Matteo Barigozzi, Marc Hallin

# Generalized dynamic factor models and volatilities estimation and forecasting

## Article (Accepted version) (Refereed)

**Original citation:**
Barigozzi, Matteo and Hallin, Marc (2017) *Generalized dynamic factor models and volatilities estimation and forecasting.* Journal of Econometrics. ISSN 0304-4076

DOI: 10.1016/j.jeconom.2017.08.010

# Generalized Dynamic Factor Models and Volatilities

## Estimation and Forecasting

MATTEO BARIGOZZI[†] AND MARC HALLIN[‡]

[†]*LSE, Department of Statistics, Houghton Street, London WC2A 2AE, UK.*
E-Mail: m.barigozzi@lse.ac.uk

[‡] *ECARES, Université libre de Bruxelles CP114/4 B-1050 Bruxelles, Belgium*
E-Mail: mhallin@ulb.ac.be

July 27, 2016

**Abstract**

In large panels of financial time series with dynamic factor structure on the levels or returns, the volatilities of the common and idiosyncratic components often exhibit strong correlations, indicating that both are exposed to the same market volatility shocks. This suggests, alongside the dynamic factor decomposition of returns, a dynamic factor decomposition of volatilities or volatility proxies. Based on this observation, Barigozzi and Hallin (2016) proposed an entirely non-parametric and model-free two-step general dynamic factor approach which accounts for a joint factor structure of returns and volatilities, and allows for extracting the market volatility shocks. Here, we go one step further, and show how the same two-step approach naturally produces volatility forecasts for the various stocks under study. In an applied exercise, we consider the panel of asset returns of the constituents of the S&P100 index over the period 2000-2009. Numerical results show that the predictors based on our two-step method outperform existing univariate and multivariate GARCH methods, as well as static factor GARCH models, in the prediction of daily high–low range—while avoiding the usual problems associated with the curse of dimensionality.

*JEL Classification*: C32, C38, C58.

*Keywords*: Volatility, Dynamic Factor Models, GARCH models.

# 1 Introduction

Decomposing asset returns and risks or volatilities into a *common*, market-driven, component and an individual, *idiosyncratic* one, is one of the main issues in financial econometrics, risk management, and portfolio optimization. Well-known theoretical results such as the Asset Pricing Theorem, indeed, show that market-driven risks cannot be diversified away, while individual ones can be eliminated

[‡]Supported by the IAP research network grant P7/06 of the Belgian government (Belgian Science Policy).

through portfolio diversification. Some of the first econometric illustrations of this are Connor and Korajczyk (1986) for returns and Engle and Marcucci (2006) for volatilities.

The very definition of a *market volatility* concept, however, calls for the analysis of a large number of individual stocks—typically, a large panel of stock volatility proxies, or a large panel of stock returns (from which volatility proxies are to be extracted)—large enough that it provides a good picture of the entire market. Such an analysis unavoidably runs into the usual challenges associated with high-dimensional observations—here, moreover, with the additional complexities of a time series context, where both auto- and cross-correlations, of all lags, play crucial roles. Since the advent of the "big data" revolution, the analysis of high-dimensional time series has attracted much interest, in conjunction with the surge of activity in the estimation of high-dimensional covariance matrices, and has become one of the most active areas of time series econometrics. A number of procedures have been proposed, of which the so-called *dynamic factor model methods*, under their various forms (exact, approximate, static, finite/infinite factor spaces, ...) so far have been the most successful.

Essentially, two distinct approaches to the analysis of large panels of volatilities can be found in the literature: *(i)* the analysis of directly observed series of volatility proxies, and *(ii)* the estimation of conditional heteroskedastic models for returns.

*(i)* When the panel under study itself is a large panel of volatility proxies (as realized volatilities or adjusted log-ranges), a factor analysis on such panels is the common way to cope with high-dimensionality issues—see Engle and Marcucci (2006), Barigozzi et al. (2014), Luciani and Veredas (2015), or Ghysels (2014), for recent contributions in that context. But the question then naturally arises of how those volatility proxies have been obtained (presumably, from some unreported primitive large panel of returns). Moreover, a direct analysis of volatility proxies only can tell one part of the story. Indeed, optimizing financial portfolios by minimizing total risk (variance) and maximizing total return, while also taking into account the existence of non-diversifiable market-driven components requires a *joint* analysis of returns *and* volatilities.

*(ii)* Multivariate conditionally heteroskedastic models do provide a unified framework for such joint analysis by defining volatilities as conditional variances of observed returns. Among those models are the multivariate stochastic volatility models by Harvey et al. (1994), the GARCH-DCC model by Engle (2002), and the composite likelihood GARCH models by Engle et al. (2008),[1] to quote only a few. However, being parametric, those models all suffer of the "curse of dimensionality": estimation, even panels of moderate size, rapidly becomes unfeasible. In order to overcome this problem, and in agreement with the Capital Asset Pricing Model (CAPM) idea of a market shock affecting all components of a financial index, factor structures on the returns have been developed jointly with GARCH modelling for the latent factors: see, for instance, Ng et al. (1992), Harvey et al. (1992), Diebold and Nerlove (1989), Van der Weide (2002), Connor et al. (2006), Sentana et al. (2008), or Rangel and Engle (2012). All those factor models, however, are *static*, and mainly of the *exact* type (strictly no idiosyncratic cross-correlations); thus, they do not fully exploit the time series nature of the data. They cannot account for idiosyncratic cross-sectional dependencies, which typically do exist in large datasets;[2] a fortiori, they cannot take into account the idiosyncratic contribution to the total volatility.

In both cases, the relation between returns and market volatility remains (fully or partially) unexplored, hence unexploited. In factor models for volatilities (approach *(i)*), common factors are

[1] We refer to the surveys by Bauwens et al. (2006), Asai et al. (2006), and Silvennoinen and Teräsvirta (2009) for comprehensive reviews of the subject.

[2] Recently, Fan et al. (2013, 2015) improved on this specific aspect by allowing for sparsity in the idiosyncratic covariance matrix of returns; but then, they just discard the idiosyncratic contribution to total volatility.

interpreted as driving “market volatility” but nothing can be said about their relation to returns, as returns are not included in the analysis. On the other hand, in conditionally heteroskedastic factor models for returns (approach *(ii)*), volatility factors are typically identified as the conditional standard errors of the return-common factors—a gross oversimplification, as factor models for returns do not carry any information on a possible factor structure for volatilities (see Barigozzi and Hallin, 2016, for details and empirical confirmation).

A global point of view, with a joint analysis of returns and volatilities in a high-dimensional setting, is therefore highly desirable. Barigozzi and Hallin (2016) propose such an analysis, with a two-step dynamic factor approach of the problem based on the *general* or *generalized dynamic factor model* introduced in Forni et al. (2000): a first dynamic factor model procedure, applied to the panel of returns, is extracting a (double) panel of volatility proxies which, in a second step, is analyzed via a second dynamic factor model procedure. Barigozzi and Hallin (2016), however, are focused on the objective of recovering volatility market shocks. Here, we go one step further, and show how the same two-step approach, possibly combined with an application of GARCH techniques, naturally produces forecasts of conditional volatilities.

Now, (conditional) volatilities are commonly defined as the (conditional) standard errors of returns (conditional on past values). This creates a tension with general dynamic factor models, which are entirely based on $L_2$ projection techniques, hence deviations from best linear predictors rather than from conditional expectations. These two points of view are usually reconciled (e.g., in ARMA-GARCH models, cfr. Francq and Zakoian, 2004) by imposing strong white noise assumptions on conditionally standardized innovations. Such assumptions are highly ad hoc and unrealistic in the high-dimensional context considered here; moreover, they are quite contrary to the spirit of general dynamic factor models. Rather than imposing such assumptions, we prefer modifying slightly the concept of (conditional) volatility, which we throughout define as square roots of (conditional) expectations of squared linear innovations (that is, squared residuals from $L_2$ projections). In the present context, that definition, moreover, naturally takes place after the decomposition of returns into common and an idiosyncratic components, yielding *two* volatilities—one for the common (market-driven) component of returns, and a second one for the idiosyncratic component. See Sections 2 and 3 for details.

In an applied exercise, we consider the panel of asset returns of the constituents of the S&P100 index from 26th January 2000 through 9th December 2009—a period comprising the recent Great Financial Crisis—and compare the forecasts produced by our two-step methods with the few feasible alternatives available in the literature: univariate and multivariate GARCH and static factor GARCH models. As a benchmark, we adopt adjusted intra-daily log range, as originally advocated by Parkinson (1980), and then also by Alizadeh et al. (2002) and Brownlees and Gallo (2010), among others. Numerical results on different time windows between 2007 and 2009 indicate that the forecasts based on our two-step methods outperform, often quite significantly, all their competitors.

The paper is organized as follows. Section 2 presents the two-step general dynamic factor procedure we are proposing. Section 3 deals with the forecasting problem, and Section 4 provides empirical results for the S&P100 panel. Finally, in Section 5, we conclude and discuss possible extensions.

# 2 The method

## 2.1 A two-step general dynamic factor model

The observation we are dealing with is an $n \times T$ panel of stock returns or levels, that is, the finite realization of a double-indexed stochastic process, of the form $\mathbf{Y} := \{Y_{it} |\ i \in \mathbb{N},\ t \in \mathbb{Z}\}$, where $t$ stands for time and $i$ for the cross-sectional index identifying the stocks. This $n \times T$ panel either can be considered a collection of $n$ observed highly interrelated time series (length $T$), or a unique observed time series in dimension $n$. As both $n$ and $T$ are "large", $(n, T)$-asymptotics, where both $n$ and $T$ tend to infinity, are considered throughout.

Let $\mathbf{Y}_n := \{\mathbf{Y}_{n,t} = (Y_{1t}, Y_{2t}, \ldots, Y_{nt})' | \, t \in \mathbb{Z}\}$ denote the $n$-dimensional subprocess of $\mathbf{Y}$, and consider the following assumptions.

ASSUMPTION (A1). *For all $n \in \mathbb{N}$, the vector process $\mathbf{Y}_n$ is strictly stationary, with mean zero and finite variances.*

ASSUMPTION (A2). *For all $n \in \mathbb{N}$, the spectral measure of $\mathbf{Y}_n$ is absolutely continuous with respect to the Lebesgue measure on $[-\pi, \pi]$, that is, $\mathbf{Y}_n$ admits a spectral density matrix $\mathbf{\Sigma}_{\mathbf{Y};n}(\theta)$, $\theta \in [-\pi, \pi]$.*

We say that $\mathbf{Y}$ admits a *dynamic factor representation* with $q$ factors if $Y_{it}$ for all $i$ and $t$ decomposes into a *"common" component* $\{X_{it}\}$, and an *"idiosyncratic" component* $\{Z_{it}\}$ such that

$$Y_{it} \quad = \quad X_{it} + Z_{it} =: \sum_{k=1}^{q} b_{ik}(L) u_{kt} + Z_{it}, \quad i \in \mathbb{N}, \quad t \in \mathbb{Z}, \tag{2.1}$$

($L$, as usual, stands for the lag operator), and

(*i*) the $q$-dimensional vector process $\mathbf{u} := \{\mathbf{u}_t = (u_{1t} u_{2t} \ldots u_{qt})' | \, t \in \mathbb{Z}\}$ is orthonormal zero-mean white noise;

(*ii*) the idiosyncratic $n$-dimensional processes $\mathbf{Z}_n := \{\mathbf{Z}_{n,t} = (Z_{1t} Z_{2t} \ldots Z_{nt})' | \, t \in \mathbb{Z}\}$ are zero-mean second-order stationary for any $n$, with $\theta$-a.e. bounded (as $n \to \infty$) dynamic eigenvalues;

(*iii*) $Z_{kt_1}$ and $u_{ht_2}$ are mutually orthogonal for any $k$, $h$, $t_1$ and $t_2$;

(*iv*) the filters $b_{ik}(L)$ are one-sided and square-summable: $\sum_{m=1}^{\infty} b_{ikm}^2 < \infty$ for all $i \in \mathbb{N}$ and $k = 1, \ldots, q$;

(*v*) $q$ is minimal with respect to (*i*)-(*iv*).

This actually defines the *general* or *generalized* dynamic factor model (GDFM), of which all other factor models (in the econometric time series literature) are particular cases; in vector notation, (2.1) also takes the form

$$\mathbf{Y}_{n,t} = \mathbf{X}_{n,t} + \mathbf{Z}_{n,t} = \mathbf{B}_n(L)\mathbf{u}_t + \mathbf{Z}_{n,t}, \quad n \in \mathbb{N}, \quad t \in \mathbb{Z}. \tag{2.2}$$

For any $\theta \in [-\pi, \pi]$, denote by $\lambda_{\mathbf{Y};n,1}(\theta), \ldots, \lambda_{\mathbf{Y};n,n}(\theta)$ the eigenvalues (in decreasing order of magnitude) of $\mathbf{\Sigma}_{\mathbf{Y};n}(\theta)$; the mappings $\theta \mapsto \lambda_{\mathbf{Y};n,i}(\theta)$ are $\mathbf{Y}_n$'s *dynamic eigenvalues*. The GDFM decomposition (2.1) can be identified by means of the following assumption.

ASSUMPTION (A3). *For some $q \in \mathbb{N}$, the $q$th dynamic eigenvalue of $\mathbf{\Sigma}_{\mathbf{Y};n}(\theta)$, $\lambda_{\mathbf{Y};n,q}(\theta)$, diverges as $n \to \infty$, $\theta$-a.e. in $[-\pi, \pi]$, while the $(q+1)$th one, $\lambda_{\mathbf{Y};n,q+1}(\theta)$, is $\theta$-a.e. bounded.*

More precisely, we know from Forni et al. (2000) and Forni and Lippi (2001) that, given Assumptions (A1) and (A2), Assumption (A3) is necessary and sufficient for the process $\mathbf{Y}$ to admit the dynamic factor representation (2.1).[3] Hallin and Lippi (2014) moreover provide very weak primitive conditions under which (2.1), hence Assumption (A3), holds for some $q < \infty$.

[3]Those references in Assumption (A1) only assume second-order stationarity, though. We are assuming strict stationarity in order to apply factor model methods to non-linear transformations of the $Y_{it}$'s.

The decomposition (2.2) of $\mathbf{Y}_n$ induces (with obvious notation) decompositions

$$\mathbf{\Gamma}_{\mathbf{Y};n,k} = \mathbf{\Gamma}_{\mathbf{X};n,k} + \mathbf{\Gamma}_{\mathbf{Z};n,k} \quad \text{and} \quad \mathbf{\Sigma}_{\mathbf{Y};n}(\theta) = \mathbf{\Sigma}_{\mathbf{X};n}(\theta) + \mathbf{\Sigma}_{\mathbf{Z};n}(\theta)$$

of $\mathbf{Y}_n$'s cross-covariance and spectral density matrices $\mathbf{\Gamma}_{\mathbf{Y};n,k} := \mathsf{E}[\mathbf{Y}_{n,t}\mathbf{Y}'_{n,t-k}]$ and $\mathbf{\Sigma}_{\mathbf{Y};n}(\theta)$, respectively.

Since $\mathbf{Y}_n$ decomposes into two components $\mathbf{X}_n$ and $\mathbf{Z}_n$ (to avoid confusion in the sequel, we call them *"level-common"* and *"level-idiosyncratic"*), where $\mathbf{X}_n$ is driven by the $q$-tuple $\{\mathbf{u}_t\}$ of *common* or *market shocks*, and $\mathbf{Z}_n$ is orthogonal to the same, two distinct sources of volatility are to be expected: the volatility originating in the shocks driving the level-common components $\mathbf{X}_n$ (volatility of level-common components), and the volatility originating in the shocks driving the level-idiosyncratic components $\mathbf{Z}_n$ (volatility of level-idiosyncratic components).

The analysis of volatilities, traditionally, is based on the autocovariance structure of some nonlinear transform of innovation processes–something the factor model decomposition (2.1) at first sight does not provide. For the common component $\mathbf{X}_n$, however, such residuals can be obtained from recent results by Forni and Lippi (2011) and Forni et al. (2015b,a). As for the idiosyncratic components $\mathbf{Z}_n$, since they are only mildly cross-correlated, componentwise residuals, without much loss of information, can be obtained via univariate AR fitting: (see Forni et al., 2005; Luciani, 2014; Luciani and Veredas, 2015).[4]

Assume, without loss of generality and for the simplicity of notation, that $n$ is an integer multiple of $(q+1)$, that is, $n = m(q+1)$ for some $m \in \mathbb{N}$. Forni and Lippi (2011) and Forni et al. (2015b) show that, under Assumptions (A1)-(A3) and the mild additional condition of a rational spectrum, there exist an $m(q+1)\times m(q+1)$ block-diagonal matrix of one-sided filters $\mathbf{A}_n(L)$ with $m$ blocks $\mathbf{A}^{(i)}(L)$ of dimension $(q+1)\times(q+1)$ such that the VAR operators $(\mathbf{I}_n - \mathbf{A}_n(L))$ are *fundamental* for $\mathbf{X}_n$, and a full-rank $n \times q$ matrix of constants $\mathbf{H}_n$ such that $\mathbf{Y}_n$ admits a VAR representation of the form

$$(\mathbf{I}_n - \mathbf{A}_n(L))\,\mathbf{Y}_{n,t} = \mathbf{H}_n\mathbf{u}_t + (\mathbf{I}_n - \mathbf{A}_n(L))\,\mathbf{Z}_{n,t} =: \mathbf{H}_n\mathbf{u}_t + \widetilde{\mathbf{Z}}_{n,t}, \quad n \in \mathbb{N}, \quad t \in \mathbb{Z}, \tag{2.3}$$

where $\widetilde{\mathbf{Z}}_n := (\mathbf{I}_n - \mathbf{A}_n(L))\,\mathbf{Z}_{n,t}$ is idiosyncratic, i.e. only has $\theta$-a.e. bounded (as $n \to \infty$) dynamic eigenvalues.

The form of the extreme-right-hand side of (2.3) is of particular importance. It shows, indeed, that the filtered panel $(\mathbf{I}_n - \mathbf{A}_n(L))\,\mathbf{Y}_{n,t}$, where the AR filters in $\mathbf{A}_n(L)$ can be estimated via $(q+1)$-dimensional AR fitting, admits a *static* factor model representation: the common shocks $\mathbf{u}_t$ in (2.3) indeed are loaded statically via the matrix loadings $\mathbf{H}_n$. Those shocks, their loadings, and the $\widetilde{\mathbf{Z}}_{n,t}$,'s therefore can be recovered from the observations by means of traditional static factor methods—as described, for instance, by Stock and Watson (2002) or Bai and Ng (2002)—applied to the filtered panel $(\mathbf{I}_n - \mathbf{A}_n(L))\,\mathbf{Y}_{n,t}$.

Denote by $\mathbf{e} := \{e_{it} := (\mathbf{H}_n\mathbf{u}_t)_i | \, i \in \mathbb{N}, \, t \in \mathbb{Z}\}$ the double-indexed process of those level-common residuals. The $n$-dimensional (but singular, being the $n$-dimensional linear transform of a $q$-dimensional white noise) subprocess $\mathbf{e}_n := \mathbf{H}_n\mathbf{u}$ of $\mathbf{e}$ is the innovation process of $\mathbf{Y}_n$'s common component $\mathbf{X}_n$, hence is zero-mean second-order white noise. Here and throughout, *linear innovation* or *innovation* is to be understood in a linear, $\mathrm{L}_2$ context: $\mathbf{e}_n$ is thus the difference between $\mathbf{X}_n$ and its projection onto its own past—which coincides with its projection onto the past of $\mathbf{Y}_n$ since $\mathbf{Z}_n$ (hence also $\widetilde{\mathbf{Z}}_n$) is orthogonal (all leads and lags) to $\mathbf{X}_n$:

$$e_{it} := X_{it} - \mathrm{Proj}^{\mathbf{X}}_{t-1}[X_{it}], \quad i \in \mathbb{N}, \quad t \in \mathbb{Z},$$

[4] Sparse VAR fitting is a feasible alternative, which we did not consider here.

where the notation $\text{Proj}_{t-1}^{\mathbf{X}}$ denotes projection onto the Hilbert space spanned up to time $(t-1)$ by the $\mathbf{X}_{nt}$'s (equivalently, the Hilbert space spanned up to time $(t-1)$ by $\mathbf{u}$ or by $\mathbf{e}_n$). The conditional expectation of $e_{it}^2$ is what we define here as the squared volatility of level-common component $X_{it}$:

$$\begin{aligned} V_{\mathbf{X};it|t-1}^2 &:= \mathsf{E}[e_{it}^2|\mathbf{X}_{n,t-1}\ldots] \\ &= \mathsf{E}[e_{it}^2|\mathbf{u}_{t-1}\ldots] = \mathsf{E}[e_{it}^2|\mathbf{e}_{n,t-1}\ldots] \quad 1 \le i \le n, \quad n \in \mathbb{N}, \quad t \in \mathbb{Z}. \end{aligned} \tag{2.4}$$

Note that (in the absence of further assumptions) $V_{\mathbf{X};it|t-1}^2$ here is not a conditional variance, as the conditional mean of $e_{it}$, unlike the unconditional one, needs not be zero. Nor is it an expectation conditional on the past of $\mathbf{Y}_n$—unless $\mathbf{X}_n$ and $\mathbf{Z}_n$ are assumed to be independent (which is the type of assumption the AR-(G)ARCH literature typically makes). Nevertheless, being the square root of the conditional expectation of the squared deviation of $X_{it}$ from its best $\text{L}_2$ predictor, $V_{\mathbf{X};it|t-1}$, in the $\text{L}_2$ context of dynamic factor models, fully qualifies as a volatility concept.

As for the $\widetilde{Z}_{it}$'s, being idiosyncratic, they are only mildly cross-correlated, and a componentwise residual analysis only overlooks negligible information. We therefore assume, for each $\{\widetilde{Z}_{it}|\, t \in \mathbb{Z}\}$, a univariate AR representation, of the form

$$\big(1 - c_i(L)\big)\widetilde{Z}_{it} = v_{it}, \quad i \in \mathbb{N}, \quad t \in \mathbb{Z}, \tag{2.5}$$

where the AR filters $c_i(L)$ are one-sided, square-summable, and such that the roots of $c(z) = 0$ all lie outside the unit disc.[5] Denote by $\mathbf{v} := \{v_{it}|\, i \in \mathbb{N}, t \in \mathbb{Z}\}$ the corresponding double-indexed process of residuals: the $v_{it}$'s are zero-mean second-order white noise, and constitute the univariate innovations of the level-idiosyncratic components $\widetilde{Z}_{it}$. The corresponding $n$-dimensional subprocess is denoted as $\mathbf{v}_n := \{\mathbf{v}_{nt} = (v_{1t}, v_{2t}, \ldots, v_{nt})'|\, t \in \mathbb{Z}\}$. By *univariate innovation* here, we mean that $\{v_{it}\}$ is the (linear) innovation of $\widetilde{Z}_{it}$ considered as a univariate process, the past of which, typically, is a strict subspace of that of $\widetilde{\mathbf{Z}}_{n,t}$

$$v_{it} := \widetilde{Z}_{it} - \text{Proj}_{t-1}^{\widetilde{Z}_i}[\widetilde{Z}_{it}] \quad i \in \mathbb{N}, \quad t \in \mathbb{Z},$$

where the notation $\text{Proj}_{t-1}^{\widetilde{Z}_i}$ denotes projection onto the Hilbert space spanned up to time $(t-1)$ by the $\widetilde{Z}_{it}$'s. The conditional (on the past until $(t-1)$ of $\widetilde{Z}_{it}$) expectation of $v_{it}^2$ is what we define here as the squared volatility of $Y_{it}$'s idiosyncratic component

$$V_{\mathbf{Z};it|t-1}^2 := \mathsf{E}[v_{it}^2|\widetilde{Z}_{i,t-1}\ldots] = \mathsf{E}[v_{it}^2|v_{i,t-1}\ldots], \quad i \in \mathbb{N}, \quad t \in \mathbb{Z}. \tag{2.6}$$

Actually, that squared volatility is an approximation to the expectation of $v_{it}^2$ conditional on the past until $(t-1)$ of the $n$-dimensional vector process $\widetilde{\mathbf{Z}}_n$.

At this point, one could think of recombining the two mutually orthogonal shocks affecting each individual return, and proceed with a volatility analysis of the $n \times T$ panel of $(e_{it} + v_{it})$'s. Merging those two sources of volatility is not good statistical practice, though, as the couples $(e_{it}, v_{it})$ clearly carry more information than the sums $(e_{it} + v_{it})$. As in Barigozzi and Hallin (2016), we therefore proceed with a joint volatility analysis of the *two* $n \times T$ panels at hand.

Classical volatility analyses are based on the autocovariance structure of volatility proxies—some non-linear transform of the residuals resulting from some second-order fit. Define, for any fixed $i \in \mathbb{N}$, the level-common and level-idiosyncratic log-volatility proxies

$$s_{it} := \log(e_{it}^2) \quad \text{and} \quad w_{it} := \log(v_{it}^2), \quad i \in \mathbb{N}, \quad t \in \mathbb{Z}. \tag{2.7}$$

[5] Sparse or low-dimensional VAR representations are also a possibility.

The advantage of logarithmic proxies over squared residuals lies in the fact that they can be analyzed via additive factor models without imposing any tricky positivity constraints (see also Engle and Marcucci, 2006, for a similar definition).

Just as the original observations, the $s_{it}$'s and $w_{it}$'s constitute double-indexed processes $\mathbf{s}$ and $\mathbf{w}$, hence, for finite $n$ and $T$, two $n \times T$ panels

$$\mathbf{s}_n := \{\mathbf{s}_{n,t} = (s_{1t}, s_{2t}, \ldots, s_{nt})' | \, t \in \mathbb{Z}\} \quad \text{and} \quad \mathbf{w}_n := \{\mathbf{w}_{n,t} = (w_{1t}, w_{2t}, \ldots, w_{nt})' | \, t \in \mathbb{Z}\}.$$

As $n$ is large, a dynamic factor model approach again naturally enters the picture.

If the two panels (2.7) and (2.7) are to be analyzed via general dynamic factor model techniques, we need the existence of spectral densities, with $q_s$ and $q_w$ exploding eigenvalues, respectively.

ASSUMPTION (B1). *The second-order moments* $\mathsf{E}[s_{it}^2]$ *and* $\mathsf{E}[w_{it}^2]$ *are finite for all* $i \in \mathbb{N}$ *and, for all* $n \in \mathbb{N}$*, the spectral densities of* $\mathbf{s}_n$ *and* $\mathbf{w}_n$ *are absolutely continuous with respect to the Lebesgue measure over* $[-\pi, \pi]$*, that is,* $\mathbf{s}_n$ *and* $\mathbf{w}_n$ *admit spectral density matrices,* $\mathbf{\Sigma}_{\mathbf{s};n}(\theta)$ *and* $\mathbf{\Sigma}_{\mathbf{w};n}(\theta)$*, respectively, for* $\theta \in [-\pi, \pi]$.

ASSUMPTION (B2).

1. *There exists* $q_s \in \mathbb{N}$ *such that the* $q_s$*th eigenvalue* $\lambda_{\mathbf{s};n,q_s}(\theta)$ *of* $\mathbf{\Sigma}_{\mathbf{s};n}(\theta)$ *diverges as* $n \to \infty$*,* $\theta$*-a.e. in* $[-\pi, \pi]$*, while the* $(q_s + 1)$*th one,* $\lambda_{\mathbf{s};n,q_s+1}(\theta)$*, is* $\theta$*-a.e. bounded.*
2. *There exists* $q_w \in \mathbb{N}$ *such that the* $q_w$*th eigenvalue* $\lambda_{\mathbf{w};n,q_w}(\theta)$ *of* $\mathbf{\Sigma}_{\mathbf{w};n}(\theta)$ *diverges as* $n \to \infty$*,* $\theta$*-a.e. in* $[-\pi, \pi]$*, while the* $(q_w + 1)$*th one,* $\lambda_{\mathbf{w};n,q_w+1}(\theta)$*, is* $\theta$*-a.e. bounded.*

As argued in Hallin and Lippi (2014), such an assumption is extremely natural and mild: why would a data-generating process with "unbounded complexity"—a weird system with increasingly many exploding dynamic eigenvalues—provide a good approximation to the finite-$(n, T)$ situation under study? Assumptions (B1) and (B2) jointly imply that each of the two panels of log-volatility proxies admit a dynamic factor representation with $q_s$ and $q_w$ common shocks, respectively. Barigozzi and Hallin (2016) show that this is empirically justified for the financial data considered in this paper, with, moreover, $q_s = q_w = 1$.

The fact that $q_s = q_w = 1$ implies a degenerate block structure which considerably simplifies the analysis described (for general $q_s$ and $q_w$) in Hallin and Liška (2011) and Barigozzi and Hallin (2016): writing $\mathring{s}_{it}$ for $s_{it} - \mathsf{E}[s_{it}]$ and $\mathring{w}_{it}$ for $w_{it} - \mathsf{E}[w_{it}]$, we have the decompositions

$$\mathring{s}_{it} = \chi_{\mathbf{s};it} + \xi_{\mathbf{s};it} = d_{\mathbf{s};i}(L)\varepsilon_t + \xi_{\mathbf{s};it}, \quad i \in \mathbb{N}, \quad t \in \mathbb{Z}, \tag{2.8}$$
$$\mathring{w}_{it} = \chi_{\mathbf{w};it} + \xi_{\mathbf{w};it} = d_{\mathbf{w};i}(L)\varepsilon_t + \xi_{\mathbf{w};it}, \quad i \in \mathbb{N}, \quad t \in \mathbb{Z}, \tag{2.9}$$

or, with obvious vector notation,

$$\mathring{\mathbf{s}}_{n,t} = \boldsymbol{\chi}_{\mathbf{s};n,t} + \boldsymbol{\xi}_{\mathbf{s};n,t} = \mathbf{D}_{\mathbf{s};n}(L)\varepsilon_t + \boldsymbol{\xi}_{\mathbf{s};n,t}, \quad n \in \mathbb{N}, \quad t \in \mathbb{Z}, \tag{2.10}$$
$$\mathring{\mathbf{w}}_{n,t} = \boldsymbol{\chi}_{\mathbf{w};n,t} + \boldsymbol{\xi}_{\mathbf{w};n,t} = \mathbf{D}_{\mathbf{w};n}(L)\varepsilon_t + \boldsymbol{\xi}_{\mathbf{w};n,t}, \quad n \in \mathbb{N}, \quad t \in \mathbb{Z}, \tag{2.11}$$

such that the same properties (*i*)-(*v*) of decomposition (2.1) hold. The $\varepsilon_t$'s here are the linear innovations of the $\boldsymbol{\chi}_{\mathbf{s};n,t}$'s and the $\boldsymbol{\chi}_{\mathbf{w};n,t}$'s.[6]

Moreover, those univariate shocks $\varepsilon_t$ naturally qualify as the *market volatility shocks*, and their impact on volatilities (estimation of impulse-response functions, etc.) is studied in detail in Barigozzi and Hallin (2016). Here instead we focus on the estimation of the four components arising from (2.10)-(2.11), which we then use for computing multi-period ahead volatility forecasts for the common and the idiosyncratic components of each individual stock return in the original panel.

[6]That is, the difference between $\boldsymbol{\chi}_{\mathbf{s};n,t}$ and $\boldsymbol{\chi}_{\mathbf{s};n,t}$ and their projections onto their respective pasts—which also coincides with their projections onto the past of $(\mathbf{s}_{n,t}, \mathbf{w}_{n,t})$.

## 2.2 Estimation

A superscript $T$ is used for estimated quantities, as opposed to population ones. While in Barigozzi and Hallin (2016) we considered estimation for arbitrary numbers of factors in returns and volatilities, here we limit ourselves to the simpler case of one factor in all panels as suggested from the empirical application in Section 4.

### 2.2.1 Step 1: estimating the level-common and level-idiosyncratic shocks

Estimation of the level-common and level-idiosyncratic innovations is in six steps.

(*i*) Start with the lag-window estimator of the spectral density matrix of the returns

$$\boldsymbol{\Sigma}^T_{\mathbf{Y};n}(\theta) := \frac{1}{2\pi}\sum_{k=-T+1}^{T-1} \mathrm{K}\left(\frac{k}{B_T}\right) e^{ik\theta}\boldsymbol{\Gamma}^T_{\mathbf{Y};n,k},$$

where $\boldsymbol{\Gamma}^T_{\mathbf{Y};n,k} := T^{-1}\sum_{t=|k|+1}^{T}\mathbf{Y}_{n,t}\mathbf{Y}'_{n,t-|k|}$ is the $k$th lag estimated autocovariance of returns and $\mathrm{K}(\cdot)$ a suitable (see Forni et al. (2015a) for details) kernel function with bandwidth $B_T$. Compute the eigenvector $\mathbf{p}^T_{\mathbf{Y};n,1}(\theta)$ corresponding to $\boldsymbol{\Sigma}^T_{\mathbf{Y};n}(\theta)$'s largest eigenvalue $\lambda^T_{\mathbf{Y};n,1}(\theta)$.

(*ii*) The estimates of the spectral density matrices of the level-common component process $\mathbf{X}_n$ and the level-idiosyncratic one $\mathbf{Z}_n$ are ($\mathbf{p}^*$ stands for the transposed complex conjugate of $\mathbf{p}$)

$$\boldsymbol{\Sigma}^T_{\mathbf{X};n}(\theta) := \lambda^T_{\mathbf{Y};n,1}(\theta)\mathbf{p}^T_{\mathbf{Y};n,1}(\theta)\mathbf{p}^{T*}_{\mathbf{Y};n,1}(\theta) \quad \text{and} \quad \boldsymbol{\Sigma}^T_{\mathbf{Z};n}(\theta) := \boldsymbol{\Sigma}^T_{\mathbf{Y};n}(\theta) - \boldsymbol{\Sigma}^T_{\mathbf{X};n}(\theta),$$

respectively.

(*iii*) By classical inverse Fourier transform of $\boldsymbol{\Sigma}^T_{\mathbf{X};n}(\theta)$, estimate the autocovariances $\boldsymbol{\Gamma}^T_{\mathbf{X};n,k}$, $k \in \mathbb{Z}$ of the level-common components.

(*iv*) Assuming, for simplicity, that $n = 2m$ for some $m \in \mathbb{N}$, consider the $m$ diagonal $2\times 2$ blocks of the $\boldsymbol{\Gamma}^T_{\mathbf{X};n,k}$'s. From each block, estimate (via standard AIC or BIC, then Yule-Walker methods) the order, and the coefficients, of a 2-dimensional VAR model. This yields, for the $i$th diagonal block, an estimator $\mathbf{A}^{(i)T}(L)$ of the autoregressive filter $\mathbf{A}^{(i)}(L)$ appearing in (2.3), hence an estimator $\mathbf{A}^T_n(L)$ of $\mathbf{A}_n(L)$: put $\widetilde{\mathbf{Y}}^T_n := \left(\mathbf{I}_n - \mathbf{A}^T_n(L)\right)\mathbf{Y}_n$ and $\boldsymbol{\Gamma}^T_{\widetilde{\mathbf{Y}};n,0} := T^{-1}\sum_{t=1}^{T}\widetilde{\mathbf{Y}}^T_{n,t}\widetilde{\mathbf{Y}}^{T\prime}_{n,t}$.

(*v*) Projecting the $\widetilde{Y}^T_i$'s onto their first largest static principal component (namely, the first principal component of $\boldsymbol{\Gamma}^T_{\widetilde{\mathbf{Y}};n,0}$) provides an estimate $\mathbf{e}^T_n$ of the level-common innovation process $\mathbf{e}_n$. Note that separate identification of $\mathbf{H}^T_n$ and $u^T$ such that $\mathbf{e}^T_n = \mathbf{H}^T_n u^T$ is not required (although possible; see Barigozzi and Hallin, 2016, for details).

(*vi*) The estimator of the idiosyncratic component $\widetilde{\mathbf{Z}}_n$ is then $\widetilde{\mathbf{Z}}^T_n := \left(\mathbf{I}_n - \mathbf{A}^T_n(L)\right)\mathbf{Y}^T_n - \mathbf{e}^T_n$. Fitting a univariate AR model (the order of which identified via standard AIC or BIC methods) to each of the $n$ components of $\widetilde{\mathbf{Z}}^T_n$, denote by $\mathbf{v}^T_n$ the resulting $n \times 1$ vector of residuals.

The results of Forni et al. (2015a) establish the consistency, as $n, T \to \infty$, of all those estimators. Note that the cross-sectional ordering of the panel has an impact on the selection of the 2-dimensional blocks in step (*iv*). Each cross-sectional permutation of the panel would lead to distinct estimators $\mathbf{e}^T_n$ and $\mathbf{v}^T_n$ sharing the same asymptotic properties. These estimators can then be aggregated into a unique one by simple averaging (after an obvious reordering of their components). Although considering all $n!$ permutations is clearly unfeasible, in practice, as stressed by Forni et al. (2015a), a few of them are enough to deliver stable averages (which therefore are matching the infeasible average over all $n!$ permutations). In Section 4, we repeat steps (*iv*)-(*v*) 100 times over randomly generated permutations.

### 2.2.2 Step 2: recovering the market volatility shocks

The estimated innovations $\mathbf{e}_n^T$ and $\mathbf{v}_n^T$ obtained in Step 1 *(v)-(vi)* are the starting point of the block-factor analysis of Step 2, itself consisting of two parts.

(*viii*) From the components of $\mathbf{e}_n^T$ and $\mathbf{v}_n^T$, compute the estimated and centered log-volatility proxies $\mathring{\mathbf{s}}_n^T$ and $\mathring{\mathbf{w}}_n^T$ as in (2.7).

(*ix*) Repeat steps *(i)-(vi)* of Section 2.2.1, on the $2n$-dimensional joint panel $\boldsymbol{\eta}_{2n}^T$ of centered log-volatility proxies $\mathring{\mathbf{s}}_n^T$ and $\mathring{\mathbf{w}}_n^T$ resulting from *(viii)*.[7] That involves a lag-window estimator

$$\boldsymbol{\Sigma}_{\boldsymbol{\eta};2n}^T(\theta) := \frac{1}{2\pi}\sum_{k=-T+1}^{T-1} \mathrm{K}\left(\frac{k}{M_T}\right) e^{ik\theta}\boldsymbol{\Gamma}_{\boldsymbol{\eta};2n,k}^T,$$

of the $2n \times 2n$ spectral density, where $\boldsymbol{\Gamma}_{\boldsymbol{\eta};2n,k}^T := T^{-1}\sum_{t=|k|+1}^T \boldsymbol{\eta}_{2n,t}^T \boldsymbol{\eta}_{2n,t-|k|}^{T\prime}$ is the $k$th lag empirical autocovariance of the $2n \times 1$ vector of log-volatility proxies and $\mathrm{K}(\cdot)$ a suitable kernel function with bandwidth $M_T$ which will depend also on the bandwidth $B_T$ of step *(i)* of Section 2.2.1. Step *(iv)* (performed on $\boldsymbol{\eta}_{2n,t}^T$) produces a $2n$-dimensional block-diagonal VAR operator (with $n$ two-dimensional diagonal blocks) of the form $(\mathbf{I}_{2n} - \mathbf{G}_{2n;\boldsymbol{\eta}}^T(L))$. Step *(vi)* eventually yields estimated common components of the log-volatility proxies[8]

$$\boldsymbol{\chi}_{\boldsymbol{\eta};2n}^T = \begin{pmatrix} \boldsymbol{\chi}_{\mathbf{s};n}^T \\ \boldsymbol{\chi}_{\mathbf{w};n}^T \end{pmatrix} := (\mathbf{I}_{2n} - \mathbf{G}_{2n;\boldsymbol{\eta}}^T(L))^{-1} \begin{pmatrix} \mathbf{H}_{\mathbf{s};n}^T \\ \mathbf{H}_{\mathbf{w};n}^T \end{pmatrix} \varepsilon^T, \qquad (2.12)$$

where $\mathbf{H}_{\mathbf{s};n}^T$ and $\mathbf{H}_{\mathbf{w};n}^T$ are $n \times 1$, and $\varepsilon^T$ is scalar. Here again, full identification of the shock $\varepsilon^T$ is not required. The estimated idiosyncratic components of the log-volatility proxies then are

$$\boldsymbol{\xi}_{\boldsymbol{\eta};2n}^T := \boldsymbol{\eta}_{2n}^T - \boldsymbol{\chi}_{\boldsymbol{\eta};2n}^T.$$

As already mentioned, the consistency, as $n$ and $T$ tend to infinity, of all estimators derived in this section is carefully established in Forni et al. (2015a), where they are computed from observed data. Here, those estimators are based on the estimated log-volatility proxies $\mathbf{s}_n^T$ and $\mathbf{w}_n^T$ obtained in Section 2.2.1. A formal consistency proof thus is needed which, with consistency rates, is the subject of ongoing research.

## 3 Forecasting

The factor decomposition (2.8)-(2.9) for the $s_{it}$'s and $w_{it}$'s yields, for the squared innovations of the level-common and level-idiosyncratic components, the multiplicative factor models

$$e_{it}^2 = \exp\left(\chi_{\mathbf{s};it} + \xi_{\mathbf{s};it} + \mathsf{E}[s_{it}]\right), \qquad i \in \mathbb{N}, \quad t \in \mathbb{Z}, \qquad (3.1)$$

$$v_{it}^2 = \exp\left(\chi_{\mathbf{w};it} + \xi_{\mathbf{w};it} + \mathsf{E}[w_{it}]\right), \qquad i \in \mathbb{N}, \quad t \in \mathbb{Z}. \qquad (3.2)$$

[7]Namely, $\eta_{it}^T$, $i = 1, \ldots, 2n$ is either $\mathring{s}_{jt}^T$ or $\mathring{w}_{jt}$ for some $j = 1, \ldots, n$.

[8]Due to block-diagonality, inverting the VAR filters only requires the inversion of two-dimensional VARs.

Those squared innovations thus each consist of a product of two components (namely, $\exp\chi_{\mathbf{s};it}$ and $\exp\xi_{\mathbf{s};it}$ for $e^2_{it}$, $\exp\chi_{\mathbf{w};it}$ and $\exp\xi_{\mathbf{w};it}$ for $v^2_{it}$) and a scale factor. Hence, we have four components containing information on volatilities, to be taken into account in the construction of volatility predictions.

To this end, we propose two approaches. The first one (Section 3.1) is entirely based on the factor models discussed in the previous sections: from these, we build linear predictors of log-volatility proxies, to which we apply the exponential transformations (3.1)-(3.2) to compute predictions of the level-common and level-idiosyncratic squared innovations $e^2_{it}$ and $v^2_{it}$, respectively, that can be interpreted as squared volatility forecasts.

In the second approach (Section 3.2), we combine our two-stage general dynamic factor model with a heuristic application of GARCH techniques:[9] after computing the transformations as in (3.1) and (3.2), we fit a GARCH model on each of the four components $\exp\chi_{\mathbf{s};it}$, $\exp\xi_{\mathbf{s};it}$, $\exp\chi_{\mathbf{w};it}$, and $\exp\xi_{\mathbf{w};it}$. Details are provided in Sections 3.1 and 3.2; both approaches, along with some competitors, are implemented in the empirical exercise of Section 4.

## 3.1 Prediction of squared volatilities (approach 1)

From representations (2.8) and (2.9), and using similar notation as in Section 2.1, we obtain the linear predictors

$$\chi_{\mathbf{s};i,t+1|t} := \text{Proj}_t^{\varepsilon}[\chi_{\mathbf{s};i,t+1}] = \sum_{k=0}^{k^*} d_{\mathbf{s};i,k+1}\varepsilon_{t-k}, \quad i\in\mathbb{N}, \quad t\in\mathbb{Z}, \tag{3.3}$$

$$\chi_{\mathbf{w};i,t+1|t} := \text{Proj}_t^{\varepsilon}[\chi_{\mathbf{w};i,t+1}] = \sum_{h=0}^{h^*} d_{\mathbf{w};i,h+1}\varepsilon_{t-h}, \quad i\in\mathbb{N}, \quad t\in\mathbb{Z}, \tag{3.4}$$

where the sums are truncated at some pre-selected lags $k^*$ and $h^*$, and the coefficients $d_{\mathbf{s};ik}$ and $d_{\mathbf{w};ik}$ are the coefficients of the impulse-response functions $d_{\mathbf{s};i}(L)$ and $d_{\mathbf{w};i}(L)$, respectively. Up to the truncation, $\chi_{\mathbf{s};i,t+1|t}$ and $\chi_{\mathbf{w};i,t+1|t}$ constitute optimal one-period ahead linear predictors, in the Hilbert spaces spanned up to time $t$ by $\log(e^2_{it})$ and $\log(v^2_{it})$, of $\chi_{\mathbf{s};i,t+1}$ and $\chi_{\mathbf{w};i,t+1}$. Moreover, when restricted to $\boldsymbol{\chi}_{\mathbf{s};t+1}$ or $\boldsymbol{\chi}_{\mathbf{w};t+1}$, the projections in (3.3)-(3.4) also coincide with the projections onto the past up to time $t$ of $(\mathbf{s}_{n,t}, \mathbf{w}_{n,t})$.

Next, with little loss, the idiosyncratic components $\xi_{\mathbf{s};it}$ and $\xi_{\mathbf{w};it}$ of the log-volatility proxies can be modeled separately as univariate AR processes (as we did in Step 1 of estimation for the level-idiosyncratic ones). This yields "univariate" linear predictors (in the sense of Section 2.1)

$$\xi_{\mathbf{s};i,t+1|t} = \sum_{\ell=0}^{\ell_i^*} \psi_{\mathbf{s};i,\ell+1}\xi_{\mathbf{s};i,t-\ell}, \quad i\in\mathbb{N}, \quad t\in\mathbb{Z}, \tag{3.5}$$

$$\xi_{\mathbf{w};i,t+1|t} = \sum_{m=0}^{m_i^*} \psi_{\mathbf{w};i,m+1}\xi_{\mathbf{w};i,t-m}, \quad i\in\mathbb{N}, \quad t\in\mathbb{Z}, \tag{3.6}$$

where the orders $\ell_i^*$ and $m_i^*$ are determined, for example, via BIC.

[9] By *heuristic* we mean that we do not impose the assumptions guaranteeing the consistency or the optimality of the method hold.

From (3.3)-(3.5) and (3.4)-(3.6), and using transformations (3.1)-(3.2), we construct predictions of the level-common and level-idiosyncratic squared innovations $e^2_{i,t+1}$ and $v^2_{i,t+1}$ as

$$e^2_{i,t+1|t} \;:=\; \exp\left(\chi_{\mathbf{s};i,t+1|t} + \xi_{\mathbf{s};i,t+1|t} + \mathsf{E}[s_{it}]\right), \quad i \in \mathbb{N}, \quad t \in \mathbb{Z}, \tag{3.7}$$

$$v^2_{i,t+1|t} \;:=\; \exp\left(\chi_{\mathbf{w};i,t+1|t} + \xi_{\mathbf{w};i,t+1|t} + \mathsf{E}[w_{it}]\right), \quad i \in \mathbb{N}, \quad t \in \mathbb{Z}. \tag{3.8}$$

In the finite-sample case, when observing a sample of length $T$ for $n$ time series, after replacing the expectations $\mathsf{E}[s_{it}]$ and $\mathsf{E}[w_{it}]$ with the corresponding sample means, and the exact coefficients $d_{\mathbf{s};ij}$, $d_{\mathbf{w};ij}$, $\psi_{\mathbf{s};ij}$ and $\psi_{\mathbf{w};ij}$ with estimated ones, the above expressions yield one-period ahead forecasts $e^{2\,T}_{i,T+1|T}$ and $v^{2\,T}_{i,T+1|T}$ of squared innovations. Those forecasts can then be recombined into a single forecast for each individual stock. Indeed, since $e_{i,t+1}$ and $v_{i,t+1}$ are mutually orthogonal, it is natural to add up the predictions:

$$\mathcal{E}^T_{\mathbf{Y};i,T+1|T} := e^{2\,T}_{i,T+1|T} + v^{2\,T}_{i,T+1|T}, \quad i = 1, \ldots, n. \tag{3.9}$$

This approach can of course be straightforwardly generalized to any multi-period ahead forecast.

From Step 2 of estimation, the estimators of the market shocks $\varepsilon_t$ and the coefficients $d$ in (3.3) and (3.4) are readily available, thus the linear predictors (3.3)-(3.4) can be immediately computed. On the other hand, estimation of the $\psi$ coefficients in (3.5)-(3.6) requires an additional step involving $n$ univariate estimations. Moreover, since log-volatilities are known to display long-memory, a possible alternative model to consider is the univariate heterogeneous AR (HAR) model by Corsi (2009) from which linear predictors can be computed in a very similar way.

Some caveats are in order, though. First, let us recall that $e^2_{it}$ and $v^2_{it}$ are just squared linear innovations—not squared deviations from conditional expectations. Hence, their expectations are not conditional variances, unless Gaussian assumptions or strong white noise assumptions on the noise driving the AR representations of $\xi_{\mathbf{s};it}$ and $\xi_{\mathbf{w};it}$ are made. Moreover, unless a further assumption is made that idiosyncratic returns are mutually strictly orthogonal, $v^2_{it}$ only takes into account the univariate past of $Z_{it}$; a similar remark holds for $\xi_{\mathbf{s};i,t+1|t}$ and $\xi_{\mathbf{w};i,t+1|t}$. Such assumptions, which are of the same nature as those underlying classical VARMA-GARCH models, are quite unlikely to hold in this context, and contradict the spirit of factor model methods. Therefore, we will refrain imposing them. Second, the optimality properties of $e^2_{i,T+1|T}$ and $v^2_{i,T+1|T}$ as predictors of $e^2_{i,T+1}$ and $v^2_{i,T+1}$, of the linear $\mathrm{L}_2$ type, hold in the space of their log-transforms, and do not resist exponentiation. For all those reasons, the forecasts proposed here should be considered somewhat heuristic. Heuristic as they may be, however, their performance quite often appears to be better than their competitors' when dealing with real data: see Section 4.

### 3.2 Prediction of squared volatilities (approach 2)

The volatility forecasts developed in Section 3.1 are based on $\mathrm{L}_2$ features in the space of the log-transforms $\log e^2_{i,t+1}$ and $\log v^2_{i,t+1}$. As an alternative, one may prefer combining the factor approach with GARCH techniques—much in the spirit of the factor GARCH models considered in the literature[10], but exploiting the more elaborate two-step dynamic factor method developed here.

For each of the four quantities appearing in (3.1)-(3.2), we can think of a conditional heteroskedastic GARCH dynamic scheme of the form

$$\exp\left(\chi_{\mathbf{s};it}\right) = \omega_{\mathbf{s};it}\nu^2_{\mathbf{s};it}, \qquad \exp\left(\xi_{\mathbf{s};it}\right) = h_{\mathbf{s};it}\epsilon^2_{\mathbf{s};it}, \qquad i \in \mathbb{N}, \quad t \in \mathbb{Z} \tag{3.10}$$

$$\exp\left(\chi_{\mathbf{w};it}\right) = \omega_{\mathbf{w};it}\nu^2_{\mathbf{w};it}, \qquad \exp\left(\xi_{\mathbf{w};it}\right) = h_{\mathbf{w};it}\epsilon^2_{\mathbf{w};it}, \qquad i \in \mathbb{N}, \quad t \in \mathbb{Z} \tag{3.11}$$

[10]See, for instance, Diebold and Nerlove (1989); Harvey et al. (1992); Sentana et al. (2008); Hafner and Preminger (2009).

with

$$\omega_{\mathbf{s};it} = \gamma_{\mathbf{s};i} + \sum_{k=1}^{p_i} \alpha_{\mathbf{s};i,k} \exp\left(\chi_{\mathbf{s};i,t-k}\right) + \sum_{\ell=1}^{q_i} \beta_{\mathbf{s};i,\ell}\omega_{\mathbf{s};i,t-\ell}, \quad i \in \mathbb{N}, \quad t \in \mathbb{Z}, \tag{3.12}$$

$$h_{\mathbf{s};it} = c_{\mathbf{s};i} + \sum_{k=1}^{p_i^*} a_{\mathbf{s};i,k} \exp\left(\xi_{\mathbf{s};i,t-k}\right) + \sum_{\ell=1}^{q_i^*} b_{\mathbf{s};i,\ell} h_{\mathbf{s};i,t-\ell}, \quad i \in \mathbb{N}, \quad t \in \mathbb{Z}, \tag{3.13}$$

$$\omega_{\mathbf{w};it} = \gamma_{\mathbf{w};i} + \sum_{k=1}^{\tilde{p}_i} \alpha_{\mathbf{w};i,k} \exp\left(\chi_{\mathbf{w};i,t-k}\right) + \sum_{\ell=1}^{\tilde{q}_i} \beta_{\mathbf{w};i,\ell}\omega_{\mathbf{w};i,t-\ell}, \quad i \in \mathbb{N}, \quad t \in \mathbb{Z}, \tag{3.14}$$

$$h_{\mathbf{w};it} = c_{\mathbf{w};i} + \sum_{k=1}^{\tilde{p}_i^*} a_{\mathbf{w};i,k} \exp\left(\xi_{\mathbf{w};i,t-k}\right) + \sum_{\ell=1}^{\tilde{q}_i^*} b_{\mathbf{w};i,\ell} h_{\mathbf{w};i,t-\ell}, \quad i \in \mathbb{N}, \quad t \in \mathbb{Z}, \tag{3.15}$$

where the orders $(p_i, q_i)$, $(p_i^*, q_i^*)$, $(\tilde{p}_i, \tilde{q}_i)$, and $(\tilde{p}_i^*, \tilde{q}_i^*)$ can be determined, for example, via BIC. The standard GARCH assumptions with the addition of independence of the common and idiosyncratic components of the volatility panel here would take the form

ASSUMPTION (C). *The processes $\nu_{\mathbf{s};i}$, $\epsilon_{\mathbf{s};i}$, $\nu_{\mathbf{w};i}$, $\epsilon_{\mathbf{w};i}$ are i.i.d. with mean zero and unit variance. Moreover, $\nu_{\mathbf{s};i}$ and $\epsilon_{\mathbf{s};j}$, $\nu_{\mathbf{w};i}$ and $\epsilon_{\mathbf{w};j}$, $\nu_{\mathbf{s};i}$ and $\nu_{\mathbf{s};j}$, $\nu_{\mathbf{w};i}$ and $\nu_{\mathbf{w};j}$, $\epsilon_{\mathbf{s};i}$ and $\epsilon_{\mathbf{s};j}$ , and $\epsilon_{\mathbf{w};i}$ and $\epsilon_{\mathbf{w};j}$, respectively, are mutually independent at all leads and lags for any $i, j \in \mathbb{N}$.*

Those assumptions, clearly, are ad hoc: independence, as a rule, is unnatural in the $L_2$ framework of factor models, which are entirely based on second-order moments. A heuristic application of the estimation techniques derived from those assumptions nevertheless yields forecasts that work quite well, and, in the empirical exercise of Section 4, outperform all existing methods.

Denote by $\mathsf{E}_{t-1}^{\mathbf{X}}$ the conditional expectation given the past until $(t-1)$ of $\mathbf{X}_n$: the conditioning space thus contains the past values of all $\nu_{\mathbf{s};i}$'s and all $\epsilon_{\mathbf{s};i}$'s. Then, under Assumption (C), (3.1), (3.10), and the GARCH specifications (3.12)-(3.13), the squared volatility of the level-common component is (see (2.4))[11], using Assumption (C) in the derivation of the last equality[12],

$$\begin{aligned}
V^2_{\mathbf{X};it|t-1} &:= \mathsf{E}_{t-1}^{\mathbf{X}}\left[e_{it}^2\right] = \mathsf{E}_{t-1}^{\mathbf{X}}\left[\exp\left(\chi_{\mathbf{s};it} + \xi_{\mathbf{s};it} + \mathsf{E}[s_{it}]\right)\right] \\
&= \mathsf{E}_{t-1}^{\mathbf{X}}\left[\omega_{\mathbf{s};it}\nu_{\mathbf{s};it}^2 h_{\mathbf{s};it}\epsilon_{\mathbf{s};it}^2\right]\exp(\mathsf{E}[s_{it}]) \\
&= \omega_{\mathbf{s};it}h_{\mathbf{s};it}\,\mathsf{E}_{t-1}^{\mathbf{X}}\left[\nu_{\mathbf{s};it}^2\epsilon_{\mathbf{s};it}^2\right]\exp(\mathsf{E}[s_{it}]) \\
&= \omega_{\mathbf{s};it}h_{\mathbf{s};it}\,\exp(\mathsf{E}[s_{it}]) \quad i \in \mathbb{N}, \quad t \in \mathbb{Z},.
\end{aligned} \tag{3.16}$$

Following the same reasoning, and under the same conditions, (3.2), (3.11), (3.14) and (3.15) yield, for the squared volatility of the level-idiosyncratic component (see (2.6)),

$$V^2_{\mathbf{Z};it|t-1} := \mathsf{E}_{t-1}^{\widetilde{Z}_i}\left[v_{it}^2\right] = \omega_{\mathbf{w};it}h_{\mathbf{w};it}\,\exp(\mathsf{E}[w_{it}]), \quad i \in \mathbb{N}, \quad t \in \mathbb{Z}, \tag{3.17}$$

where $\mathsf{E}_{t-1}^{\widetilde{Z}_i}$ stands for the conditional expectation given the past until $(t-1)$ of $\widetilde{Z}_i$. This, as explained in Section 2.1, provides an approximation of the conditional expectation given the past until $(t-1)$ of $\widetilde{\mathbf{Z}}_n$—unless of course an unrealistic exact factor structure for the returns is imposed.

[11]Note that, by construction, $\omega_{\mathbf{s};it}$ only depends on $\nu_{\mathbf{s};i,t-k}$, $k > 0$; analogously, $h_{\mathbf{s};it}$ only depends on $\epsilon_{\mathbf{s};i,t-k}$, $k > 0$.

[12]Without Assumption (C), the term $\mathsf{E}_{t-1}^{\mathbf{X}}[\nu_{\mathbf{s};it}^2\epsilon_{\mathbf{s};it}^2]$ does not disappear.

When observing a sample of length $T$ for $n$ time series, and after replacing expectations with sample means and parameters with their estimators, the right-hand sides of (3.16) and (3.17) constitute heuristic one-period ahead predictors of squared volatilities which we denote as $V^{2\,T}_{\mathbf{X};i,T+1|T}$ and $V^{2\,T}_{\mathbf{Z};i,T+1|T}$. As before, these two predictions can be recombined into a unique forecast

$$V^{2\,T}_{\mathbf{Y};i,T+1|T} := V^{2\,T}_{\mathbf{X};i,T+1|T} + V^{2\,T}_{\mathbf{Z};i,T+1|T}, \quad i = 1, \dots, n. \tag{3.18}$$

This approach of course straightforwardly generalizes to any multi-period ahead forecast.

The parameters in (3.12)-(3.15) are classically estimated by Gaussian Quasi Maximum Likelihood (Bollerslev, 1986) computed from the estimated common and idiosyncratic components of log-volatility proxies obtained in Section 2.2.2. Due to the symmetry of the standard GARCH model, we do not need information about the sign of innovations. On the other hand, if we were to consider leverage effects and therefore asymmetric GARCH specifications, as, for instance, the TARCH model by Zakoian (1994), the sign of the return process would be needed; in this case, we could use the sign of the estimated level-common and level-idiosyncratic residuals, respectively, which are available from Section 2.2.1.

Let us stress once more that we do not require Assumption (C) to hold, so that our approach essentially is a heuristic one. An asymptotic study, on the model of, e.g., Francq and Zakoian (2004) or Hafner and Preminger (2009), could be performed by imposing, on top of (3.10)-(3.11) and (C), mutually independent $v_{it}$'s (hence an exact factor structure). Again, such assumptions are extremely strong and unrealistic, and are not in line with the spirit of the general dynamic factor approach; we will not make them, and prefer an empirical evaluation of our forecasts. Such evaluation is provided in Section 4, and looks quite favorable.

# 4 Forecasting the volatility of S&P100

As an application, we consider the panel of stocks, based on daily adjusted closing prices, used in the construction of the Standard & Poor's 100 (S&P100) index. Since we are interested in forecasting volatilities during the Great Financial Crisis, we limit our study to daily log-returns from 26th January 2000 to 9th December 2009. We have thus an observation period of $2500$ days. Since not all constituents of the index were traded during the observation period, we end up with a panel of $n = 90$ time series.[13]

We estimate the factor models for returns and volatilities as described in Section 2.2. In accordance with the results from the Hallin and Liška (2007) criterion, we set $q = q_s = q_w = 1$. The VAR orders of the 2-dimensional blocks in both estimation steps, and the AR orders for the level-idiosyncratic components are selected by means of BIC.[14]

From this, as explained in Section 3, we can build forecasts in two ways. First, as in Section 3.1, we compute forecasts of the common and idiosyncratic components of log-volatility proxies. The truncation lags for the common components forecasts (3.3)-(3.4) were set to $k^* = h^* = 20$, while for the idiosyncratic components the AR orders $\ell^*_i$ and $m^*_i$ in (3.5)-(3.6) were chosen via BIC. Alternatively, we also adopt HAR specifications for all components, thus taking into account possible long-memory, as suggested by Corsi (2009). We then obtain forecasts of squared volatilities according

[13] The dataset is downloadable from Yahoo Finance and a list of the series used is provided in the Appendix.

[14] Volatilities are know to display long-memory (see for example Andersen et al., 2003); however, as shown in Barigozzi and Hallin (2016) on the same dataset, the fractional differencing parameter $d$ seems to be well below 0.5, thus posing no problem for stationarity.

TABLE 1: *Values of estimated GARCH parameters.*

| level-common common volatility $\omega^T_{\mathbf{s};it}$ | | | level-common idiosyncratic volatility $h^T_{\mathbf{s};it}$ | | |
|---|---|---|---|---|---|
| $\alpha^T_{\mathbf{s};i}$ | $\beta^T_{\mathbf{s};i}$ | $(\alpha^T_{\mathbf{s};i}+\beta^T_{\mathbf{s};i})$ | $a^T_{\mathbf{s};i}$ | $b^T_{\mathbf{s};i}$ | $(a^T_{\mathbf{s};i}+b^T_{\mathbf{s};i})$ |
| 0.055 | 0.939 | 0.994 | 0.595 | 0.317 | 0.912 |
| (0.015) | (0.014) | (0.001) | (0.274) | (0.299) | (0.038) |
| level-idiosyncratic common volatility $\omega^T_{\mathbf{w};it}$ | | | level-idiosyncratic idiosyncratic volatility $h^T_{\mathbf{w};it}$ | | |
| $\alpha^T_{\mathbf{w};i}$ | $\beta^T_{\mathbf{w};i}$ | $(\alpha^T_{\mathbf{w};i}+\beta^T_{\mathbf{w};i})$ | $a^T_{\mathbf{w};i}$ | $b^T_{\mathbf{w};i}$ | $(a^T_{\mathbf{w};i}+b^T_{\mathbf{w};i})$ |
| 0.011 | 0.597 | 0.608 | 0.046 | 0.923 | 0.969 |
| (0.013) | (0.409) | (0.410) | (0.046) | (0.145) | (0.118) |

In each row we report the cross-sectional mean and standard deviation (in parentheses) of the estimated parameters of the GARCH models for the conditional variances. We also report the mean and standard deviation of persistences defined as the sum of the two parameters.

to the two approaches described in the previous section. In particular, we have the combined forecasts $\mathcal{E}^T_{\mathbf{Y};i,T+h|T}$ as defined in (3.9). Multi-period ahead forecast are also defined straightforwardly.

In a second exercise, we estimate, as in Section 3.2, GARCH models for (3.10)-(3.11). Given a sample of length $T$, we then obtain four sets of estimators, $\omega^T_{\mathbf{s};it}$, $h^T_{\mathbf{s};it}$, $\omega^T_{\mathbf{w};it}$, and $h^T_{\mathbf{w};it}$. GARCH orders are selected by BIC, which mostly yields GARCH(1,1) models. Therefore, in Table 1, we report some descriptive statistics of the estimated parameters when considering a GARCH(1,1) model for all series. In particular, it is interesting to look at the values of the cross-sectional average persistence (defined as the sum of the GARCH parameters) in each panel of estimated volatilities. We see that parameter estimates for the level-common volatilities, $\omega^T_{\mathbf{s};it}$ and $h^T_{\mathbf{s};it}$, and level-idiosyncratic idiosyncratic volatilities, $h^T_{\mathbf{w};it}$, display the typical behavior of GARCH models with average persistence very close to one. If we look at cross-sectional standard deviations of persistence, the panel $h^T_{\mathbf{w};it}$, which is idiosyncratic both for levels and volatilities, seems to be quite heterogeneous. On the contrary, the panel $\omega^T_{\mathbf{s};it}$, which is common to levels and volatilities, is highly homogeneous. Finally, the level-idiosyncratic common volatility presents an exception, with lower persistence, 0.61 on average, thus indicating a faster mean reversion in conditional variance with respect to the three other panels.

From (3.12)-(3.15), we build four sets of one-period ahead forecasts: $\omega_{\mathbf{s};i,T+1|T}$, $h_{\mathbf{s};i,T+1|T}$, $\omega_{\mathbf{w};i,T+1|T}$, and $h_{\mathbf{w};i,T+1|T}$. These forecasts are then recombined using (3.16), (3.17), and (3.18), yielding the squared volatility forecast $V^{2\,T}_{\mathbf{Y};i,T|T+1}$. Multi-period ahead forecast are also defined straightforwardly.

Following standard practice, for each series $i = 1, \dots, n$, we compare conditional variance forecasts with the adjusted intra-daily log range, defined by Parkinson (1980) as

$$\rho_{it} := \frac{(\log p_{\text{high};it} - \log p_{\text{low};it})^2}{4 \log 2}, \quad i = 1, \dots, n, \quad t = 1, \dots, T, \tag{4.1}$$

where $p_{\text{high};it}$ and $p_{\text{low};it}$ denote the maximum and the minimum prices of stock $i$ on day $t$, respectively. It has been shown by Alizadeh et al. (2002) and Brownlees and Gallo (2010) that theoretically, numerically, and empirically the adjusted intra-daily log range is a highly efficient volatility proxy robust to microstructure noise and hence at least equally as good as more sophisticated alternatives such as, for example, realized volatilities (Andersen et al., 2003).

We repeat estimation and forecast of the model 500 times corresponding to the 500 days in the period from 14th December 2007 to 9th December 2009, thus including the Great Financial Crisis. Each forecast is computed from the estimation on a rolling sample of $T = 2000$ observations. We consider $h$-period ahead forecasts with $h = 1, 2, 5, 10$. We compute three different forecasts based on our model:

(*i*) the squared innovations forecast $\mathcal{E}_{\mathbf{Y};i,T+h|T}$, as defined in (3.9) and based on HAR dynamics,[15]

(*ii*) the total squared volatility forecast $V^2_{\mathbf{Y};i,T+h|T}$, as defined in (3.18), and

(*iii*) the market squared volatility forecast, $(\omega_{\mathbf{s};i,T+h|T} + \omega_{\mathbf{w};i,T+h|T})$, which can be computed directly from (3.12)-(3.14).

In addition, we also consider the following three competitor models, taken from the classical literature:

(*iv*) componentwise univariate GARCH,

(*v*) multivariate composite likelihood vech-GARCH, and

(*vi*) *static* factor GARCH models.

The univariate GARCH in (*iv*) is performed on the residuals of univariate AR models estimated on each individual series. The vech-GARCH in (*v*) is estimated by means of composite maximum likelihood as in Engle et al. (2008) and is the only multivariate GARCH model able to cope with the high-dimensionality of the dataset at hand, since classical models, as for example BEKK (Engle and Kroner, 1995) or DCC (Engle, 2002), cannot be estimated in a reasonable amount of time (convergence of the maximization algorithm of these models moreover seems problematic even when based on a composite likelihood). Finally, the static factor GARCH model in (*vi*) is in the spirit of Diebold and Nerlove (1989), Van der Weide (2002), Alessi et al. (2009), and Aramonte et al. (2013): in a first step, $r$ factors are extracted by means of static principal components (as in Stock and Watson, 2002, for instance); in a second step, an AR-GARCH model is adjusted to each estimated principal component.[16] The number $r$ of static factors is selected by means of the criteria of Bai and Ng (2002) or Alessi et al. (2010); both criteria identify three static common factors. However, estimating more than one factor might imply a larger estimation error in the second step, and we therefore also consider the case of a single common factor. Finally, we also estimate univariate AR-GARCH models for the idiosyncratic component of the static model.[17]

We know from Patton (2011) that the use of a conditionally unbiased, but imperfect, conditional variance proxy (as the adjusted intra-daily log range used here) can lead to undesirable outcomes in standard methods for comparing conditional variance forecasts. To assess the relative forecasting performances of various methods, a loss function must be chosen such that the ranking of competing forecasts is robust to the presence of noise in the volatility proxy. The root-mean-squared-error (RMSE) satisfies this property and, for the total squared volatility forecast (3.18), is defined as

$$\text{RMSE}_i(h) = \left[\frac{1}{\tau}\sum_{t=0}^{\tau-1}\left(V^2_{\mathbf{Y};i,T+h+t|T+t-1} - \rho_{i,T+h+t}\right)^2\right]^{1/2}, \quad i = 1,\ldots,n,\ h = 1,2,5,$$

[15] Results based on AR specifications are very similar.

[16] The results are unaffected if we estimate the model using the various versions of Kalman filter proposed by Diebold and Nerlove (1989), Harvey et al. (1992), and Sentana et al. (2008).

[17] Results are very similar when we assume homoskedastic idiosyncratic components.

where $\rho_{it}$ is given in (4.1) and $\tau$ is the out-of-sample size considered.

For all forecasts built from the models listed under *(i)-(vi)*, we report, in Tables 2, 3, 4, and 5, RMSEs for the periods 14/12/2007-9/12/2009, 14/12/2007-11/12/2008 which contains the surge and the peak of the Great Financial Crisis, and 11/12/2008-9/12/2009 which contains the aftermath of the Crisis. RMSEs are averaged across all 90 series, and averaged across each of the following sectors—Finance, Energy, Information and Technology, Consumer Discretionary, Consumer Staples, Health Care, and Industry. All values are reported relative to the RMSE of univariate GARCH forecasts.

TABLE 2: *One-period ahead relative RMSEs.*

| 14/12/2007-9/12/2009 | All series | FINA | ENER | INFT | COND | CONS | HEAL | INDU |
|---|---|---|---|---|---|---|---|---|
| $\mathcal{E}^T_{\mathbf{Y};i,T+1\vert T}$ | 0.745 | 0.845 | 0.719 | 0.705 | 0.776 | 0.808 | 0.759 | 0.572 |
| $V^{2\,T}_{\mathbf{Y};i,T+1\vert T}$ | 0.670 | 0.703 | 0.599 | 0.950 | 0.664 | 0.911 | 0.735 | 0.390 |
| $(\omega^T_{\mathbf{s};i,T+1\vert T}+\omega^T_{\mathbf{w};i,T+1\vert T})$ | 0.651 | 0.733 | 0.532 | 0.824 | 0.654 | 0.825 | 0.739 | 0.392 |
| static factor GARCH ($r=1$) | 0.964 | 0.893 | 0.800 | 1.346 | 0.975 | 1.145 | 1.101 | 0.919 |
| static factor GARCH ($r=3$) | 0.963 | 0.852 | 0.762 | 1.431 | 0.999 | 1.254 | 1.150 | 0.934 |
| vech-GARCH | 0.778 | 0.764 | 0.675 | 0.911 | 0.996 | 1.003 | 0.726 | 0.696 |
| 14/12/2007-11/12/2008 | All series | FINA | ENER | INFT | COND | CONS | HEAL | INDU |
| $\mathcal{E}^T_{\mathbf{Y};i,T+1\vert T}$ | 0.921 | 0.999 | 0.965 | 0.828 | 0.971 | 0.993 | 0.868 | 0.829 |
| $V^{2\,T}_{\mathbf{Y};i,T+1\vert T}$ | 0.828 | 0.814 | 0.804 | 1.130 | 0.830 | 1.115 | 0.824 | 0.544 |
| $(\omega^T_{\mathbf{s};i,T+1\vert T}+\omega^T_{\mathbf{w};i,T+1\vert T})$ | 0.782 | 0.830 | 0.704 | 0.961 | 0.807 | 0.941 | 0.791 | 0.535 |
| static factor GARCH ($r=1$) | 0.985 | 0.925 | 0.816 | 1.317 | 0.981 | 1.082 | 0.993 | 1.016 |
| static factor GARCH ($r=3$) | 0.992 | 0.887 | 0.787 | 1.405 | 0.998 | 1.206 | 1.047 | 1.044 |
| vech-GARCH | 0.909 | 0.886 | 0.898 | 0.931 | 1.090 | 1.052 | 0.812 | 0.912 |
| 11/12/2008-9/12/2009 | All series | FINA | ENER | INFT | COND | CONS | HEAL | INDU |
| $\mathcal{E}^T_{\mathbf{Y};i,T+1\vert T}$ | 0.494 | 0.744 | 0.359 | 0.217 | 0.407 | 0.365 | 0.329 | 0.478 |
| $V^{2\,T}_{\mathbf{Y};i,T+1\vert T}$ | 0.675 | 0.793 | 0.441 | 0.928 | 0.548 | 0.863 | 0.689 | 0.529 |
| $(\omega^T_{\mathbf{s};i,T+1\vert T}+\omega^T_{\mathbf{w};i,T+1\vert T})$ | 0.682 | 0.764 | 0.458 | 0.815 | 0.613 | 0.872 | 0.776 | 0.625 |
| static factor GARCH ($r=1$) | 1.008 | 0.970 | 0.849 | 1.370 | 1.070 | 1.058 | 0.971 | 1.009 |
| static factor GARCH ($r=3$) | 1.015 | 0.952 | 0.821 | 1.438 | 1.097 | 1.152 | 1.030 | 1.031 |
| vech-GARCH | 0.573 | 0.695 | 0.360 | 0.519 | 0.729 | 0.654 | 0.361 | 0.608 |

In each row we report, for one-period ahead forecasts, the average relative RMSE across all 90 series or across the series of a given sector. FINA: Financials; ENER: Energy; INFT: Information and Technology; COND: Consumer Discretionary; CONS: Consumer Staples; HEAL: Health Care; INDU: Industrials.

To fully appreciate the performance of the different forecasting methods in Figure 1, we show the distribution of relative RMSEs across all series and computed on rolling windows of $\tau = 20$ days.

To conclude, in Figures 2 (financial sector) and 3 (other sectors), we compare, for selected stocks, our one-period ahead total conditional variance forecasts $V^2_{\mathbf{Y};i,T+1|T}$ with the forecasts resulting from

TABLE 3: *Two-period ahead relative RMSEs.*

| 14/12/2007-9/12/2009 | All series | FINA | ENER | INFT | COND | CONS | HEAL | INDU |
|---|---|---|---|---|---|---|---|---|
| $\mathcal{E}^T_{\mathbf{Y};i,T+2\vert T}$ | 0.616 | 0.826 | 0.559 | 0.379 | 0.555 | 0.591 | 0.561 | 0.509 |
| $V^{2\,T}_{\mathbf{Y};i,T+2\vert T}$ | 0.582 | 0.692 | 0.471 | 0.791 | 0.448 | 0.795 | 0.615 | 0.363 |
| $(\omega^T_{\mathbf{s};i,T+2\vert T} + \omega^T_{\mathbf{w};i,T+2\vert T})$ | 0.583 | 0.728 | 0.448 | 0.654 | 0.456 | 0.752 | 0.661 | 0.407 |
| static factor GARCH ($r = 1$) | 0.974 | 0.904 | 0.800 | 1.412 | 1.004 | 1.116 | 1.081 | 0.917 |
| static factor GARCH ($r = 3$) | 0.972 | 0.861 | 0.759 | 1.493 | 1.034 | 1.216 | 1.136 | 0.933 |
| vech-GARCH | 0.671 | 0.756 | 0.518 | 0.655 | 0.835 | 0.858 | 0.562 | 0.643 |
| 14/12/2007-11/12/2008 | All series | FINA | ENER | INFT | COND | CONS | HEAL | INDU |
| $\mathcal{E}^T_{\mathbf{Y};i,T+2\vert T}$ | 0.597 | 0.675 | 0.609 | 0.412 | 0.664 | 0.609 | 0.605 | 0.588 |
| $V^{2\,T}_{\mathbf{Y};i,T+2\vert T}$ | 0.597 | 0.511 | 0.546 | 0.951 | 0.557 | 0.890 | 0.689 | 0.419 |
| $(\omega^T_{\mathbf{s};i,T+2\vert T} + \omega^T_{\mathbf{w};i,T+2\vert T})$ | 0.570 | 0.518 | 0.490 | 0.765 | 0.549 | 0.808 | 0.710 | 0.490 |
| static factor GARCH ($r = 1$) | 0.987 | 0.876 | 0.832 | 1.399 | 1.048 | 1.054 | 0.972 | 1.008 |
| static factor GARCH ($r = 3$) | 0.985 | 0.807 | 0.796 | 1.478 | 1.073 | 1.165 | 1.035 | 1.032 |
| vech-GARCH | 0.624 | 0.595 | 0.547 | 0.577 | 0.862 | 0.787 | 0.597 | 0.720 |
| 11/12/2008-9/12/2009 | All series | FINA | ENER | INFT | COND | CONS | HEAL | INDU |
| $\mathcal{E}^T_{\mathbf{Y};i,T+2\vert T}$ | 0.410 | 0.638 | 0.304 | 0.218 | 0.353 | 0.240 | 0.314 | 0.379 |
| $V^{2\,T}_{\mathbf{Y};i,T+2\vert T}$ | 0.615 | 0.700 | 0.396 | 0.941 | 0.507 | 0.837 | 0.665 | 0.462 |
| $(\omega^T_{\mathbf{s};i,T+2\vert T} + \omega^T_{\mathbf{w};i,T+2\vert T})$ | 0.620 | 0.661 | 0.418 | 0.801 | 0.565 | 0.849 | 0.748 | 0.578 |
| static factor GARCH ($r = 1$) | 0.999 | 0.943 | 0.843 | 1.363 | 1.068 | 1.047 | 0.970 | 1.009 |
| static factor GARCH ($r = 3$) | 1.003 | 0.912 | 0.815 | 1.432 | 1.091 | 1.147 | 1.031 | 1.031 |
| vech-GARCH | 0.503 | 0.586 | 0.309 | 0.512 | 0.681 | 0.575 | 0.354 | 0.559 |

In each row we report, for two-period ahead forecasts, the average relative RMSE across all 90 series or across the series of a given sector. FINA: Financials; ENER: Energy; INFT: Information and Technology; COND: Consumer Discretionary; CONS: Consumer Staples; HEAL: Health Care; INDU: Industrials.

a static factor GARCH model—which, according to the results presented, seems to be the best competitor. Forecasts are plotted together with the adjusted intra-daily log range.

Inspection of results reveals that, overall, we tend to outperform all competing models considered in regard to the total squared volatility forecast $V^2_{\mathbf{Y};i,T+h|T}$ and sometimes also when considering the squared innovations forecasts $\mathcal{E}_{\mathbf{Y};i,T+h|T}$. In detail, when focussing on different time windows, we notice that our method strongly outperforms the others during periods of relative quiet in the market while during crisis it tends to be slightly worse than the static factor GARCH model. A possible explanation could be that, due to the high collinearity and lesser persistence (caused by continuous and abrupt fluctuations in the market) of the series under consideration during the Financial Crisis, co-movements can easily explained with just one principal component as in the static factor GARCH with one factor. On the other hand, during quieter periods, the role of idiosyncratic returns and volatilities

TABLE 4: *Five-period ahead relative RMSEs.*

| 14/12/2007-9/12/2009 | All series | FINA | ENER | INFT | COND | CONS | HEAL | INDU |
|---|---|---|---|---|---|---|---|---|
| $\mathcal{E}^T_{\mathbf{Y};i,T+5\|T}$ | 0.908 | 1.044 | 0.905 | 0.573 | 0.917 | 0.991 | 0.968 | 0.785 |
| $V^{2\,T}_{\mathbf{Y};i,T+5\|T}$ | 0.828 | 0.924 | 0.793 | 0.838 | 0.826 | 0.985 | 0.893 | 0.586 |
| $(\omega^T_{\mathbf{s};i,T+5\|T} + \omega^T_{\mathbf{w};i,T+5\|T})$ | 0.782 | 0.950 | 0.728 | 0.639 | 0.787 | 0.803 | 0.800 | 0.529 |
| static factor GARCH ($r = 1$) | 0.979 | 0.920 | 0.828 | 1.448 | 0.977 | 1.129 | 1.139 | 0.894 |
| static factor GARCH ($r = 3$) | 0.978 | 0.885 | 0.785 | 1.544 | 0.990 | 1.231 | 1.186 | 0.905 |
| vech-GARCH | 0.914 | 0.952 | 0.851 | 0.770 | 1.028 | 1.147 | 0.932 | 0.857 |
| **14/12/2007-11/12/2008** | **All series** | **FINA** | **ENER** | **INFT** | **COND** | **CONS** | **HEAL** | **INDU** |
| $\mathcal{E}^T_{\mathbf{Y};i,T+5\|T}$ | 1.084 | 1.155 | 1.145 | 0.675 | 1.071 | 1.196 | 1.123 | 1.152 |
| $V^{2\,T}_{\mathbf{Y};i,T+5\|T}$ | 1.010 | 1.033 | 1.021 | 1.057 | 0.977 | 1.210 | 1.059 | 0.860 |
| $(\omega^T_{\mathbf{s};i,T+5\|T} + \omega^T_{\mathbf{w};i,T+5\|T})$ | 0.916 | 1.032 | 0.905 | 0.781 | 0.924 | 0.903 | 0.906 | 0.736 |
| static factor GARCH ($r = 1$) | 1.005 | 0.956 | 0.864 | 1.446 | 0.982 | 1.064 | 1.052 | 0.974 |
| static factor GARCH ($r = 3$) | 1.007 | 0.920 | 0.823 | 1.548 | 0.990 | 1.176 | 1.100 | 0.990 |
| vech-GARCH | 1.041 | 1.054 | 1.059 | 0.729 | 1.094 | 1.210 | 1.061 | 1.166 |
| **11/12/2008-9/12/2009** | **All series** | **FINA** | **ENER** | **INFT** | **COND** | **CONS** | **HEAL** | **INDU** |
| $\mathcal{E}^T_{\mathbf{Y};i,T+5\|T}$ | 0.516 | 0.758 | 0.381 | 0.264 | 0.375 | 0.343 | 0.319 | 0.478 |
| $V^{2\,T}_{\mathbf{Y};i,T+5\|T}$ | 0.647 | 0.776 | 0.440 | 0.866 | 0.468 | 0.755 | 0.609 | 0.492 |
| $(\omega^T_{\mathbf{s};i,T+5\|T} + \omega^T_{\mathbf{w};i,T+5\|T})$ | 0.672 | 0.766 | 0.464 | 0.758 | 0.552 | 0.844 | 0.729 | 0.605 |
| static factor GARCH ($r = 1$) | 1.001 | 0.961 | 0.845 | 1.352 | 1.059 | 1.045 | 0.971 | 1.001 |
| static factor GARCH ($r = 3$) | 1.006 | 0.941 | 0.816 | 1.422 | 1.082 | 1.144 | 1.033 | 1.016 |
| vech-GARCH | 0.580 | 0.689 | 0.382 | 0.520 | 0.699 | 0.630 | 0.352 | 0.608 |

In each row we report, for five-period ahead forecasts, the average relative RMSE across all 90 series or across the series of a given sector. FINA: Financials; ENER: Energy; INFT: Information and Technology; COND: Consumer Discretionary; CONS: Consumer Staples; HEAL: Health Care; INDU: Industrials.

becomes important, and our model seems to better disentangle those dynamics specific to the single series from those related to the market. Summing up, our losses with respect to other models are limited during periods of high volatility while our gains are quite substantial in the other periods and therefore, over all days considered, our approach delivers a better performance both on average across stocks and for many individual stocks—in particular, the Financial ones.

## 5 Conclusion

In this paper, we propose a two-step general dynamic factor method for the analysis of financial volatilities in large panels of stock returns. Our focus throughout is to produce measures of squared

TABLE 5: *Ten-period ahead relative RMSEs.*

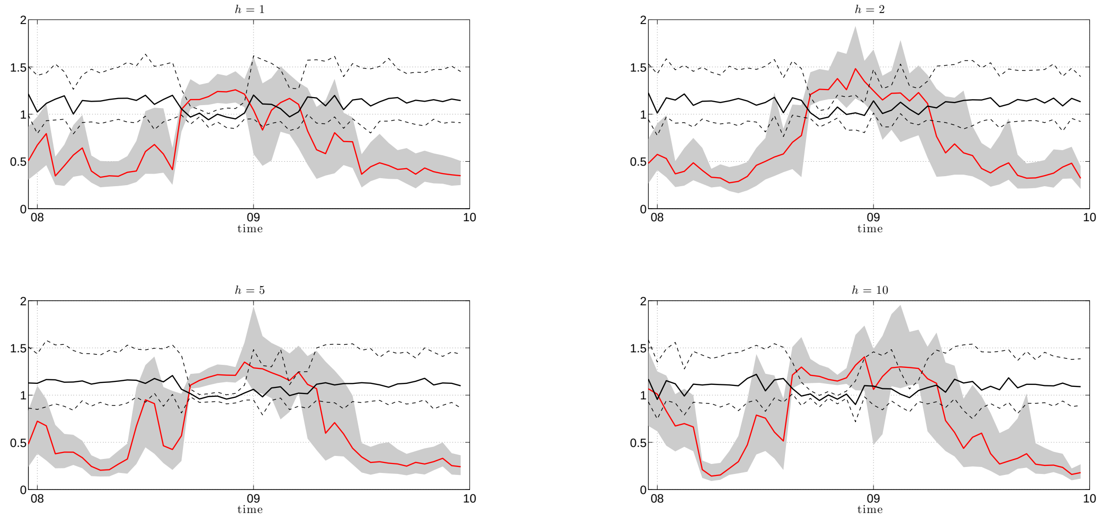
| 14/12/2007-9/12/2009 | All series | FINA | ENER | INFT | COND | CONS | HEAL | INDU |
|---|---|---|---|---|---|---|---|---|
| $\mathcal{E}^T_{\mathbf{Y};i,T+10\vert T}$ | 0.801 | 0.849 | 0.797 | 0.703 | 0.810 | 0.790 | 0.838 | 0.772 |
| $V^{2\ T}_{\mathbf{Y};i,T+10\vert T}$ | 0.762 | 0.803 | 0.699 | 0.899 | 0.749 | 0.839 | 0.803 | 0.649 |
| $(\omega^T_{\mathbf{s};i,T+10\vert T} + \omega^T_{\mathbf{w};i,T+10\vert T})$ | 0.739 | 0.804 | 0.673 | 0.744 | 0.720 | 0.845 | 0.792 | 0.620 |
| static factor GARCH ($r = 1$) | 0.954 | 0.918 | 0.801 | 1.374 | 0.955 | 1.100 | 1.061 | 0.895 |
| static factor GARCH ($r = 3$) | 0.947 | 0.884 | 0.757 | 1.455 | 0.974 | 1.178 | 1.095 | 0.894 |
| vech-GARCH | 0.819 | 0.799 | 0.745 | 0.885 | 0.979 | 0.997 | 0.811 | 0.826 |
| 14/12/2007-11/12/2008 | All series | FINA | ENER | INFT | COND | CONS | HEAL | INDU |
| $\mathcal{E}^T_{\mathbf{Y};i,T+10\vert T}$ | 0.912 | 0.900 | 0.945 | 0.822 | 0.984 | 0.901 | 0.848 | 1.041 |
| $V^{2\ T}_{\mathbf{Y};i,T+10\vert T}$ | 0.888 | 0.867 | 0.847 | 1.089 | 0.924 | 0.976 | 0.843 | 0.890 |
| $(\omega^T_{\mathbf{s};i,T+10\vert T} + \omega^T_{\mathbf{w};i,T+10\vert T})$ | 0.847 | 0.859 | 0.792 | 0.890 | 0.887 | 0.930 | 0.812 | 0.844 |
| static factor GARCH ($r = 1$) | 0.972 | 0.945 | 0.839 | 1.349 | 0.965 | 1.037 | 0.975 | 0.966 |
| static factor GARCH ($r = 3$) | 0.965 | 0.916 | 0.791 | 1.426 | 0.978 | 1.116 | 1.010 | 0.961 |
| vech-GARCH | 0.897 | 0.859 | 0.871 | 0.884 | 1.046 | 0.986 | 0.825 | 1.043 |
| 11/12/2008-9/12/2009 | All series | FINA | ENER | INFT | COND | CONS | HEAL | INDU |
| $\mathcal{E}^T_{\mathbf{Y};i,T+10\vert T}$ | 0.406 | 0.646 | 0.271 | 0.192 | 0.284 | 0.313 | 0.336 | 0.358 |
| $V^{2\ T}_{\mathbf{Y};i,T+10\vert T}$ | 0.532 | 0.671 | 0.335 | 0.741 | 0.367 | 0.631 | 0.571 | 0.369 |
| $(\omega^T_{\mathbf{s};i,T+10\vert T} + \omega^T_{\mathbf{w};i,T+10\vert T})$ | 0.579 | 0.668 | 0.385 | 0.676 | 0.471 | 0.780 | 0.700 | 0.522 |
| static factor GARCH ($r = 1$) | 0.988 | 0.949 | 0.824 | 1.333 | 1.042 | 1.052 | 0.967 | 0.986 |
| static factor GARCH ($r = 3$) | 0.990 | 0.920 | 0.790 | 1.404 | 1.065 | 1.141 | 1.025 | 0.998 |
| vech-GARCH | 0.515 | 0.602 | 0.281 | 0.531 | 0.682 | 0.643 | 0.404 | 0.561 |

In each row we report, for five-period ahead forecasts, the average relative RMSE across all 90 series or across the series of a given sector. FINA: Financials; ENER: Energy; INFT: Information and Technology; COND: Consumer Discretionary; CONS: Consumer Staples; HEAL: Health Care; INDU: Industrials.

innovations and of their conditional mean as proxies of squared volatilities and to produce multi-period ahead forecasts of the same.

In a previous paper, we showed that the decomposition into "common" and "idiosyncratic" component of the returns does not necessarily coincide with the corresponding decomposition for volatilities, in the sense that level-idiosyncratic components, just as much as the the level-common ones, are affected by market volatility shocks (Barigozzi and Hallin, 2016). Here, based on this finding, we propose a "divide and rule" analysis of volatilities by decomposing them into four different components: common and idiosyncratic of level-common innovations and common and idiosyncratic of level-idiosyncratic innovations. In Section 4 we show that, for the assets composing the S&P100 index, GARCH forecasts based on those four components are generally better than univariate and other factor based forecasts when compared with the adjusted intra-daily log range.

FIGURE 1: *Relative RMSEs over the period December 2007– December 2009.*




Median for our model (red) and static factor GARCH with one factor (black) with related 25th and 75th percentiles.

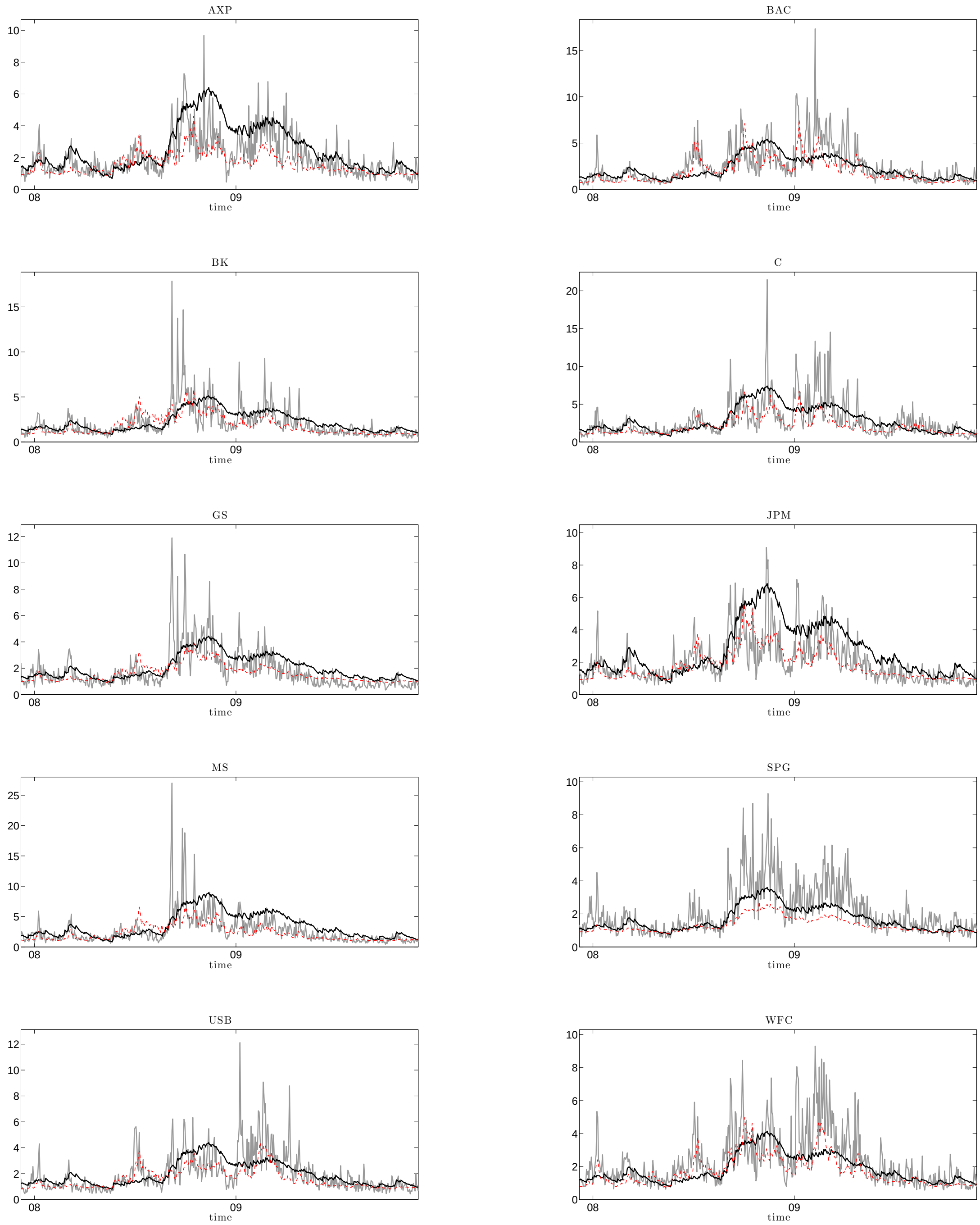
The present framework can be extended in many directions of potential interest in financial econometrics and risk management. In particular, two extensions are under study: (*i*) the construction of conditional prediction intervals for returns, providing estimated conditional Value at Risk values, and (*ii*) the estimation of optimal portfolios in the Markowitz sense.

TABLE 6: *Relative RMSEs for the selected stocks displayed in Figures 2 and 3.*

| Ticker | $\mathcal{E}^T_{\mathbf{Y};i,T+h}$ | $V^{2\,T}_{\mathbf{Y};i,T+h}$ | static factor GARCH ($r=1$) | Ticker | $\mathcal{E}^T_{\mathbf{Y};i,T+h}$ | $V^{2\,T}_{\mathbf{Y};i,T+h}$ | static factor GARCH ($r=1$) |
|---|---|---|---|---|---|---|---|
| | | | $h=1$ | | | | |
| AXP | 0.572 | 0.394 | 0.871 | FDX | 0.444 | 0.319 | 0.964 |
| BAC | 0.710 | 0.508 | 0.889 | UNP | 0.763 | 0.473 | 1.170 |
| BK | 0.868 | 0.696 | 0.847 | DELL | 0.471 | 0.285 | 0.749 |
| C | 0.920 | 0.729 | 0.853 | IBM | 0.996 | 0.853 | 1.605 |
| GS | 0.908 | 0.745 | 0.963 | MSFT | 0.583 | 0.558 | 0.764 |
| JPM | 0.465 | 0.313 | 0.945 | DIS | 1.072 | 0.926 | 1.019 |
| MS | 0.693 | 0.517 | 0.934 | HD | 0.830 | 0.478 | 1.661 |
| SPG | 0.694 | 0.571 | 0.668 | SBUX | 0.568 | 0.546 | 1.106 |
| USB | 0.689 | 0.473 | 0.837 | WMT | 0.482 | 0.292 | 0.719 |
| WFC | 0.963 | 0.687 | 0.844 | COF | 0.569 | 0.345 | 0.779 |
| | | | $h=2$ | | | | |
| AXP | 0.491 | 0.359 | 0.920 | FDX | 0.491 | 0.381 | 0.960 |
| BAC | 0.740 | 0.655 | 0.946 | UNP | 0.546 | 0.292 | 1.220 |
| BK | 0.729 | 0.570 | 0.941 | DELL | 0.382 | 0.238 | 0.746 |
| C | 0.997 | 0.871 | 0.866 | IBM | 0.650 | 0.743 | 1.714 |
| GS | 0.654 | 0.445 | 0.967 | MSFT | 0.524 | 0.490 | 0.741 |
| JPM | 0.477 | 0.368 | 0.937 | DIS | 0.398 | 0.764 | 1.350 |
| MS | 0.772 | 0.563 | 0.968 | HD | 0.364 | 0.341 | 1.516 |
| SPG | 0.697 | 0.542 | 0.580 | SBUX | 0.311 | 0.413 | 1.153 |
| USB | 0.428 | 0.328 | 0.815 | WMT | 0.402 | 0.243 | 0.724 |
| WFC | 0.756 | 0.550 | 0.841 | COF | 0.581 | 0.382 | 0.789 |
| | | | $h=5$ | | | | |
| AXP | 0.804 | 0.693 | 0.880 | FDX | 0.639 | 0.490 | 0.940 |
| BAC | 1.024 | 0.834 | 0.933 | UNP | 0.734 | 0.446 | 1.166 |
| BK | 0.986 | 0.851 | 0.935 | DELL | 0.518 | 0.361 | 0.726 |
| C | 1.091 | 0.995 | 0.909 | IBM | 1.074 | 0.910 | 1.469 |
| GS | 0.865 | 0.700 | 0.850 | MSFT | 0.778 | 0.710 | 0.772 |
| JPM | 0.902 | 0.779 | 0.906 | DIS | 0.440 | 0.692 | 1.366 |
| MS | 1.199 | 1.056 | 0.996 | HD | 0.985 | 0.708 | 1.535 |
| SPG | 0.951 | 0.808 | 0.745 | SBUX | 0.456 | 0.399 | 1.131 |
| USB | 1.056 | 0.990 | 0.962 | WMT | 0.651 | 0.476 | 0.706 |
| WFC | 1.085 | 0.903 | 0.907 | COF | 0.948 | 0.785 | 0.799 |
| | | | $h=10$ | | | | |
| AXP | 1.065 | 0.988 | 0.883 | FDX | 0.523 | 0.451 | 0.932 |
| BAC | 0.722 | 0.607 | 0.891 | UNP | 0.860 | 0.754 | 1.068 |
| BK | 1.007 | 0.995 | 0.997 | DELL | 0.802 | 0.693 | 0.791 |
| C | 0.688 | 0.624 | 0.846 | IBM | 0.936 | 0.873 | 1.360 |
| GS | 0.970 | 0.938 | 0.979 | MSFT | 0.787 | 0.709 | 0.740 |
| JPM | 0.556 | 0.473 | 0.931 | DIS | 0.832 | 0.809 | 1.316 |
| MS | 0.998 | 0.992 | 0.995 | HD | 0.908 | 0.725 | 1.314 |
| SPG | 0.859 | 0.781 | 0.753 | SBUX | 0.760 | 0.647 | 1.095 |
| USB | 0.719 | 0.571 | 0.750 | WMT | 0.484 | 0.373 | 0.712 |
| WFC | 0.912 | 0.794 | 0.846 | COF | 0.839 | 0.709 | 0.783 |

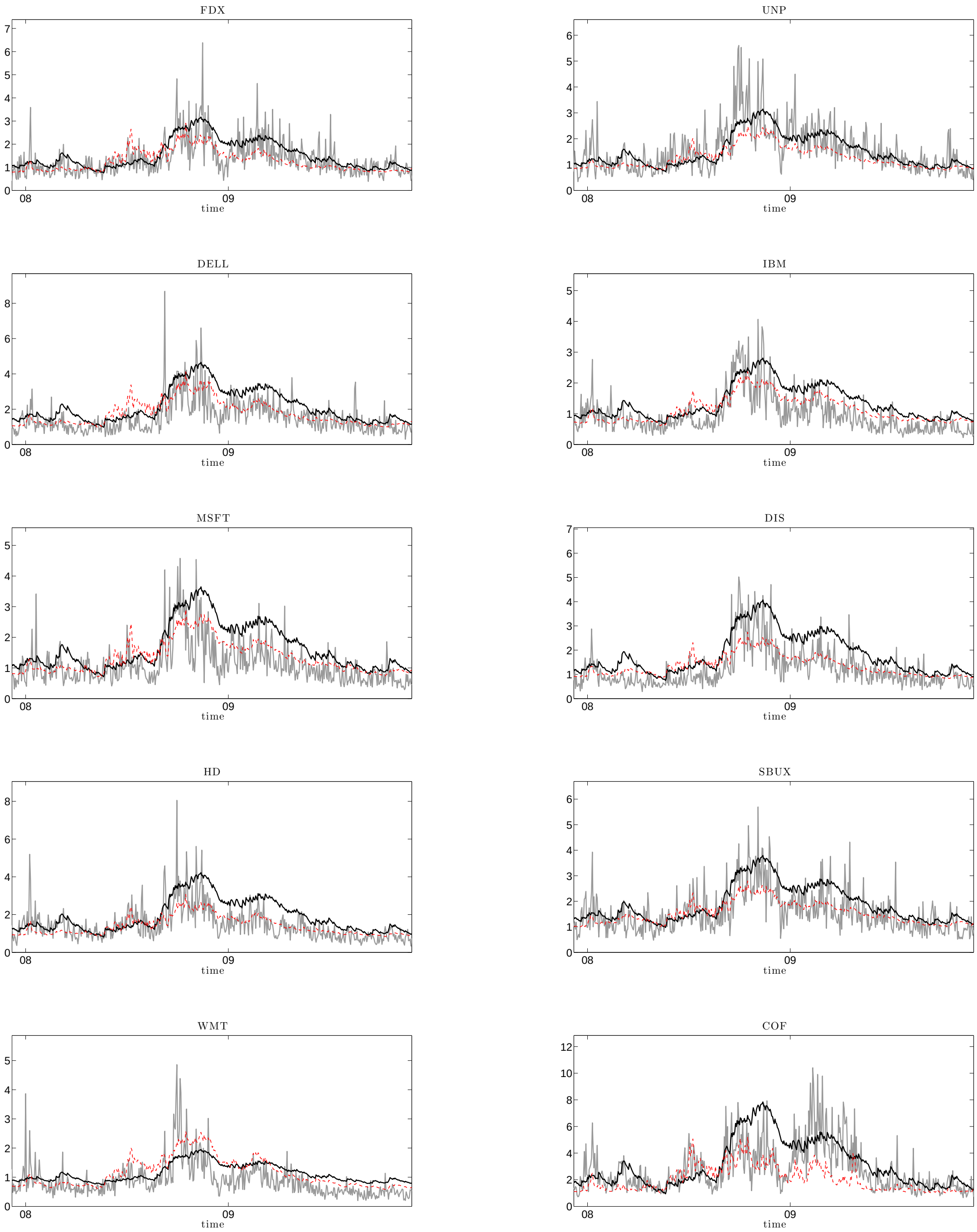
For each stock we report, for $h$-period ahead forecast with $h = 1, 2, 5, 10$, the RMSE (relative to univariate GARCH forecasts) for our two approaches and for the static factor GARCH model. See the Appendix for tickers' definitions.

FIGURE 2: *Squared volatility forecasts – Financial sector.*




One-period ahead forecasts of squared volatility obtained from our model (red) and from a static factor GARCH with one factor (black) for selected stocks from the financial sector, along with the observed adjusted intra-daily log range (light grey). See Appendix for tickers' definitions.

FIGURE 3: *Squared volatility forecasts – Other sectors.*




One-period ahead forecasts of squared volatility obtained from our model (red) and from a static factor GARCH with one factor (black) for selected stocks from the financial sector, along with the observed adjusted intra-daily log range (light grey). See Appendix for tickers' definitions.

## References

Alessi, L., M. Barigozzi, and M. Capasso (2009). Estimation and forecasting in large datasets with conditionally heteroskedastic dynamic common factors. Working Paper 1115, European Central Bank.

Alessi, L., M. Barigozzi, and M. Capasso (2010). Improved penalization for determining the number of factors in approximate static factor models. *Statistics and Probability Letters 80*, 1806–1813.

Alizadeh, S., M. W. Brandt, and F. X. Diebold (2002). Range-based estimation of stochastic volatility models. *The Journal of Finance 57*, 1047–1091.

Andersen, T. G., T. Bollerslev, F. X. Diebold, and P. Labys (2003). Modeling and forecasting realized volatility. *Econometrica 71*, 579–625.

Aramonte, S., M. del Giudice Rodriguez, and J. Wu (2013). Dynamic factor value-at-risk for large heteroskedastic portfolios. *Journal of Banking & Finance 37*, 4299–4309.

Asai, M., M. McAleer, and J. Yu (2006). Multivariate stochastic volatility: A review. *Econometric Reviews 25*, 145–175.

Bai, J. and S. Ng (2002). Determining the number of factors in approximate factor models. *Econometrica 70*, 191–221.

Barigozzi, M., C. T. Brownlees, G. M. Gallo, and D. Veredas (2014). Disentangling systematic and idiosyncratic dynamics in panels of volatility measures. *Journal of Econometrics 182*, 364–384.

Barigozzi, M. and M. Hallin (2016). Generalized dynamic factor models and volatilities: recovering the market volatility shocks. *The Econometrics Journal 19*, 33–60.

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

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics 31*, 307–327.

Brownlees, C. T. and G. M. Gallo (2010). Comparison of volatility measures: A risk management perspective. *Journal of Financial Econometrics 8*, 29–56.

Connor, G. and R. A. Korajczyk (1986). Performance measurement with the arbitrage pricing theorem. a new framework for analysis. *Journal of Financial Economics 15*, 373–394.

Connor, G., R. A. Korajczyk, and O. Linton (2006). The common and specific components of dynamic volatility. *Journal of Econometrics 132*, 231–255.

Corsi, F. (2009). A simple approximate long-memory model of realized volatility. *Journal of Financial Econometrics 7*, 174–196.

Diebold, F. X. and M. Nerlove (1989). The dynamics of exchange rate volatility: a multivariate latent factor ARCH model. *Journal of Applied Econometrics 4*, 1–21.

Engle, R. F. (2002). Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *Journal of Business & Economic Statistics 20*, 339–350.

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

Engle, R. F. and J. Marcucci (2006). A long–run pure variance common features model for the common volatilities of the Dow Jones. *Journal of Econometrics 132*, 7–42.

Engle, R. F., N. Shephard, and K. Sheppard (2008). Fitting vast dimensional time–varying covariance models. mimeo.

Fan, J., Y. Liao, and M. Mincheva (2013). Large covariance estimation by thresholding principal orthogonal complements. *Journal of the Royal Statistical Society, Series B 75*, 603–680.

Fan, J., Y. Liao, and X. Shi (2015). Risk of large portfolios. *Journal of Econometrics 186*, 367–387.

Forni, M., M. Hallin, M. Lippi, and L. Reichlin (2000). The Generalized Dynamic Factor Model: identification and estimation. *The Review of Economics and Statistics 82*, 540–554.

Forni, M., M. Hallin, M. Lippi, and L. Reichlin (2005). The Generalized Dynamic Factor Model: one-sided estimation and forecasting. *Journal of the American Statistical Association 100*, 830–840.

Forni, M., M. Hallin, M. Lippi, and P. Zaffaroni (2015a). Dynamic factor models with infinite-dimensional factor space: Asymptotic analysis. ECARES Working Paper 2015-23, Université libre de Bruxelles, Belgium.

Forni, M., M. Hallin, M. Lippi, and P. Zaffaroni (2015b). Dynamic factor models with infinite-dimensional factor spaces: one-sided representations. *Journal of Econometrics 185*, 359–371.

Forni, M. and M. Lippi (2001). The Generalized Dynamic Factor Model: representation theory. *Econometric Theory 17*, 1113–1141.

Forni, M. and M. Lippi (2011). The unrestricted Dynamic Factor Model: one-sided representation results. *Journal of Econometrics 163*, 23–28.

Francq, C. and J.-M. Zakoian (2004). Maximum likelihood estimation of pure GARCH and ARMA-GARCH processes. *Bernoulli 10*, 605–637.

Ghysels, E. (2014). Factor analysis with large panels of volatility proxies. Available at ssrn: http://ssrn.com/abstract=2412988 or http://dx.doi.org/10.2139/ssrn.2412988.

Hafner, C. and A. Preminger (2009). Asymptotic theory for a factor GARCH model. *Econometric Theory 25*, 336–363.

Hallin, M. and M. Lippi (2014). Factor models in high–dimensional time series. A time-domain approach. *Stochastic Processes and their Applications 123*, 2678–2695.

Hallin, M. and R. Liška (2007). Determining the number of factors in the general dynamic factor model. *Journal of the American Statistical Association 102*, 603–617.

Harvey, A., E. Ruiz, and E. Sentana (1992). Unobserved component time series models with arch disturbances. *Journal of Econometrics 52*, 129–157.

Harvey, A., E. Ruiz, and N. Shephard (1994). Multivariate stochastic variance models. *The Review of Economic Studies 61*, 247–264.

Luciani, M. (2014). Forecasting with approximate dynamic factor models: the role of non- pervasive shocks. *International Journal of Forecasting 30*, 20–29.

Luciani, M. and D. Veredas (2015). Estimating and forecasting large panels of volatilities with approximate dynamic factor models. *Journal of Forecasting 34*, 163–176.

Ng, V., R. F. Engle, and M. Rothschild (1992). A multi-dynamic-factor model for stock returns. *Journal of Econometrics 52*, 245–266.

Parkinson, M. (1980). The extreme value method for estimating the variance of the rate of return. *The Journal of Business 53*, 61–65.

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

Rangel, J. G. and R. F. Engle (2012). The Factor–Spline–GARCH model for high and low frequency correlations. *Journal of Business & Economic Statistics 30*, 109–124.

Sentana, E., G. Calzolari, and G. Fiorentini (2008). Indirect estimation of large conditionally heteroskedastic factor models, with an application to the Dow 30 stocks. *Journal of Econometrics 146*, 10–25.

Silvennoinen, A. and T. Teräsvirta (2009). Multivariate garch models. In *Handbook of Financial Time Series*, pp. 201–229. Springer.

Stock, J. H. and M. W. Watson (2002). Forecasting using principal components from a large number of predictors. *Journal of the American Statistical Association 97*, 1167–1179.

Van der Weide, R. (2002). GO–GARCH: A multivariate generalized orthogonal GARCH model. *Journal of Applied Econometrics 17*, 549–564.

Zakoian, J.-M. (1994). Threshold heteroskedastic models. *Journal of Economic Dynamics and Control 18*, 931–955.

# A Data

TABLE 7: *S&P100 consituents.*

| Ticker | Name | | |
|---|---|---|---|
| AAPL | Apple Inc. | HPQ | Hewlett Packard Co. |
| ABT | Abbott Laboratories | IBM | International Business Machines |
| AEP | American Electric Power Co. | INTC | Intel Corporation |
| AIG | American International Group Inc. | JNJ | Johnson & Johnson Inc. |
| ALL | Allstate Corp. | JPM | JP Morgan Chase & Co. |
| AMGN | Amgen Inc. | KO | The Coca-Cola Company |
| AMZN | Amazon.com | LLY | Eli Lilly and Company |
| APA | Apache Corp. | LMT | Lockheed-Martin |
| APC | Anadarko Petroleum Corp. | LOW | Lowe’s |
| AXP | American Express Inc. | MCD | McDonald’s Corp. |
| BA | Boeing Co. | MDT | Medtronic Inc. |
| BAC | Bank of America Corp. | MMM | 3M Company |
| BAX | Baxter International Inc. | MO | Altria Group |
| BK | Bank of New York | MRK | Merck & Co. |
| BMY | Bristol-Myers Squibb | MS | Morgan Stanley |
| BRK.B | Berkshire Hathaway | MSFT | Microsoft |
| C | Citigroup Inc. | NKE | Nike |
| CAT | Caterpillar Inc. | NOV | National Oilwell Varco |
| CL | Colgate-Palmolive Co. | NSC | Norfolk Southern Corp. |
| CMCSA | Comcast Corp. | ORCL | Oracle Corporation |
| COF | Capital One Financial Corp. | OXY | Occidental Petroleum Corp. |
| COP | ConocoPhillips | PEP | Pepsico Inc. |
| COST | Costco | PFE | Pfizer Inc. |
| CSCO | Cisco Systems | PG | Procter & Gamble Co. |
| CVS | CVS Caremark | QCOM | Qualcomm Inc. |
| CVX | Chevron | RTN | Raytheon Co. |
| DD | DuPont | SBUX | Starbucks Corporation |
| DELL | Dell | SLB | Schlumberger |
| DIS | The Walt Disney Company | SO | Southern Company |
| DOW | Dow Chemical | SPG | Simon Property Group, Inc. |
| DVN | Devon Energy | T | AT&T Inc. |
| EBAY | eBay Inc. | TGT | Target Corp. |
| EMC | EMC Corporation | TWX | Time Warner Inc. |
| EMR | Emerson Electric Co. | TXN | Texas Instruments |
| EXC | Exelon | UNH | UnitedHealth Group Inc. |
| F | Ford Motor | UNP | Union Pacific Corp. |
| FCX | Freeport-McMoran | UPS | United Parcel Service Inc. |
| FDX | FedEx | USB | US Bancorp |
| GD | General Dynamics | UTX | United Technologies Corp. |
| GE | General Electric Co. | VZ | Verizon Communications Inc. |
| GILD | Gilead Sciences | WAG | Walgreens |
| GS | Goldman Sachs | WFC | Wells Fargo |
| HAL | Halliburton | WMB | Williams Companies |
| HD | Home Depot | WMT | Wal-Mart |
| HON | Honeywell | XOM | Exxon Mobil Corp. |