xi_1 \xi_3 \\ \gamma_1 \xi_2 \xi_3 + \tau_3 & \tau_2 \xi_3 & \tau_1 \xi_2 \xi_3 & \delta_1 \xi_2 \xi_3 + \rho_3 \xi_2 \xi_3 & \rho_2 \xi_2 \xi_3 & \rho_1 \xi_2 \xi_3 \end{pmatrix},$$

with $\xi_i := \delta_i + \gamma_i, i = \{2, 3\}$. Therefore, as we already did for (2.31), we take advantage of the fact that for any matrix norm $\| \cdot \|$ the inequality $\rho(A_3 A_2 A_1) \leq \|A_3 A_2 A_1\|$ is satisfied and hence it suffices to require that $\|A_3 A_2 A_1\| < 1$ to ensure that $\rho(A_3 A_2 A_1) < 1$. We implement this condition by using, for example, the maximum row sum norm (see [Horn 13]), in which case the inequality $\|A\| < 1$ amounts to the following three conditions:

$$\begin{align*}
\delta_1 + \gamma_1 + \rho_1 + \tau_1 & < 1, \\
(\delta_2 + \gamma_2)(\delta_1 + \gamma_1 + \rho_1 + \tau_1) + \rho_2 + \tau_2 & < 1, \\
(\delta_3 + \gamma_3)(\delta_2 + \gamma_2)(\delta_1 + \gamma_1 + \rho_1 + \tau_1) + \rho_2 + \tau_2 + \tau_3 + \rho_3 & < 1.
\end{align*}$$

**Forecasting using the intraday extended information sets.** In order to forecast using the different intraday information sets, we start by estimating heteroscedastic models of the type 1 or 2 on the filtered GST $\{\tilde{\varepsilon}_t\}$ and on the corresponding residuals $\{u_i\}$, that yield predictable variance processes with respect to them. More specifically, we distinguish three cases:

(i) **Forecasting with respect to $F_{t_i}^*$:** when we carry this out, we assume that the index returns $r_{1,t}$ and $r_{2,t}$ quoted at the instants $t_1$ and $t_2$ of day $t$ have already been observed, that the corresponding GST returns $\tilde{\varepsilon}_{1,t}$ and $\tilde{\varepsilon}_{2,t}$ have been filtered, and that the corresponding residuals $u_{1,t}$ and $u_{2,t}$ are hence available. If we construct models of the type 1 or 2 for $\{\tilde{\varepsilon}_t\}$ and $\{u_t\}$ we
Volatility forecasting using global stochastic financial trends extracted from non-synchronous data

obtain \( \mathcal{F}_{t_3} \)-predictable conditional variance processes \( \{\sigma^\xi_t\} \) and \( \{\sigma^u_t\} \), that is, for any fixed day \( t \), the components of \( \sigma^\xi_t \) and \( \sigma^u_t \) are \( \mathcal{F}_{t_3-1} = \mathcal{F}_{t_2} \)-measurable and

\[
\text{var}(\tilde{\xi}_t \mid \mathcal{F}_{t_2}) = \text{diag}\left(\varepsilon_{1,t}^2, \varepsilon_{2,t}^2, \varepsilon_{3,t}^2\right) \quad \text{and} \quad \text{var}(u_t \mid \mathcal{F}_{t_2}) = \text{diag}\left(u_{1,t}^2, u_{2,t}^2, u_{3,t}^2\right).
\]

(ii) **Forecasting with respect to \( \mathcal{F}_{t_1}^s \):** in this case we use models of the type 1 or 2 for \( \{\tilde{\xi}_t\} \) and \( \{u_t\} \) but we fix the coefficient \( \gamma_3 = 0 \). These restrictions produce \( \mathcal{F}_{t_2}^s \)-predictable conditional variance processes \( \{\sigma^\xi_t\} \) and \( \{\sigma^u_t\} \) for which

\[
\text{var}(\tilde{\xi}_t \mid \mathcal{F}_{t_2}^s) = \text{diag}\left(\varepsilon_{1,t}^2, \varepsilon_{2,t}^2, \varepsilon_{3,t}^2\right) \quad \text{and} \quad \text{var}(u_t \mid \mathcal{F}_{t_2}^s) = \text{diag}\left(u_{1,t}^2, u_{2,t}^2, u_{3,t}^2\right).
\]

(iii) **Forecasting with respect to \( \mathcal{F}^{(t-1)}_{t_3} = \mathcal{F}^{(t-1)}_{t_2} \):** in this case we use models of the type 1 or 2 for \( \{\tilde{\xi}_t\} \) and \( \{u_t\} \) but we fix the coefficients \( \gamma_3 = \gamma_2 = 0 \). These restrictions produce \( \mathcal{F}^{(t-1)}_{t_2} \)-predictable conditional variance processes \( \{\sigma^\xi_t\} \) and \( \{\sigma^u_t\} \) for which

\[
\text{var}(\tilde{\xi}_t \mid \mathcal{F}^{(t-1)}_{t_2}) = \text{diag}\left(\varepsilon_{1,t}^2, \varepsilon_{2,t}^2, \varepsilon_{3,t}^2\right) \quad \text{and} \quad \text{var}(u_t \mid \mathcal{F}^{(t-1)}_{t_2}) = \text{diag}\left(u_{1,t}^2, u_{2,t}^2, u_{3,t}^2\right).
\]

We subsequently use these conditional variance processes to construct standardized vectors out of which we specify DCC models as in (3.3)-(3.4) or in (3.5). This procedure yields conditional covariance matrices \( \{\Sigma^\xi_{\mathcal{F}_2}\} \) and \( \{\Sigma^u_{\mathcal{F}_2}\} \) that have the same predictability properties as the conditional variance processes used to construct them. More specifically, they are in general \( \mathcal{F}_{t_2} \)-measurable, if \( \gamma_3 = 0 \) in the conditional variance models then they are \( \mathcal{F}_{t_2} \)-measurable, and if \( \gamma_2 = \gamma_3 = 0 \), then they are \( \mathcal{F}^{(t-1)}_{t_2} \)-measurable. Additionally, these covariance matrices are such that

\[
\tilde{\xi}_t = L^\xi_t \xi^\xi_t \quad \text{and} \quad u_t = L^u_t \xi^u_t, \quad \text{with} \quad \{\xi^\xi_t\}, \{\xi^u_t\} \sim \text{IN}(0, \mathbb{I}_3),
\]

and \( L^\xi_t \) (respectively, \( L^u_t \)) the lower triangular Cholesky factor of \( H^\xi_t \) (respectively \( H^u_t \)), that is, \( H^\xi_t = L^\xi_t L^\xi_t^\top \) (respectively, \( H^u_t = L^u_t L^u_t^\top \)). The lower triangularity of \( L^\xi_t \) and \( L^u_t \) imply that the extended intraday information sets generated by the components of \( \tilde{\xi}_t \) are identical to those spanned by the corresponding entries of \( \xi^\xi_t \). This important observation allows us to explicitly write down in the following paragraphs covariance forecasting formulas with respect to the different intraday extended information sets.

(i) **Covariance forecasting with respect to \( \mathcal{F}_{t_2} \) using models 1 or 2 with no restrictions.** In that case:

\[
E\left[\tilde{\xi}_t \tilde{\xi}_t^\top \mid \mathcal{F}_{t_2}\right] = E\left[L^\xi_t \xi^\xi_t \xi^\xi_t^\top L^\xi_t^\top \mid \mathcal{F}_{t_2}\right] = L^\xi_t E\left[\xi^\xi_t \xi^\xi_t^\top \mid \mathcal{F}_{t_2}\right] L^\xi_t^\top
\]

\[
= L^\xi_t \begin{pmatrix} \xi^2_{1,1,2} & \xi^2_{1,1,3} & 0 \\ \xi^2_{1,2,1} & \xi^2_{2,2,2} & \xi^2_{2,2,3} \\ 0 & 0 & 1 \end{pmatrix} L^\xi_t^\top =: \Sigma^\xi_{\tilde{\xi},t|t_2}. \tag{3.17}
\]

Analogously,

\[
E\left[u_t u_t^\top \mid \mathcal{F}_{t_2}\right] = L^u_t \begin{pmatrix} \xi^u_{1,1,1} & \xi^u_{1,1,2} & \xi^u_{1,1,3} \\ \xi^u_{1,2,1} & \xi^u_{2,2,2} & \xi^u_{2,2,3} \\ 0 & 0 & 1 \end{pmatrix} L^u_t^\top =: \Sigma^u_{u,t|t_2}. \tag{3.18}
\]
(ii) Covariance forecasting with respect to $\mathcal{F}_{t_1}^*$ using models 1 or 2 with the restriction $\gamma_3 = 0$. In that case:

$$E\left[\hat{\varepsilon}_t\hat{\varepsilon}_t^\top | \mathcal{F}_{t_1}^*\right] = L_t^\varepsilon\begin{pmatrix} \xi_{t,1}^2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}L_t^{\varepsilon\top} =: \Sigma_{\varepsilon, t|t_1}^*, \quad (3.19)$$

$$E\left[u_tu_t^\top | \mathcal{F}_{t_1}^*\right] = L_t^u\begin{pmatrix} \xi_{u,1}^2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}L_t^{u\top} =: \Sigma_{u, t|t_1}^*, \quad (3.20)$$

(iii) Covariance forecasting with respect to $\mathcal{F}_{(t-1)_3}^*$ using models 1 or 2 with the restrictions $\gamma_2 = \gamma_3 = 0$. In that case:

$$E\left[\hat{\varepsilon}_t\hat{\varepsilon}_t^\top | \mathcal{F}_{(t-1)_3}^*\right] = L_{t-1}^\varepsilon L_t^{\varepsilon\top} = H_t^{\varepsilon} =: \Sigma_{\varepsilon, t|(t-1)_3}^*, \quad (3.21)$$

$$E\left[u_tu_t^\top | \mathcal{F}_{(t-1)_3}^*\right] = L_{t-1}^u L_t^{u\top} = H_t^u =: \Sigma_{u, t|(t-1)_3}^*, \quad (3.22)$$

A straightforward computation shows that in all three cases, these forecasts of the covariance matrices of the processes $\{\hat{\varepsilon}_t\}$ and $\{u_t\}$ yield the following forecasts for the covariance matrices $H_{t|t_1}^*$ of the returns based on the different intraday extended information sets:

$$H_{t|t_1}^* := E[(r_t - \alpha)(r_t - \alpha)^\top | \mathcal{F}_t^*] = BP_{t|t_1}^* B^\top + \Sigma_{u, t|t_1}^*, \quad (3.23)$$

where

$$P_{t|t_1}^* := E\left[\hat{e}_t\hat{e}_t^\top | \mathcal{F}_t^*\right] = \begin{pmatrix} \Sigma_{\varepsilon, t|t_1}^* \\
0 \\
0 \end{pmatrix} \begin{pmatrix} \xi_{t,1}^2 & \xi_{2,t-1} & \xi_{2,t-1}\xi_{3,t-1} \\
0 & \xi_{2,t-1}\xi_{3,t-1} & \xi_{2,t-1}\xi_{3,t-1} \end{pmatrix}, \quad (3.24)$$

$\Sigma_{\varepsilon, t|t_1}^*$ is given by (3.17), (3.19), or (3.21), and $\Sigma_{u, t|t_1}^*$ is provided in (3.18), (3.20), and (3.22).

## 4 Empirical performance of the GST-based volatility forecasting schemes

In this section we carry out an empirical study in order to evaluate the one-day ahead volatility forecasting performances of the proposed linear and nonlinear state-space models concerning the log-returns of three major market indices with non-synchronous closing times.

### 4.1 Dataset and competing models

**Dataset.** We use as dataset the daily closing values of three major stock market indices, namely, NIKKEI 225, FTSE, and S&P500$^5$. These markets are geographically located in different time zones and have asynchronous closing times: NIKKEI 225 is an index based on the quotes of the Tokyo Stock Exchange that closes at 6:00 UTC. FTSE and S&P500 are based on the quotes of the London and the New York stock exchanges that close at 16:30 and 21:00 UTC, respectively. The closing values are adjusted for dividend payments and stock splits and the resulting data is synchronized by taking into account all the holidays of the different markets. The daily log-returns for the three indices are computed between January 5, 1996 and April 1, 2015 which yields a dataset with $T := 4581$ observations. The

$^5$Data were downloaded from the Yahoo Finance database.
whole log-returns sample is demeaned and it is divided into two parts. The first one corresponds to
the period between January 5, 1996 and April 1, 2013; it has length $T_{\text{est}} := 4095$ and it is reserved for
estimation purposes. The remaining $T_{\text{out}} := 486$ observations from April 2, 2013 to April 1, 2015 are
reserved for an out-of-sample study consisting on one-day ahead volatility forecasting.

In order to illustrate the robustness of the results obtained in this empirical study, we have included
in the technical appendix in Section A.3 a similar analysis based on the same dataset but using a shorter
estimation period (January 5, 1996 – December 4, 2006) that does not contain the volatility events in
the Fall 2008. The out-of-sample study in that case comprises the entire Great Recession (December 5,
2006 – April 1, 2015).

Competing models. We consider three groups of models whose comparative one-day ahead volatility
forecasting performance will be assessed, namely:

(i) Linear state space model combined with Models 1 and 2 (LSS Model 1/Model 2):
We use the linear state-space setup (2.5a)-(2.5b) and we estimate the components $\{\hat{\varepsilon}_t\}$ of the
GST via the Kalman recursions (2.10)-(2.15), together with the model parameters by minimizing
minus the log-likelihood function provided in (2.16) associated with the considered $T_{\text{est}}$ in-sample
observations. The model has been properly identified using Proposition 2.1 by setting $\beta_1 = 1$.
We proceed by estimating Hadamard DCC models on the filtered estimates $\{\hat{\varepsilon}_t\}$ of the GST and
on the associated linear state-space model residuals $\{u_t\}$. In order to construct the standardized
returns for the estimates $\{\hat{\varepsilon}_t\}$ that are needed in the first step of the DCC estimation, we use
either Model 1 in (3.8)-(3.10) or Model 2 in (3.11)-(3.13). These dynamical prescriptions model
the conditional variances of the GST components taking into account the particular temporal
dependance between them due to the chronology of the market closings they originate from. In
this particular empirical experiment, the conditional variances of the residuals $\{u_t\}$ are modeled
using standard individual GARCH(1,1) models even though an approach analogous to the one
followed for the GST estimates based on Model 1 or Model 2 could be adopted. Hadamard DCC
models of the type (3.4) are used at the time of modeling the conditional variances of both for the
Volatility forecasting using global stochastic financial trends extracted from non-synchronous data

The one-step ahead forecast of the volatility is constructed following the indications provided in Section 3. We now carry out volatility forecasting using the observations reserved for the out-of-sample study following the indications provided in Section 3. The one-step ahead volatility forecast is constructed in a straightforward manner, while the Hadamard prescription (3.4) presents some complications due to the presence of positivity constraints to which the model parameters are subjected and that we handle using the tools presented in [Bauw 15].

4.2 Volatility forecasting study

We now carry out volatility forecasting using the $T_{out}$ observations reserved for the out-of-sample study following the indications provided in Section 3. The one-step ahead volatility forecast is constructed in a standard way by constructing and estimating both scalar and Hadamard three dimensional models that use exclusively the daily quoted information on the closing values of the indices under consideration and ignore their non-synchronicity. The first stage of the model construction is common for both the scalar and the Hadamard setups. The deGARCHing of the returns is accomplished with the conditional deviations provided by standard one-dimensional GARCH(1,1) models that are estimated on the individual daily returns. The estimation procedure in the case of the scalar DCC (3.5) is straightforward (see for instance [Engl 09]), while the Hadamard prescription (3.4) presents some complications due to the presence of positivity constraints to which the model parameters are subjected and that we handle using the tools presented in [Bauw 15].

(ii) Nonlinear state space model combined with Models 1 (NSS Model 1): A procedure identical to the one presented in the previous point is followed but, in this case, using the nonlinear state-space model (2.7a)-(2.7b) for the GST components $\{\varepsilon_t\}$ in which the Model 1 (2.17)-(2.19) prescribes their conditional variances using a GARCH-type functional dependence adapted to their chronology. The model has been properly identified using Proposition 2.2 by setting $\beta_1 = 1$. As in the previous case, two Hadamard DCC models are then used for both the filtered $\{\varepsilon_t\}$ and the residuals $\{u_t\}$. The deGARCHing of the GST components is performed by using the conditional deviations implied by the Model 1 in (2.17)-(2.19). Again, as in the linear state-space model case, a standard GARCH(1,1) model prescription is used to determine the conditional variances of the residuals $\{u_t\}$ that is subsequently used to standardize them.

(iii) Scalar and Hadamard DCC models: These families of models are designed and widely used for volatility forecasting and we hence choose them to serve as a benchmark for the forecasting tasks that we perform in this empirical section. We proceed in a standard way by constructing and estimating both scalar and Hadamard three dimensional models that use exclusively the daily quoted information on the closing values of the indices under consideration and ignore their non-synchronicity. The first stage of the model construction is common for both the scalar and the Hadamard setups. The deGARCHing of the returns is accomplished with the conditional deviations provided by standard one-dimensional GARCH(1,1) models that are estimated on the individual daily returns. The estimation procedure in the case of the scalar DCC (3.5) is straightforward (see for instance [Engl 09]), while the Hadamard prescription (3.4) presents some complications due to the presence of positivity constraints to which the model parameters are subjected and that we handle using the tools presented in [Bauw 15].
Volatility forecasting using global stochastic financial trends extracted from non-synchronous data

(iii) GST based volatility forecasting with the linear state-space models: Since in the linear setup the extended information sets can be updated more frequently, we use this feature to improve the forecasts. As it has been discussed earlier, we distinguish three separate cases in the context of the LSS Model 1/Model 2 depending on the extended intraday information sets involved:

1. **Forecasting with respect to extended information set** $\mathcal{F}_{t_{-2}}^*$: Estimation of models 1 and 2 subjected to no parametric restrictions. The forecast $H_{t_{-2|t}}^*$ is computed using the expression (3.23) together with (3.17) and (3.18).

2. **Forecasting with respect to extended information set** $\mathcal{F}_{t_{-1}}^*$: Estimation of models 1 and 2 is subjected to the parametric restriction $\gamma_3 = 0$. The forecast $H_{t_{-1|t}}^*$ is computed using the expression (3.23) together with (3.19) and (3.20).

3. **Forecasting with respect to extended information set** $\mathcal{F}_{t_{-(t-1)}}^* = \mathcal{F}_{t-1}^*$: Estimation of models 1 and 2 is subjected to the parametric restrictions $\gamma_2 = \gamma_3 = 0$. The forecast $H_{t_{-(t-1)|t}}^*$ is computed using the expression (3.23) together with (3.21) and (3.22).

### 4.3 Model confidence sets based on covariance and KLIC loss functions

The different models are compared using the model confidence set (MCS) approach introduced in [Hans 03, Hans 11] with loss functions that involve the daily log-returns of the three indices under consideration and the forecasts of the conditional covariance matrices associated to each of the models under consideration.

**Covariance loss functions.** We will use three different covariance loss functions in the implementation of the MCS approach depending on the specific intraday extended information set used at the time of forecasting, namely:

$$d_{\mathcal{F}_{t_{-2}}^*}^{cov} := (r_{3,3}^2 - h_{33,t}^2)^2,$$

$$d_{\mathcal{F}_{t_{-1}}^*}^{cov} := \frac{1}{3} \sum_{i \leq j \leq 2, 3} (r_{i,t}r_{j,t} - h_{ij,t})^2,$$

$$d_{\mathcal{F}_{t_{-(t-1)}}^*}^{cov} := \frac{1}{6} \sum_{i \leq j \leq 1, 2, 3} (r_{i,t}r_{j,t} - h_{ij,t})^2,$$

where $t \in \{T_{est} + 1, \ldots, T_{est} + T_{out}\}$ and $h_{ij,t}$ are the ($i, j$)-entries of the corresponding model dependent forecasts for the conditional covariance matrices at $t$. More specifically, when using the scalar/Hadamard DCC model, we will consider the conditional covariance matrix $H_t$ at $t$ with respect to the information set $\mathcal{F}_{t-1}$. In the nonlinear state-space model case, we will consider $H_{t|t-1}^*$ associated to $\mathcal{F}_{t-1}^*$ and determined by (3.1). Finally, when dealing with the linear state-space model case, we will use the forecasts $H_{t|t}^*$ determined by (3.23) and based on the different intraday extended information sets.

The MCS approach identifies, from a set of competing models, the subset of models that are equivalent in terms of out-of-sample conditional covariance predictive ability and which outperform all the other models at a considered significance level $\alpha$ for the so called equivalence test. We set this significance level at 10% and 25%, and use 100 000 block bootstrap replicates with block length two in order to obtain the distribution of the relevant test statistic under the null of the equal predictive ability.

Tables 4.1 and 4.2 contain the MCS results associated to the values of the covariance loss functions obtained in 36 different out-of-sample time intervals of the form $\{T_{est} + 1, \ldots, T_{est} + 136 + 10k\}$ with $k = \{0, 1, \ldots, 35\}$. The first 136 elements in the out-of-sample period are included in all these intervals in order to ensure that there are enough values available for the bootstrapping process that is necessary in the estimation of the distribution of the model equivalence test statistic. The date corresponding to the end of this offset interval is October 18, 2013.
Volatility forecasting using global stochastic financial trends extracted from non-synchronous data

DCC models

<table>
<thead>
<tr>
<th>DCC models</th>
<th>Linear state-space models</th>
<th>Nonlinear state-space model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scalar</td>
<td>Hadamard</td>
<td>Model 1 (γ₃ = 0)</td>
</tr>
<tr>
<td>MCS dₓₜ</td>
<td>44</td>
<td>36</td>
</tr>
<tr>
<td>Sum p-values</td>
<td>1.0384</td>
<td>1.8603</td>
</tr>
<tr>
<td>MCS dₓₜ₋₁</td>
<td>44</td>
<td>34</td>
</tr>
<tr>
<td>Sum p-values</td>
<td>1.4948</td>
<td>29.5786</td>
</tr>
</tbody>
</table>

Table 4.1: Model confidence sets (MCS) constructed using the covariance based loss functions (4.1), (4.2), and (4.3), respectively, for 36 different out-of-sample lengths l(k), namely for l(k) = Tₑst + 136 + 10k, k ∈ {0, 1, ..., 35}. For each model and information set under consideration the corresponding value indicates the number of times that model has been included in the MCS at a 90% confidence level; the value underneath indicates the sum of all the MCS p-values obtained by a given model in the 36 tests. The best performing models for the considered information set are marked in bold red.

MCS dₓₜ₋₁₃₄ (t = 136) | 22                         | 22                         | 22                         | 22                         | 36                         | 2       |
| Sum p-values | 16.3389                    | 17.2042                    | 11.9633                     | 11.9633                     | 35.2726                    | 0.4470  |

Table 4.2: Model confidence sets (MCS) constructed using the covariance based loss functions (4.1), (4.2), and (4.3), respectively, for 36 different out-of-sample lengths l(k), namely for l(k) = Tₑst + 136 + 10k, k ∈ {0, 1, ..., 35}. For each model and information set under consideration the corresponding value indicates the number of times that model has been included in the MCS at a 75% confidence level; the value underneath indicates the sum of all the MCS p-values obtained by a given model in the 36 tests. The best performing models for the considered information set are marked in bold red.

These tables report, for each model, the number of times that it is included in the model confidence set with a significance level of 10% or 25%. The second figure represents the total sum of the 36 obtained MCS p-values corresponding to each model. Figures 3-5 depict the evolution of the MCS p-values when the number of the out-of-sample observations considered grows with a step equal to ten observations. These results show that:

(i) The group of Kalman-based models significantly outperforms the standard DCC models regardless the information sets involved.

(ii) The linear state-space models significantly outperform their nonlinear state-space counterparts regardless the information sets involved.

The appendix A.3 contains the results of an analogous experiment with a shorter estimation period that does not contain the high volatility events of the Fall 2008 period. In that situation the empirical study shows that: first, the conclusion (i) above holds and, second, the nonlinear state-space model based approach outperforms the linear one when the information set involved is the largest available.
Volatility forecasting using global stochastic financial trends extracted from non-synchronous data

Figure 3: Evolution of the $p$-values of the MCS test based on the covariance loss function $d_{\mathcal{F}^*_{F_t^2}}$ in (4.1) in terms of the out-of-sample length. The significance level of the MCS test is $\alpha = 0.25$.

Figure 4: Evolution of the $p$-values of the MCS test based on the covariance loss function $d_{\mathcal{F}^*_{F_t^1}}$ in (4.2) in terms of the out-of-sample length. The significance level of the MCS test is $\alpha = 0.25$. 
Volatility forecasting using global stochastic financial trends extracted from non-synchronous data

Figure 5: Evolution of the p-values of the MCS test based on the covariance loss function $d_{F^*(t-1)3}^{cov}$ in (4.3) in terms of the out-of-sample length. The significance level of the MCS test is $\alpha = 0.25$.

**KLIC loss functions.** We also implement the MCS approach using loss functions based on the Kullback-Leibler Information Criterion (KLIC) [Kull 51]. We recall that the KLIC divergence $D_{T_{\text{est}}, t}(\phi \parallel \psi)$ of a density $\psi$ that depends on the parameters $\theta$ from another density $\phi$, is defined as:

$$D_{T_{\text{est}}, t}(\phi \parallel \psi) = \frac{1}{t - T_{\text{est}}} \sum_{i = T_{\text{est}} + 1}^{t} \ln \left[ \frac{\phi_i(\mathbf{r}_i)}{\psi_i(\mathbf{r}_i; \theta)} \right], \quad t \in \{T_{\text{est}} + 1, \ldots, T_{\text{est}} + T_{\text{out}}\}$$

(4.4)

where $\phi_i(\mathbf{r}_i)$ is the real underlying conditional density associated to the data under consideration and $\psi_i(\mathbf{r}_i; \theta)$ is the one corresponding to the competing model of interest.

We use this information criterion in order to construct a loss function to evaluate the out-of-sample density forecasting abilities of the considered models, that we subsequently use in the MCS context (see [Bao 06] and [Banu 15]). Since the terms having to do with the real density $\phi(\mathbf{r})$ are common to all the models and appear as an additive constant, we then disregard the numerator in (4.4) at the time of constructing the KLIC loss functions. Additionally, in order to account for the specific intraday extended information sets used at the time of forecasting, we use again three different KLIC loss functions adapted to these different filtrations, namely:

$$d_{F_2}^{KLIC} := - \frac{1}{t - T_{\text{est}}} \sum_{i = T_{\text{est}} + 1}^{t} \ln \left[ \frac{1}{\sqrt{2\pi h_{33,i}}} \exp \left( - \frac{r_{3,i}^2}{2h_{33,i}} \right) \right],$$

(4.5)
Volatility forecasting using global stochastic financial trends extracted from non-synchronous data

\[
\begin{align*}
&d_{F_{t_i}}^{\text{KLIC}} := -\frac{1}{t - T_{\text{est}}} \sum_{i=T_{\text{est}}+1}^{t} \ln \left[ \frac{1}{2\pi} \left| \det \begin{pmatrix} h_{22,i} & h_{23,i} \\ h_{32,i} & h_{33,i} \end{pmatrix} \right| \right]^{-\frac{1}{2}} \times \\
&\times \exp \left( -\frac{1}{2} \begin{pmatrix} r_{2,i} \\ r_{3,i} \end{pmatrix}^\top \begin{pmatrix} h_{22,i} & h_{23,i} \\ h_{32,i} & h_{33,i} \end{pmatrix}^{-1} \begin{pmatrix} r_{2,i} \\ r_{3,i} \end{pmatrix} \right), \\
&d_{F_{(1-13)}}^{\text{KLIC}} := -\frac{1}{t - T_{\text{est}}} \sum_{i=T_{\text{est}}+1}^{t} \ln \left[ \frac{1}{(2\pi)^{3/2} \det(H_i)^{1/2}} \exp \left( -\frac{1}{2} r_i^\top H_i^{-1} r_i \right) \right],
\end{align*}
\]

where \( t \in \{T_{\text{est}}+1, \ldots, T_{\text{est}}+T_{\text{out}}\} \) and \( h_{i,j,t} \) are the \((i,j)\)-entries of the model dependent forecasts for the conditional covariance matrices at \( t \) explained above in the context of the covariance loss functions (4.1)-(4.3).

Tables 4.3 and 4.4 contain the MCS results at significance levels 10% and 25% associated to the values of the covariance loss functions obtained in 36 different out-of-sample time intervals of the form \( \{T_{\text{est}}+1, \ldots, T_{\text{est}}+136+10k\} \) with \( k = \{0, 1, \ldots, 35\} \).

<table>
<thead>
<tr>
<th>DCC models</th>
<th>Linear state-space models</th>
<th>Nonlinear state-space model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scalar</td>
<td>Hadamard</td>
<td>Model 1 ((\gamma_3 = 0))</td>
</tr>
<tr>
<td>MCS (d_{F_{t_i}}^{\text{KLIC}})</td>
<td>0 0</td>
<td>0 0</td>
</tr>
<tr>
<td>Sum p-values</td>
<td>0 0</td>
<td>0 0</td>
</tr>
<tr>
<td>MCS (d_{F_{(1-13)}}^{\text{KLIC}})</td>
<td>0 0</td>
<td>36 0</td>
</tr>
<tr>
<td>Sum p-values</td>
<td>0 0</td>
<td>36.0000</td>
</tr>
</tbody>
</table>

Table 4.3: Model confidence sets (MCS) constructed using the KLIC loss functions (4.5), (4.6), and (4.7), for 36 different out-of-sample length intervals \( l(k) \), namely for \( l(k) = T_{\text{est}}+136+10k \), \( k = \{0, 1, \ldots, 35\} \). For each model and information set under consideration the corresponding value indicates the number of times that model has been included in the MCS at a 90% confidence level; the value underneath indicates the sum of all the MCS p-values obtained by a given model in the 36 tests. The best performing models for the considered information set are marked in bold red.

<table>
<thead>
<tr>
<th>DCC models</th>
<th>Linear state-space models</th>
<th>Nonlinear state-space model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scalar</td>
<td>Hadamard</td>
<td>Model 1 ((\gamma_3 = 0))</td>
</tr>
<tr>
<td>MCS (d_{F_{t_i}}^{\text{KLIC}})</td>
<td>0 0</td>
<td>0 0</td>
</tr>
<tr>
<td>Sum p-values</td>
<td>0 0</td>
<td>0 0</td>
</tr>
<tr>
<td>MCS (d_{F_{(1-13)}}^{\text{KLIC}})</td>
<td>0 0</td>
<td>36 0</td>
</tr>
<tr>
<td>Sum p-values</td>
<td>0 0</td>
<td>36.0000</td>
</tr>
</tbody>
</table>

Table 4.4: Model confidence sets (MCS) constructed using the KLIC loss functions (4.5), (4.6), and (4.7), for 36 different out-of-sample length intervals \( l(k) \), namely for \( l(k) = T_{\text{est}}+136+10k \), \( k = \{0, 1, \ldots, 35\} \). For each model and information set under consideration the corresponding value indicates the number of times that model has been included in the MCS at a 75% confidence level; the value underneath indicates the sum of all the MCS p-values obtained by a given model in the 36 tests. The best performing models for the considered information set are marked in bold red.
Volatility forecasting using global stochastic financial trends extracted from non-synchronous data

Figure 6: The values of the KLIC loss function $d_{F \alpha}^{KLIC}$ in (4.5) for the competing models in terms of the out-of-sample length.

Figure 7: The values of the KLIC loss function $d_{F \alpha}^{KLIC}$ in (4.6) for the competing models in terms of the out-of-sample length.
Volatility forecasting using global stochastic financial trends extracted from non-synchronous data

Figures 6-8 depict the evolution of the loss function values (4.5)-(4.7) with 10 observations stepwise increase in the out-of-sample length. The lowest values correspond to the best performing models. The results are robust with respect to the number of the out-of-sample observations and as for the previous MCS experiment we can conclude that:

(i) The group of Kalman-based models significantly outperforms the standard DCC models regardless the information sets involved.

(ii) The linear state-space models significantly outperform their nonlinear state-space counterparts regardless the information sets involved.

The appendix A.3 contains the results of an analogous experiment with a shorter estimation period that does not contain the high volatility events of the Fall 2008 period. In that situation the empirical study shows that: first, the conclusion (i) above holds and, second, the nonlinear state-space model based approach significantly outperforms the linear one when the information set involved is the largest available.

5 Conclusions

In this work we have used linear and nonlinear state-space models that extract global stochastic financial trends out of asynchronous daily data. These models are specifically constructed to take advantage of the intraday arrival of closing information coming from different international markets located in lagged time zones in order to enhance volatility and correlation forecasting performance.

The state-space models considered incorporate nonlinearities at various levels capable of capturing the heteroscedasticity that global trends empirically exhibit. This feature is of much importance since
Volatility forecasting using global stochastic financial trends extracted from non-synchronous data

correlation forecasting is the main application developed. The identification of these models, as well as
the constraints that their parameters need to satisfy in order to exhibit stationary solutions and positive
semidefinite conditional correlation matrices, are carefully studied.

A volatility forecasting empirical study using the adjusted closing values of three major indices
(NIKKEI, FTSE, and S&P500) has been conducted using the models introduced in the theoretical
part and two different estimation periods. In this experiment, we use the model confidence set (MCS)
approach of [Hans 03, Hans 11] implemented with loss functions constructed with the conditional co-
variance matrices implied by the different models under consideration. The results show that the proposed Kalman-based forecasting scheme exhibits excellent and statistically significant performance improvements when compared to the use of standard multivariate parametric correlation models (scalar and non-scalar DCC).

References


Volatility forecasting using global stochastic financial trends extracted from non-synchronous data


Volatility forecasting using global stochastic financial trends extracted from non-synchronous data


A Appendices and supplementary material

A.1 Proof of Proposition 2.1

Let \( A \) be an element of the general linear group of order five, that is, \( A \in GL_5(\mathbb{R}) \). Consider the model prescription (2.5a)-(2.5b) and transform it according to the following prescription. First, replace the parameter \( B \) by \( BA^{-1}A \) in (2.5a) and second, apply \( A \) to both sides of (2.5b). This yields:

\[
\begin{align*}
\{ \mathbf{r}_t &= \alpha + BA^{-1}A\mathbf{e}_t + \mathbf{u}_t, \quad (A.1a) \\
A\mathbf{e}_t &= AT\mathbf{e}_{t-1} + AR\mathbf{v}_{t-1}. \quad (A.1b)
\end{align*}
\]

The model (2.5a)-(2.5b) remains invariant under this transformation if the following conditions hold:

(i) \( BA^{-1} \) has the same entries structure as \( B \).

(ii) The matrices \( A \) and \( T \) commute, that is, \( AT = TA \).

(iii) \( \bar{Q} := AQA^\top \) with \( Q := RR^\top \) is a matrix of the same entries structure as \( Q \), namely, \( \bar{Q} = \text{diag}(\bar{\sigma}_v^2, \bar{\sigma}_v^2, \bar{\sigma}_v^2, 0, 0) \), for some \( \bar{\sigma}_v^2, \bar{\sigma}_v^2, \bar{\sigma}_v^2 \in \mathbb{R}^+ \).
Indeed, under these hypotheses, the transformed equations (A.1a)-(A.1b) become:

\[
\begin{cases}
    r_t = \alpha + (BA^{-1})(Ae_t) + u_t, \\
    (Ae_t) = T(Ae_{t-1}) + ARv_{t-1}.
\end{cases}
\tag{A.2a}
\tag{A.2b}
\]

It is hence easy to see that the model (A.2a)-(A.2b) has the same structure as the original model (2.5a)-(2.5b) with the variables \( e_t \) replaced by \( (Ae_t) \), provided that \( BA^{-1} \) has the same entries structure as \( B \) and that the covariance matrix \( \Sigma(v) \) with the variables \( v \) is of the form \( \bar{Q} = \text{diag}(\bar{\sigma}^2_{v,1}, \bar{\sigma}^2_{v,2}, \bar{\sigma}^2_{v,3}, 0, 0) \), with some \( \bar{\sigma}_{v,1}, \bar{\sigma}_{v,2}, \bar{\sigma}_{v,3} \in \mathbb{R}^+ \). This covariance matrix equals

\[
\Sigma(v) := E\left[ v_t v_t^\top \right] = E\left[ ARv_t v_t^\top R^\top A^\top \right] = ARR^\top A^\top = AQA^\top
\tag{A.3}
\]

with \( Q := RR^\top \).

We first study what the implications that conditions (i)-(iii) have in the structure of \( A \in \text{GL}_5(\mathbb{R}) \). First, by point (iii) suppose that \( A \in \text{GL}_5(\mathbb{R}) \) is such that for any \( \sigma_{v,1}, \sigma_{v,2}, \sigma_{v,3} \in \mathbb{R}^+ \) there exist \( \bar{\sigma}_{v,1}, \bar{\sigma}_{v,2}, \bar{\sigma}_{v,3} \in \mathbb{R}^+ \) such that

\[
A \cdot \begin{pmatrix} \sigma_{v,1}^2 & 0 & 0 & 0 & 0 \\
0 & \sigma_{v,2}^2 & 0 & 0 & 0 \\
0 & 0 & \sigma_{v,3}^2 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 \end{pmatrix} \cdot A^\top = \begin{pmatrix} \bar{\sigma}_{v,1}^2 & 0 & 0 & 0 & 0 \\
0 & \bar{\sigma}_{v,2}^2 & 0 & 0 & 0 \\
0 & 0 & \bar{\sigma}_{v,3}^2 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 \end{pmatrix}.
\tag{A.4}
\]

Define now \( \Sigma_v := \begin{pmatrix} \sigma_{v,1}^2 & 0 & 0 \\
0 & \sigma_{v,2}^2 & 0 \\
0 & 0 & \sigma_{v,3}^2 \end{pmatrix} \), \( \Sigma_{\bar{v}} := \begin{pmatrix} \bar{\sigma}_{v,1}^2 & 0 & 0 \\
0 & \bar{\sigma}_{v,2}^2 & 0 \\
0 & 0 & \bar{\sigma}_{v,3}^2 \end{pmatrix} \), and let \( K, P \in \mathbb{M}_3, C, W \in \mathbb{M}_{3,2}, D, X \in \mathbb{M}_{2,3}, E, U \in \mathbb{M}_2 \) be such that \( A = \begin{pmatrix} K & C \\
D & E \end{pmatrix}, A^\top = \begin{pmatrix} K^\top & D^\top \\
C^\top & E^\top \end{pmatrix}, \) and \( A^{-1} = \begin{pmatrix} P & W \\
X & U \end{pmatrix} \). Condition (A.4), namely, \( AQA^\top = \bar{Q} \) is equivalent to \( A^{-1}AQA^\top = A^{-1}\bar{Q} \) or to \( QA^\top = A^{-1}\bar{Q} \) which in the notation that we just introduced amounts to

\[
\begin{pmatrix} \Sigma_v & 0 \\
0 & 0 \end{pmatrix} \begin{pmatrix} K^\top & D^\top \\
C^\top & E^\top \end{pmatrix} = \begin{pmatrix} P & W \\
X & U \end{pmatrix} \begin{pmatrix} \Sigma_{\bar{v}} & 0 \\
0 & 0 \end{pmatrix}.
\tag{A.5}
\]

Expression (A.5) is equivalent to the following three conditions:

\[
\begin{align*}
\Sigma_v K^\top &= P\Sigma_{\bar{v}}, \\
\Sigma_v D^\top &= 0, \\
X\Sigma_{\bar{v}} &= 0.
\end{align*}
\tag{A.6}
\tag{A.7}
\tag{A.8}
\]

We continue by noticing that since \( \Sigma_v \) and \( \Sigma_{\bar{v}} \) are invertible, the expressions (A.7) and (A.8) amount to \( D^\top = 0 \) and \( X = 0 \), respectively. This shows that \( A = \begin{pmatrix} K & C \\
0 & E \end{pmatrix} \) and \( A^{-1} = \begin{pmatrix} P & W \\
0 & U \end{pmatrix} \). We now impose the condition (ii), that is, \( AT = TA \):

\[
\begin{pmatrix} K \\
0 \\
M \end{pmatrix} \begin{pmatrix} C \\
0 \\
E \end{pmatrix} = \begin{pmatrix} 0 \\
M \end{pmatrix} \begin{pmatrix} K \\
0 \end{pmatrix}.
\tag{A.9}
\]
with \( M := \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \). This relation implies that
\[
CM = 0, \quad (A.10)
\]
\[
EM = MK, \quad (A.11)
\]
\[
0 = MC. \quad (A.12)
\]
The expressions (A.10) and (A.12) imply that \( C = 0 \), which yields that
\[
A = \begin{pmatrix} K & 0 \\ 0 & E \end{pmatrix}. \quad (A.13)
\]
As \( A \) is by definition invertible, in view of (A.13) so are the submatrices \( K \) and \( E \), and hence in the block structure of \( A^{-1} \) we can set
\[
W = 0, \quad P = K^{-1}, \quad U = E^{-1}, \quad (A.14)
\]
At the same time it is easy to verify that the relation (A.11) implies that
\[
K = \begin{pmatrix} k_{11} & k_{12} & k_{13} \\ 0 & 0 & E \end{pmatrix}. \quad (A.15)
\]
Let now denote by \( k_{ij}^* \) and by \( e_{ij}^* \) with \( i, j \in \{1, 2, 3\} \) the generic entries of the matrices \( K^{-1} \) and \( E^{-1} \), respectively. We may hence write by (A.15) that
\[
K^{-1} = \begin{pmatrix} k_{11} & k_{12} & k_{13} \\ 0 & 0 & E^{-1} \end{pmatrix} = \begin{pmatrix} k_{11} & k_{12} & k_{13} \\ 0 & 0 & e_{11}^* & e_{12}^* \\ 0 & 0 & e_{21}^* & e_{22}^* \end{pmatrix}. \quad (A.16)
\]
We now use the fact that condition (i) requires that the matrix \( BA^{-1} \) has the same structure as \( B \), that is, there exist some \( \bar{\beta}_1, \bar{\beta}_2, \bar{\beta}_3 \in \mathbb{R} \) such that
\[
BA^{-1} = \begin{pmatrix} \bar{\beta}_1 & 0 & 0 & \bar{\beta}_1 \\ \bar{\beta}_2 & \bar{\beta}_2 & 0 & \bar{\beta}_2 \\ \bar{\beta}_3 & \bar{\beta}_3 & \bar{\beta}_3 & 0 \end{pmatrix}. \quad (A.17)
\]
We first partition the matrix \( B \) and write it as \( B := (B_1|B_2) \), with
\[
B_1 = \begin{pmatrix} \beta_1 & 0 & 0 \\ \beta_2 & \beta_2 & 0 \\ \beta_3 & \beta_3 & \beta_3 \end{pmatrix}, \quad B_2 = \begin{pmatrix} \beta_1 & \beta_1 \\ 0 & \beta_2 \\ 0 & 0 \end{pmatrix}. \quad (A.18)
\]
We now use (A.14), (A.18) and write
\[
BA^{-1} = (B_1|B_2) \cdot \begin{pmatrix} K^{-1} & 0 \\ 0 & E^{-1} \end{pmatrix} = (B_1K^{-1}|B_2E^{-1})
\]
which by (A.17) requires both
\[
\begin{pmatrix} \beta_1 & 0 & 0 \\ \beta_2 & \beta_2 & 0 \\ \beta_3 & \beta_3 & \beta_3 \end{pmatrix} \cdot \begin{pmatrix} \frac{1}{k_{11}} & k_{12}^* & k_{13}^* \\ 0 & e_{11}^* & e_{12}^* \\ 0 & e_{21}^* & e_{22}^* \end{pmatrix} = \begin{pmatrix} \bar{\beta}_1 & 0 & 0 \\ \bar{\beta}_2 & \bar{\beta}_2 & 0 \\ \bar{\beta}_3 & \bar{\beta}_3 & \bar{\beta}_3 \end{pmatrix} \quad (A.19)
\]
We therefore compute:

\[
\begin{pmatrix}
\beta_1 & \beta_1 \\
0 & \beta_2
\end{pmatrix}
\begin{pmatrix}
e^{*}_{11} & e^{*}_{12} \\
e^{*}_{21} & e^{*}_{22}
\end{pmatrix}
= \begin{pmatrix}
\bar{\beta}_1 & \bar{\beta}_1 \\
0 & \bar{\beta}_2
\end{pmatrix}.
\] (A.20)

The relations implied by the matrix equation (A.19) for the entries (1,2), (1,3), and (2,3) yield that $k_{12}^* = 0$, $k_{13}^* = 0$, and $e_{22}^* = 0$, respectively. At the same time, the relation for the component (2,1) of the matrix equation (A.20) yields that $e_{21}^* = 0$. Consequently, both $K^{-1}$ and $E^{-1}$ are diagonal matrices. Finally, the relation (A.19) computed for the corresponding diagonal elements of $K^{-1}$ implies that that $e_{11}^* = \frac{1}{k_{11}}, e_{22}^* = \frac{1}{k_{11}}$ and we can hence write that

\[K = \lambda I_3, \quad \lambda \in \mathbb{R}.\]

Consequently, by (A.15)

\[E = \lambda I_2,\]

which automatically guarantees by (A.13) that $A = \lambda I_5$, necessarily. This implies that the matrix $B$ in the model (2.5a)-(2.5b) is defined up to multiplication by a homothety and hence the model is well identified provided that one of the elements $\beta_i$ or $\sigma_{v,i}^2$, $i \in \{1, 2, 3\}$ is set to a constant (positive in the case of $\sigma_{v,i}^2$).

A.2 Proof of Proposition 3.1

Proof of part (i). The recursion that defines the tvGARCH(1,1) model in (3.15) implies that (see for example formula (2.2) in [Ambr 08]):

\[\sigma_{t_i}^2 = \sum_{j=0}^{\infty} a_{t_i-j} \prod_{k=1}^{j} (\gamma_{t_i-k+1}v_{t_i-k}^2 + \delta_{t_i-k+1}) = \sum_{j=0}^{\infty} a_{t_i-j} \prod_{k=1}^{j} b_{t_i, t_{i-k+1}},\] (A.21)

where $b_{t_i, t_{i-k+1}} := (\gamma_{t_i-k+1}v_{t_i-k}^2 + \delta_{t_i-k+1})$. We notice that the process \( \{b_{t_i}\} \) is made of positive independent random variables. Moreover, by the Cauchy rule for series with non-negative terms, expression (A.21) converges if

\[\lambda := \lim_{j \to \infty} [b_{t_i} b_{t_i-1} \cdots b_{t_i-j+1}]^{1/j} < 1.\]

We therefore compute:

\[
\lim_{j \to \infty} [b_{t_i} b_{t_i-1} \cdots b_{t_i-j+1}]^{1/j} = \lim_{j \to \infty} \exp \left[ \frac{1}{j} \sum_{k=1}^{j} \log (b_{t_i-k+1}) \right] = \exp \lim_{j \to \infty} \left[ \frac{1}{j} \sum_{k=1}^{j} \log (b_{t_i-k+1}) \right]
\]

\[= \exp \frac{1}{3} \sum_{l=1}^{\infty} \mathbb{E} \left[ \log (\gamma_l v_{t_i}^2 + \delta_l) \right] \leq \exp \frac{1}{3} \sum_{l=1}^{\infty} \mathbb{E} \left[ (\gamma_l v_{t_i}^2 + \delta_l) \right](A.22)\]

\[= [(\delta_1 + \gamma_1)(\delta_2 + \gamma_2)(\delta_3 + \gamma_3)]^{1/3}, \]

(A.23)

where the first equality in (A.22) follows from the strong law of large numbers and the relation that follows it is a consequence of Jensen’s inequality. The inequality in the statement implies hence by (A.23) that $\lambda := \lim_{j \to \infty} [b_{t_i} b_{t_i-1} \cdots b_{t_i-j+1}]^{1/j} < 1$. A strategy mimicking, for example, the proof of Theorem 2.1 in [Fran 10], shows that in that situation model 1 has a unique stationary solution.
Volatility forecasting using global stochastic financial trends extracted from non-synchronous data

Proof of part (ii). By Theorem 2.4 in [Fran 10], it suffices to show that the top Lyapunov exponent \( \gamma \) of the sequence \( \{A_t\} \) is smaller than zero. By Theorem 2.3 in [Fran 10]:

\[
\gamma = \lim_{t_i \to \infty} \frac{1}{t_i} E \left[ \log \| A_{t_i}, A_{t_i-1} \cdots A_1 \| \right],
\]

with \( \| \cdot \| \) any matrix norm. We now use the norm \( \| A \| = \sum_{i,j} |a_{ij}| \) and notice that if all the elements of \( A \) are positive then

\[
E[\|A\|] = \|E[A]\|. \tag{A.24}
\]

Consequently,

\[
\gamma = \lim_{t_i \to \infty} \frac{1}{t_i} E \left[ \log \| A_{t_i}, A_{t_i-1} \cdots A_1 \| \right] \leq \lim_{t_i \to \infty} \frac{1}{t_i} \log \left( E[\| A_{t_i}, A_{t_i-1} \cdots A_1 \|] \right) \tag{A.25}
\]

\[
= \lim_{t_i \to \infty} \frac{1}{t_i} \log \left( E \left[ \| A_{t_i}, A_{t_i-1} \cdots A_1 \| \right] \right) = \frac{1}{3} \log \left( \rho(A_3 A_2 A_1) \right) = \log \left( \rho(A_3 A_2 A_1)^{1/3} \right). \tag{A.26}
\]

The relation in (A.25) follows from Jensen’s inequality, the first equality in (A.26) is a consequence of (A.24) and the second one of the independence of the elements in the process \( \{A_t\} \). Finally, in (A.27) we use Gelfand’s formula for the characterization of the spectral radius of a matrix. The inequality \( \gamma \leq \log \left( \rho(A_3 A_2 A_1)^{1/3} \right) \) that we just proved guarantees that the condition in the statement ensures that \( \log \left( \rho(A_3 A_2 A_1)^{1/3} \right) < 0 \) and hence \( \gamma < 0 \), as required.

In both cases, the arguments that we provided show that the unconditional variance \( E[\sigma^2_i] \) depends only on \( i \) and hence establishes the periodic stationarity claimed in the statement. \( \blacksquare \)

A.3 Empirical performance of the GST-based volatility forecasting schemes using a smaller estimation sample

In this appendix we illustrate the robustness of the results obtained in the empirical study in Section 4 by performing a similar analysis based on the same dataset but using a shorter estimation period (January 5, 1996 – December 4, 2006) that does not contain the volatility events in the Fall 2008. The out-of-sample study in that case comprises the entire Great Recession (December 5, 2006 – April 1, 2015). This choice yields a dataset with a length of \( T := 4581 \) observations and for which \( T_{\text{est}} := 2600 \) and \( T_{\text{out}} := 1981 \).

The study follows the same scheme as the one in Section 4. In particular, we consider the same competing models and the same loss functions at the time of implementing the MCS strategy. In the case of the covariance loss functions (4.1)-(4.3), the results of the corresponding MCS comparison are contained in the Tables A.5 and A.6. These results correspond to the values of the covariance loss functions obtained in 185 different out-of-sample time intervals of the form \( \{T_{\text{est}}+1, \ldots, T_{\text{est}}+141+10k\} \) with \( k = \{0, 1, \ldots, 184\} \). The first 141 elements in the out-of-sample period are included in all these intervals in order to ensure that there are enough values available for the bootstrapping process that is necessary in the estimation of the distribution of the model equivalence test statistic. The date corresponding to the end of this offset interval is July 10, 2007.

Figures 9-11 depict the evolution of the MCS \( p \)-values when the number of the out-of-sample observations considered grows with a step equal to 10 observations and the significance level is set to \( \alpha = 0.25 \). Notice that only the figures corresponding to the DCC benchmarks and to the models that exhibit the best performances for each of the intraday extended information sets are represented in the relevant figures.
It is worth to be noted that the most visible changes in the behavior of the models illustrated in Figures 9-11 takes place in the interval given by the observations 431 and 481, which corresponds to the period of time between September 25, 2008 and December 10, 2008 in which well-known market events created high levels of volatility and noticeable dynamical changes.

The results corresponding to the KLIC loss functions (4.5)-(4.7) are contained in Tables A.7 and A.8 at a significance levels 10% and 25%, respectively. These values are obtained out of 185 different out-of-sample time intervals of the form \( \{T_{est} + 1, \ldots, T_{est} + 141 + 10k\} \) with \( k \in \{0, 1, \ldots, 184\} \). As the previous MCS experiment constructed using covariance based loss functions already showed, the forecasting approaches based on the linear state-space Model 1/Model 2 and nonlinear state-space Model 1 setups significantly outperform the standard DCC models in this context. Figures 12-8 depict the evolution of the loss function values (4.5)-(4.7) using a 10 observations stepwise increase in the out-of-sample length. The lowest values correspond to the best performing models. The results that we just obtained show the robustness of the study conducted in Section 4.3 with respect to the choice of estimation period and the number of the out-of-sample observations.
Volatility forecasting using global stochastic financial trends extracted from non-synchronous data

Figure 9: Evolution of the $p$-values of the MCS test based on the covariance loss function $\sigma_{F_t^2}^{cov}$ in (4.1) in terms of the out-of-sample length. The significance level of the MCS test is $\alpha = 0.25$.

Figure 10: Evolution of the $p$-values of the MCS test based on the covariance loss function $\sigma_{F_{t_1}^1}^{cov}$ in (4.2) in terms of the out-of-sample length. The significance level of the MCS test is $\alpha = 0.25$. 
Volatility forecasting using global stochastic financial trends extracted from non-synchronous data

Figure 11: Evolution of the $p$-values of the MCS test based on the covariance loss function $d_{F^*}^{cov}$ in (4.3) in terms of the out-of-sample length. The significance level of the MCS test is $\alpha = 0.25$.

Table A.7: Model confidence sets (MCS) constructed using the KLIC loss functions (4.5), (4.6), and (4.7), for 185 different out-of-sample lengths $l(k)$, namely for $l(k) = T_{\text{est}} + 141 + 10k$, $k \in \{0,1,\ldots,184\}$. For each model and information set under consideration the corresponding value indicates the number of times that model has been included in the MCS at the 90% confidence level; the value underneath indicates the sum of all the MCS $p$-values obtained by a given model in the 185 tests. The best performing models for the considered information set are marked in bold red.
Volatility forecasting using global stochastic financial trends extracted from non-synchronous data

Table A.8: Model confidence sets (MCS) constructed using the KLIC loss functions (4.5), (4.6), and (4.7), for 185 different out-of-sample lengths \( l(k) \), namely for \( l(k) = T_{\text{out}} + 141 + 10k \), \( k \in \{0, 1, \ldots, 184\} \). For each model and information set under consideration the corresponding value indicates the number of times that model has been included in the MCS at the 75% confidence level; the value underneath indicates the sum of all the MCS \( p \)-values obtained by a given model in the 185 tests. The best performing models for the considered information set are marked in bold red.

![Figure 12: The values of the KLIC loss function \( d_{KLIC}^{F^*} \) in (4.5) for the competing models in terms of the out-of-sample length.](image)
Volatility forecasting using global stochastic financial trends extracted from non-synchronous data

Figure 13: The values of the KLIC loss function $d_{\mathcal{F}_t}^{KLIC}$ in (4.6) for the competing models in terms of the out-of-sample length.

Figure 14: The values of the KLIC loss function $d_{\mathcal{F}_{(t-1)_3}}^{KLIC}$ in (4.7) for the competing models in terms of the out-of-sample length.